Universidade do Minho
Escola de Engenharia

Michael de Oliveira

The Interplay between Quantum Foundations and Circuit Complexity

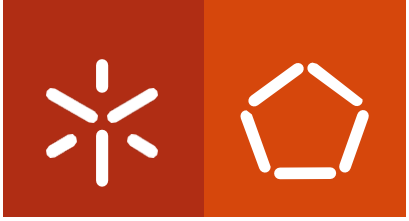
The Interplay between Quantum Foundations and Circuit Complexity
Michael de Oliveira
UMinho| 2025

The MAP-i Doctoral Programme in Informatics, of the Universities of Minho, Aveiro and Porto

universidade de aveiro
Universidade do Minho
U.PORTO

February 2025

INTERNATIONAL IBERIAN NANOTECHNOLOGY LABORATORY
SORBONNE UNIVERSITÉ
Fundação para a Ciência e a Tecnologia



Universidade do Minho
Escola de Engenharia

Michael de Oliveira

The Interplay between Quantum Foundations and Circuit Complexity

The MAP-i Doctoral Programme in Informatics, of the Universities of Minho, Aveiro and Porto



universidade de aveiro



Universidade do Minho

U.PORTO

Supervised by
Ernesto Fagundes Galvão
Elham Kashefi
Luís Soares Barbosa

February 2025

Copyright and Terms of Use for Third Party Work

This dissertation reports on academic work that can be used by third parties as long as the internationally accepted standards and good practices are respected concerning copyright and related rights.
This work can thereafter be used under the terms established in the license below.
Readers needing authorization conditions not provided for in the indicated licensing should contact the author through the RepositóriUM of the University of Minho.

License granted to users of this work:

[Caso o autor pretenda usar uma das licenças Creative Commons, deve escolher e deixar apenas um dos seguintes ícones e respetivo lettering e URL, eliminando o texto em itálico que se lhe segue. Contudo, é possível optar por outro tipo de licença, devendo, nesse caso, ser incluída a informação necessária adaptando devidamente esta minuta]



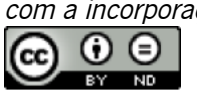
CC BY

`https://creativecommons.org/licenses/by/4.0/` *[Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original. É a licença mais flexível de todas as licenças disponíveis. É recomendada para maximizar a disseminação e uso dos materiais licenciados.]*



CC BY-SA

`https://creativecommons.org/licenses/by-sa/4.0/` *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito e que licenciem as novas criações ao abrigo de termos idênticos. Esta licença costuma ser comparada com as licenças de software livre e de código aberto «copyleft». Todos os trabalhos novos baseados no seu terão a mesma licença, portanto quaisquer trabalhos derivados também permitirão o uso comercial. Esta é a licença usada pela Wikipédia e é recomendada para materiais que seriam beneficiados com a incorporação de conteúdos da Wikipédia e de outros projetos com licenciamento semelhante.]*



CC BY-ND

`https://creativecommons.org/licenses/by-nd/4.0/` *[Esta licença permite que outras pessoas usem o seu trabalho para qualquer fim, incluindo para fins comerciais. Contudo, o trabalho, na forma adaptada, não poderá ser partilhado com outras pessoas e têm que lhe ser atribuídos os devidos créditos.]*



CC BY-NC

https://creativecommons.org/licenses/by-nc/4.0/ *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho para fins não comerciais, e embora os novos trabalhos tenham de lhe atribuir o devido crédito e não possam ser usados para fins comerciais, eles não têm de licenciar esses trabalhos derivados ao abrigo dos mesmos termos.]*



CC BY-NC-SA

https://creativecommons.org/licenses/by-nc-sa/4.0/ *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho para fins não comerciais, desde que lhe atribuam a si o devido crédito e que licenciem as novas criações ao abrigo de termos idênticos.]*



CC BY-NC-ND

https://creativecommons.org/licenses/by-nc-nd/4.0/ *[Esta é a mais restritiva das nossas seis licenças principais, só permitindo que outros façam download dos seus trabalhos e os compartilhem desde que lhe sejam atribuídos a si os devidos créditos, mas sem que possam alterá- los de nenhuma forma ou utilizá-los para fins comerciais.]*

Acknowledgements

This thesis would not have been possible without the support, guidance, and collaboration of many individuals. Their contributions—through discussions, mentorship, and encouragement—have shaped both this work and my own growth as a researcher. It is only fitting to take a moment to acknowledge them here.

I begin with my deepest gratitude to my supervisors, who played a fundamental role in this journey. First, Professor Luis, who has supported me since my master's studies—not only scientifically but in many other aspects as well. His constant encouragement and guidance have profoundly influenced my approach to research. To Professor Ernesto, I am deeply grateful for introducing me to the world of quantum foundations and for his invaluable insights throughout. And finally, Professor Elham Kashefi, whose impact on me has been transformative from our very first meeting. Her ability to bring clarity through "productive chaos" in our discussions has shaped my thinking in ways that are difficult to articulate. Beyond scientific knowledge, I have grown in ways I had not anticipated, and for that, I am truly grateful.

I now turn to those with whom I had the privilege of working—a group of exceptional researchers whose collaboration made this journey both stimulating and enjoyable. Though it would be impossible to fully capture everything they have taught me, I can only express my deepest gratitude to Alex Grilo, Andrea d'Urbano, Damian Markham, Dominik Leichtle, Johannes Frank, Min-Hsiu Hsieh, and Sathya. Our exchanges of ideas, problem-solving sessions, and shared enthusiasm made this process immensely rewarding. A special thanks goes to Sathya for carefully reading through my thesis and offering invaluable feedback that greatly enhanced it.

To the research groups that welcomed me—I could not have asked for better environments. The entire QLOC group provided me with kindness, friendship, and brought an incredible energy to my local playing ground. Then, my second group, the "Paris gang" at LIP, from the PI's to the interns, was not only an inspiring academic environment but also an incredible community that made the final stretch of this journey particularly special on the seine. And, of course, my extended family in Edinburgh—my "remote cousins"—who always welcomed me with warmth and enthusiasm. I must also mention the enriching research visits I had with Robert König's group at TUM and Hans Briegel's group at the University of Innsbruck, which were both invaluable experiences.

Beyond academia, I want to acknowledge those who have supported me regardless of my pursuits. First, my family and closest relatives, who have always been there since the very beginning. Then, my "Bonitos e Grossos" group friends and all their extensions, who never failed to bring laughter and perspective when I needed it most. And, of course, my "Mecputos" family, whose bond remains as strong as ever.

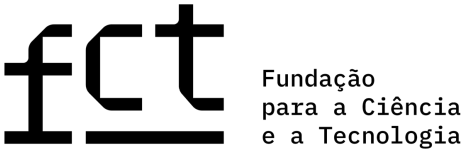
Finally, last but certainly not least—Margarida. I could not have made it to the end of this journey, or any other, with the same energy and enthusiasm without her by my side. The best company I could ever ask for, she endured endless murmurs of scattered and unorganized ideas with infinite patience. I do wonder how much quantum has seeped into her brain from all the background noise, but I can only thank her for your unwavering support. I hope that the great moments have outweighed the difficult ones, as they have for me.

Michael de Oliveira
Guimarães, 4th of February 2025

This work was supported by the International Iberian Nanotechnology Laboratory (INL) and Sorbonne University, each of which hosted the author for a period of two years. Additionally, it was funded by FCT – Fundação para a Ciência e a Tecnologia through the Ph.D. Research Scholarship 2020.05861.BD.







Statement of Integrity

I hereby declare having conducted this academic work with integrity.
I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.
I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, Braga, February 2025



Michael de Oliveira

Abstract

Quantum computers were initially proposed to surpass the limitations of classical computers in simulating physical systems. Soon after, the field of quantum computing advanced through landmark discoveries of quantum algorithms, with some conjectured to provide exponential advantages over classical counterparts. These include Shor's factoring algorithm, which has profound implications in cryptography. Additionally, moderate yet significant polynomial separations have further increased interest in the field. However, despite this progress, unconditional proofs confirming quantum superiority on the theoretical side remain elusive. At the same time, large, noise-resilient quantum computations on the experimental side, which could challenge the extended Church-Turing thesis by potentially breaking classical computational barriers, have yet to be achieved. Meanwhile, our understanding of the differences between classical and quantum mechanics has deepened. We now possess rigorous theoretical tools, such as Bell inequalities and non-local games, to describe behaviors unique to these theories. Moreover, the main loopholes for observing non-classical correlations have been closed experimentally, showing that correlations in the physical world exceed those predicted by classical mechanics.

This thesis investigates computational complexity differences between quantum and classical models by studying shallow-depth circuits. The primary focus lies in identifying, elucidating, and proving computational quantum advantages. Additionally, the analysis delves into the bijection between these circuits and foundational non-local game constructions, exploring how quantum-classical deviations in multi-party correlation settings translate to computational advantages.

A key question addressed is how to synthesize constant-depth quantum circuits for various problems to leverage quantum advantages. To this end, circuit synthesis techniques tailored to measurement-based quantum computation are developed, enabling simultaneously optimal quantum strategies for certain non-local games based on a particular bijection between this model and XOR non-local games.

Building on this work, we confront a pivotal challenge: If a candidate quantum circuit potentially provides a computational advantage, how can we prove that no classical circuit surpasses it? This work presents new unconditional quantum-classical separations within shallow-depth settings, distinguishing for the first time constant-depth quantum circuits with one- and two-qubit gates from biased polynomial threshold circuits, which incorporate neural-network-type functionalities. These separations extend to prime-dimensional quantum systems, and we conjecture to have identified Razborov-Smolensky-type problems uniquely solvable by specific prime dimensions. The thesis also differentiates constant-depth quantum circuits with multi-qubit Toffoli gates from their classical counterparts, emphasizing the power of parallel quantum computation.

Finally, verifying shallow-depth quantum computations is challenging due to conjectured exponential separations from classical computations, limiting correctness verification through simulation. While specific quantum algorithms within BQP have been addressed, we explore whether a quantum computer can reliably execute any instance from a fixed class of quantum circuits. Moving beyond heuristic-based solutions, we propose a benchmarking protocol that certifies, with exponentially high confidence, the execution of arbitrary circuits of specified dimensions, widths, and depths.

Keywords Non-local Games, Quantum Benchmarks, Quantum Circuit Complexity.

Resumo

Os computadores quânticos foram inicialmente propostos para ultrapassar as limitações dos computadores clássicos na simulação de sistemas físicos. Pouco depois, a área da computação quântica avançou com descobertas de novos algoritmos, alguns dos quais se conjectura que proporcionem vantagens exponenciais face aos clássicos. Entre eles, destaca-se o algoritmo de factorização de Shor, com profundas implicações na criptografia. Além disso, separações polinomiais moderadas, mas significativas, aumentaram ainda mais o interesse na área. No entanto, apesar deste progresso, provas incondicionais que confirmem a superioridade quântica no domínio teórico continuam a ser elusivas. Paralelamente, grandes computações quânticas resilientes ao ruído, que poderiam romper barreiras computacionais clássicas, ainda não foram alcançadas. Entretanto, a nossa compreensão das diferenças entre a mecânica clássica e quântica aprofundou-se. Dispomos agora de ferramentas teóricas rigorosas, como desigualdades de Bell e jogos não-locais, para descrever comportamentos exclusivos destas teorias. Além disso, as principais lacunas na observação de correlações não-clássicas foram fechadas experimentalmente, demonstrando que as correlações no mundo físico ultrapassam as previstas pela mecânica clássica.

Esta tese investiga as diferenças na complexidade computacional entre modelos quânticos e clássicos através do estudo de circuitos de profundidade reduzida. O foco principal é a identificação, explicação e demonstração das vantagens computacionais quânticas. Além disso, a análise explora a bijeção entre estes circuitos e construções fundamentais de jogos não-locais, examinando como desvios quântico-clássicos em cenários de correlação multipartida se traduzem em vantagens computacionais.

Uma questão-chave abordada é como sintetizar circuitos quânticos de profundidade constante para diversos problemas, a fim de explorar vantagens quânticas. Para isso, são desenvolvidas técnicas de síntese de circuitos adaptadas ao modelo de computação quântica baseado em medições, permitindo estratégias quânticas simultaneamente ótimas para certos jogos não-locais.

Com base neste trabalho, enfrentamos um desafio central: se um circuito quântico candidato oferece potencialmente uma vantagem computacional, como podemos provar que nenhum circuito clássico o supera? Este trabalho distingue, pela primeira vez, a capacidade computacional circuitos quânticos de profundidade constante com portas de um e dois qubits de circuitos dos "biased polynomial threshold circuits", que integram funcionalidades do tipo redes neuronais. Estas separações estendem-se a sistemas quânticos de dimensão prima. A tese também distingue circuitos quânticos de profundidade constante com portas Toffoli de múltiplos qubits dos seus equivalentes clássicos, enfatizando o poder da computação quântica paralela.

Por fim, a verificação de computações quânticas de profundidade reduzida é desafiadora devido às separações exponenciais conjecturadas face às computações clássicas, dificultando a validação da correção por simulação. Embora algoritmos quânticos específicos dentro de BQP já tenham sido estudados, exploramos se um computador quântico pode executar de forma fiável qualquer instância de uma classe fixa de circuitos quânticos. Para além das soluções heurísticas habituais, propomos um protocolo de benchmarking que certifica, com confiança exponencialmente elevada, a execução de circuitos arbitrários de dimensões, larguras e profundidades especificadas.

Palavras-chave Benchmarks Quânticos, Complexidade de Circuitos Quânticos, Jogos Não-locais.

Contents

I Introduction **1**

1 Context and motivation **2**
1.1 Research Questions . . . 3
1.2 Related work . . . 5
1.2.1 Synthesis of shallow-depth quantum circuits . . . 7
1.2.2 Quantum computational advantages in constant-depth . . . 7
1.2.3 Benchmarking quantum computers . . . 10

2 Contributions and thesis structure **12**
2.1 Synthesizing quantum circuits of constant-depth . . . 12
2.1.1 Bounded fan-in qubit quantum circuits . . . 12
2.1.2 Multi-qudit quantum gates . . . 15
2.2 Extending the advantage of parallel quantum computation . . . 17
2.2.1 Constant-depth qupit quantum circuits . . . 17
2.2.2 Constant depth multi-qubit Toffoli quantum circuits . . . 24
2.3 Verification-inspired quantum benchmarking . . . 26
2.4 List of publications . . . 28

II Results **29**

3 Synthesizing quantum circuits of constant-depth **30**
3.2 Introduction . . . 31
3.2.1 Our results . . . 31
3.2.2 Related work . . . 33
3.2.3 Chapter organization . . . 35
3.3 Background . . . 35
3.3.1 The non-adaptive measurement-based quantum computation model . . . 35
3.3.2 Measurement assignments . . . 36
3.4 Evaluation of Boolean functions in the $\mathrm{NMQC}_{\oplus}$ model . . . 41
3.4.1 Symmetric Boolean functions . . . 41
3.4.2 General Boolean functions . . . 48
3.5 Bounding the resources for the $\mathrm{NMQC}_{\oplus}$ model . . . 56
3.5.1 Clifford hierarchy of the measurement operators . . . 56
3.5.2 Lower bound conjecture for symmetric Boolean functions . . . 59
3.5.3 Qubit count for symmetric Boolean functions . . . 64
3.6 Circuit realization of the $\mathrm{NMQC}_{\oplus}$ model . . . 65

3.6.1 $\mathrm{NMQC}_{\oplus}$ circuits . . . 65
3.6.2 Quantum vs. classical circuits . . . 71
3.7 Discussion and conclusions . . . 75

4 Quantum circuits surpass biased thresholds circuits in constant-depth 77
4.2 Introduction . . . 78
4.2.1 Main results . . . 78
4.2.2 Technical contributions . . . 82
4.2.3 Related work . . . 83
4.2.4 Discussion and Outlook . . . 85
4.3 Overview of Proofs and Techniques . . . 86
4.3.1 Proof of Theorem 4.2.1: Separation for the qubit case . . . 88
4.3.2 Proof of Theorem 4.2.2: Separations for higher dimensions (qupits) . . . 90
4.3.3 Proof of Theorem 4.2.3: Noise-resilient separations . . . 92
4.3.4 New switching lemma for biased PTF circuits . . . 94
4.4 Preliminaries and Notation . . . 96
4.4.1 Low-depth complexity classes . . . 96
4.4.2 Random restrictions and Switching lemmas . . . 98
4.4.3 Non-local games and Correlation . . . 99
4.4.4 Noise and quantum error correction in higher dimensions . . . 100
4.5 New multi-switching lemma for biased PTF circuits . . . 103
4.5.1 Tight multi-switching lemma . . . 103
4.5.2 Depth reduction lemmas . . . 107
4.6 Quantum advantage in the noiseless case . . . 109
4.6.1 Separations for qubit cases . . . 110
4.6.2 Separation for qupit cases . . . 121
4.7 Noise-resilient quantum advantage . . . 143
4.7.1 Noise-resilient qupit Clifford circuits with quantum advice . . . 144
4.7.2 Noise-resilient quantum advantage with magic state injection . . . 160
4.8 Discussion . . . 168
4.8.1 Resource estimation . . . 169
4.8.2 Neural networks as biased threshold circuits . . . 170

5 The power of constant-depth Toffoli and qudit quantum circuits 175
5.2 Introduction . . . 176
5.2.1 Our contributions . . . 176
5.2.2 Chapter organization . . . 179
5.3 Preliminaries . . . 179
5.3.1 Classical circuit classes . . . 179
5.3.2 Quantum computation with qudits . . . 181
5.3.3 Quantum circuit classes . . . 183
5.3.4 Modular Relation Problems and Supporting Materials . . . 183
5.4 Shallow depth quantum circuits with finite-size gate sets . . . 185
5.4.1 Quantum upper bounds . . . 185
5.4.2 Classical lower bounds . . . 189
5.4.3 Separations . . . 191
5.4.4 Classical upper bounds for modular relation problems . . . 192
5.5 Shallow depth quantum circuits with infinite-size gate sets . . . 193

5.5.1 Constant-depth qudit circuits with quantum modular operators 193
5.5.2 Constant-depth qudit circuits with classical modular operators 200

6 Verifcation-inspired quantum circuit benchmarking 203
6.2 Introduction 204
6.2.1 Related work 204
6.2.2 New paradigms for quantum benchmarking 205
6.2.3 Chapter Organization 206
6.3 Preliminaries 206
6.3.1 Quantum devices 206
6.3.2 Quantum Verification Protocols 208
6.3.3 Measurement Based Quantum Computing and Universal Blind Quantum Computation 209
6.4 A formal treatment of benchmarking 211
6.4.1 Definitions 211
6.4.2 Properties of a CICC Benchmark 213
6.4.3 CICC Benchmark from Verification 214
6.4.4 Optimized Benchmarking Protocol from Robust VBQC 216
6.5 A Concrete Efficient CICC Benchmarking Protocol 219
6.5.1 Graph State based Quantum Device Characterization 220
6.6 Discussion 220

7 Conclusions and outlook 223

III Appendices 255

A Mathematical transformations 256
A.1 Walsh-Hadamard transform 256
A.2 Krawtchouk transform 257

B Concentration Inequalities 258
B.1 Hoeffding's Inequality 258

List of Figures

2 Representation of a k-ReLU gate within bPTF$^0(k)$. This gate is equivalent to a ReLU gate, defined as $f(x) = \max\{0, x - c\}$ (where the center is shifted from 0 to c), up to an integer precision w, for any input string with a Hamming weight bounded by k. Our scheme considers ReLU : $\{0,1\}^n \mapsto \{0, n - c\}$, which takes n-bit strings as input and interprets their Hamming weight as the input x. 18

3 Relationships between the key classical and quantum circuit (complexity) classes considered. Notably, Majority is in TC^0 but not AC^0. The ISMR family of problems, introduced here, separates constant-depth quantum circuits from biased polynomial threshold circuits (bias $k = \omega(\log n)$). The latter class contains AC^0, and can solve NC^1-complete problems for super-logarithmic values of the bias and input lengths, suggesting non-trivial overlap with TC^0. 19

4 Representation of the Modular XOR games $\mathcal{G}_p$ involving n parties. The plot illustrates the upper bounds on the (winning) correlation of classical strategies for $\mathcal{G}_2$, $\mathcal{G}_3$, $\mathcal{G}_5$, and $\mathcal{G}_7$, whereas the optimal quantum strategy attains a maximum correlation of 1 for all $\mathcal{G}_p$. 20

5 Once a quantum device is certified for various graph states, one can ask which computations, i.e., unitaries, can be certifiably run on the device. Unitaries can be compiled onto a graph state for the equivalent MBQC pattern. However, the same unitary can be mapped onto different graph states. Some of these might be certified on the device; others might not. Therefore, one can attempt to rewrite the computation to find a graph state that is certified on the device so it can implement the computation. 28

6 Illustration of the circuit classes for which we prove a separation in Theorem 3.2.2. For the classical class, the input can be copied only in the copy section, which precedes the logical operator section. For the quantum circuits, the same occurs. This prevents the circuits from employing the copy operation at any point in time, which would replicate the effect of memory in the circuits. 32

7 Circuit for evaluating the symmetric Boolean functions with degree two built from the FPHP relation by the composition of a classical circuit that computes the parity. 34

8 (a) The complete NMQC$_\oplus$ procedure. The input string $x \in \{0,1\}^n$ is shared with each linear function of the pre-processing stage. Then the binary values s_i are forwarded to the measurement process, determining a specific measurement operator on all qubits. In the end, the measurement results are fed into a linear function to compute the result of the NMQC$_\oplus$ computation. (b) Representation of a digital circuit composed of NAND gates that are implemented using individual NMQC$_\oplus$ computations. 36

9 Representation of the relations between the individual CSF (ordered by degree) and the asymptotic growth rate of the GHZ state necessary for deterministic evaluation with the $\mathsf{NMQC}_{\oplus}$ model. The continuous black line represents the most compact polynomials identified in [1]. The dashed dark blue line represents the most compact polynomials constructed using the decomposition from Proposition 1 and the multi-linear polynomials from Lemma 2. Additionally, the red triangles are the CSF computed exactly with the polynomials from Lemma 2. Finally, the dotted grey line represents our conjectured lower bound for the set of CSF (Conjecture 2). 66

10 a) The general form and dimensions of a circuit computing a linear function with $|S|$ inputs. b) Parallel composition of the linear functions in the linear pre-processing routine. 67

11 a) Logarithmic depth circuit for the creation of an 8-qubit GHZ state with a recursive CNOT application pattern. b) Constant quantum depth circuit with extra linear measurement post-processing the creation of an 8-qubit GHZ state. 68

12 Exact and approximated circuit synthesis of the measurement operator. The former is based on arbitrary R_Z rotations, the latter on the Clifford+ T gate set. 68

13 A schematic representation of the circuit realization proposed for the $\mathsf{NMQC}_{\oplus}$ model. Vertically aligned processes are performed in parallel, while horizontally aligned processes are performed sequentially. 71

14 This illustration depicts the primary classical and quantum circuit classes analyzed in this paper, highlighting the inclusion of relevant problems that characterize the relationships and distinctions between these classes. Notably, the Parity function is computable in $\mathsf{AC}^0[2]$ but not in AC^0, while the Majority function is not computable in $\mathsf{AC}^0[p]$ but is in TC^0. We also indicate the ISMR family of problems introduced in this work (see Definition 62), which differentiate QNC^0 from $\mathsf{bPTF}^0[k]$ with $k = \omega(\log n)$. The latter class, which strictly includes AC^0, can solve instances of NC^1-complete problems for super-logarithmic size input strings. It can also solve computational problems conjectured to be strictly beyond NC^1. This suggests a potential non-trivial overlap with TC^0. We also list more general and practically relevant computational processes known to be computable within each class, particularly neural network-type circuits. Transformers, known for their substantial use as crucial components of modern large language models (LLMs), can be simulated by AC^0 circuits when the attention mechanism in the transformer is restricted to hard attention functions [2, 3]. Furthermore, TC^0, recognized as the standard class for theoretically representing discretized neural networks [4, 5], can simulate LLMs under realistic operational conditions [6], such as restricting the precision of internal variables and auto-regression or the number of time-steps the model is run for at inference. This highlights a key distinction between AC^0 and TC^0, in that the latter class has enhanced expressivity for a wide range of algebraic operations, such as integer multiplications, division, and sorting [7, 8]. The $\mathsf{bPTF}^0[k]$ circuits that we study in this work are capable of simulating a variety of variants of standard neural networks. This class of circuits also supports attention functions that are likely to be beyond TC^0 and may accommodate substantially different LLM architectures. We refer the reader to Section 4.8.2 for a more detailed discussion on these points. 87

15 Illustration of the respective contributions to the correlations function based on the deviation between two arbitrary function f and g over $\mathbb{F}_3$ and $\mathbb{F}_5$. 100

16 Illustration of A_v and B_p type surface code stabilizers on a 3×3 lattice. 103

17 Representation of the logical operators $\overline{\mathsf{X}}$ and $\overline{\mathsf{Z}}$ for the surface code. 103

18 Illustration of the MBQC-based constant-depth quantum circuit solving an instance of the inverted strict modular relation problem. Stages 1 and 2 involve creating the graph states. In Stage $3a$, the edge qubits of the graph are measured, producing the poor-man's cat state and fanning out the measurement outcomes of the edge measurements. In Stage $3b$, phase operations are applied based on the input, and the entire state is measured. Stage 4 involves the classical post-processing. 118
19 First horizontal entangling routine 118
20 Second vertical entangling routine 118
22 Illustration of a constant-depth quantum circuit for generating a generalized poor man's cat state, accompanied by the classical information needed to describe the specific instance created. 134
23 Illustration of a constant-depth quantum circuit solving all instances of the ISMR problems $\mathcal{R}_p^m$ with parameter p prime. Note that the last wire does carry the information about the edge measurements in the GPM state creation process $E = \{e_{i,j}, \ldots, e_{l,k}\}$. . . . 135
24 Constant-depth non-adaptive Clifford quantum circuit with quantum advice states. . . . 145
25 On the top left, we have a representation of the distribution of the stabilizer generators over layers (planes) parallel to the u_1 and u_2 axes, corresponding to even and odd values of u_3 in the lattice $\mathcal{C}$. The layers termed "even*" and "odd*" are equivalent, having the same pattern defined by the distribution of stabilizers, to the "even" and "odd" layers, respectively, except for the translation by two units in the same plane formed by u_1 and u_2. The cross (×) at the origin of the 2D graphs indicates that the u_3 axis is inward-directed. The top right panel displays a representation of the stabilizer generators K_u from Equation (4.7.43) with their respective spatial orientation. At the bottom, we have a depiction of the entire lattice $\mathcal{C}$, indicating the layers and defining stabilizer positions after the application of the constant-depth circuit $\mathcal{Z}$ to the lattice. Note that the first and last layers constitute the surface (Sur), where the logical qupits reside, while the remaining middle layers define the bulk (Bulk). 155
26 noise-resilient constant-depth Clifford quantum circuit with quantum advice states. . . . 166
28 Graph representation of $\mathsf{AC}^0[p]$ classes with p prime. Each edge denotes that the connected classes are distinct, without one being a subset of the other. Along the edges, examples highlight problems that are in one class but excluded from the other. 180
29 Parameterized quantum circuit class for values q, p, and n, incorporating the advice quantum state $\left|\mathsf{GHZ}^0_{q,n}\right\rangle$, which solves the $\mathcal{R}^{q\cdot n}_{q,p}$ with bounded one-sided error. 187
30 Representation of $\mathcal{B}_6$. 188
31 Representation of a classical $\mathsf{AC}^0[q]$ circuit, parameterized by q and n, that reduces a solution to the $\mathcal{R}^{q\cdot n}_{q,2}$ problem to a solution of $\mathsf{Mod}_2(x)$, where x is the initial input string of size n. 190
32 Quantum verification and extraction as interactive algorithms. 208
33 Benchmarking Algorithm **B** interacting with a quantum device **Q** m times outputting information $z \in [0, 1]$ that is based on **Q**. 212
34 Verification-inspired CICC benchmarking algorithms $\mathbf{B_V}$ utilizing m rounds of the verification algorithm **V** 215
35 After a quantum device is certified for different graph states, one can ask which computations, i.e., unitaries, can be certifiably run on the device. Unitaries can be compiled onto a graph state for the equivalent MBQC pattern. However, one unitary can be mapped onto different graph states. Some of these might be certified on the device, others might not. Therefore, one can attempt to rewrite the computation to find a graph state that is certified on the device so it can implement the computation. 221

List of Tables

1 Summary of related work on unconditional separations of quantum and classical shallow-depth circuits. Please refer to the main text of Section 4.2.3 for problem abbreviations. . 9

2 Estimates of the input sizes n required to demonstrate quantum advantage using the previously identified constant-depth quantum circuits, with all-to-all connectivity, in both noise-free and noise-affected settings, based on the quantum upper and classical lower bounds determined. We also examine the minimal fan-in (F) scenario by comparing quantum circuits with classical circuits of equal locality—keeping the same depth and fan-in at each level as the quantum circuits. Additionally, we analyze exact-case hardness bounds for the qubit setting to establish lower resource estimates. For the remaining qudit dimensions, we rely on average-case hardness to derive comparable estimates. . 23

3 Size lower bounds for the circuit classes $\mathsf{bPTF}^0[k]$. By "exact hardness" we mean that the classical circuit is required to produce a valid output string with certainty on all inputs, and in this setting the expressions bounds the size s of the respective $\mathsf{bPTF}^0[k]$ circuits. 80

4 Summary of related work on unconditional separations of quantum and classical shallow-depth circuits. Please refer to the main text of Section 4.2.3 for problem abbreviations. . 85

6 Size lower bounds for the circuit classes $\mathsf{bPTF}^0[k]$ without simplifications and optimal values for the parameter q. 170

Part I
Introduction

Chapter 1
Context and motivation

Landmark quantum algorithms such as Shor's factoring algorithm—which is renowned for its potential to compromise widely-used cryptographic protocols—demand extensive quantum resources [9, 10, 11] in what concerns number of qubits, coherence times and error correction. They also necessitate a substantial number of elementary operations that must be executed within a fault-tolerant framework, a prerequisite common to many complex information-processing tasks. This technological constraint has sparked a growing interest in the exploration of shallow-depth quantum circuits, which aim to harness quantum computational advantages within limited operation depths or run times. This approach seeks to maximize the computational power with coherence times achievable on contemporary quantum hardware, thereby enabling the demonstration of quantum advantages on today's noisy intermediate-scale quantum (NISQ) devices [12].

Building on this shift towards less resource-demanding quantum computations, a series of pioneering studies have demonstrated the capability of shallow-depth quantum circuits to perform computations that can be proven to require exponential run time or circuit size on classical computers, conditioned on certain computational assumptions. Key examples include IQP circuits [13], Boson sampling experiments [14], and random circuit sampling [15], all of which require fewer resources than previous quantum candidates for demonstrating computational advantage. Despite these circuits not being directly tied to standard computational problems, their practicality has been the subject of much investigation [16, 17]. They are posited as potential challenges to the extended Church-Turing thesis, which hypothesizes that physically realizable computation based on quantum mechanics could be fundamentally more powerful than our standard *classical* computation models. While these findings mark significant progress towards establishing quantum computational superiority, they are not devoid of limitations, similar to those faced by more extensive quantum algorithms. First and foremost, there is no definitive proof that an unknown classical algorithm could not efficiently solve the same problems, as these claims rest on computational assumptions. Although these assumptions are generally regarded as robust, they remain unproven conjectures. Furthermore, many of these theoretical examples of quantum advantage suffer from difficulties in verifying their correct execution and from not adequately accounting for the noise present in current experimental setups. This highlights a gap between theoretical models and practical implementation.

This has motivated a line of research focused on shallow-depth quantum circuits where direct comparisons are made with their classical shallow-depth counterparts. The goal of this approach is to establish unconditional separations between complexity classes, which, even in classical computational complexity, are typically only known in the constant-depth regime. The seminal work of Bravyi et al. [18] generated significant attention in this line of work by demonstrating that classical circuits confined to the NC^0 class—characterized by constant depth and bounded fan-in gates—cannot perform certain tasks that constant-depth quantum Clifford circuits with constant fan-in can. This work connected shallow-depth quantum circuits with classical circuits of constant or logarithmic depth, an area that has been extensively studied since the 1980s [19, 20, 21, 22], and which remains an active topic of research [23, 24, 25, 26]. Importantly, these comparisons help to clearly distinguish quantum from classical computing, offering foundational insights into quantum computation, particularly in circuits capable of parallel processing,

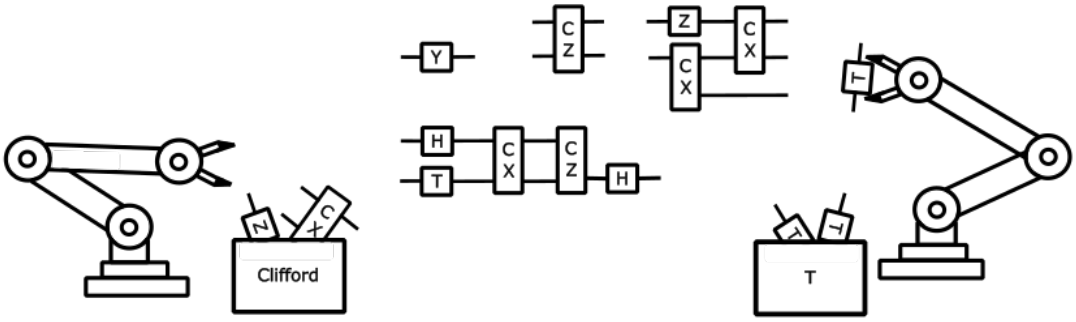
which could pave the way for practically relevant quantum algorithms in the near future.

1.1 Research Questions

By considering circuits as our computational objects—an approach that encompasses a wide range of computational models in computer science and is regarded by many as a potential framework for addressing important complexity questions, such as the P vs. NP problem—we benefit from their well-defined notions of space and time complexity. Our challenge was to examine the quantum advantage, primarily in terms of circuit size comparisons unless otherwise specified, of constant-depth quantum circuits over their classical counterparts. This task was divided into three primary components: identifying potential candidate problems for which quantum circuits achieve an advantage over classical circuits, formally proving this advantage, and certifying the accuracy of empirical attempts to validate quantum advantage. These aspects were addressed through the following key research questions.

Starting from a set of computational problems—informed by previously demonstrated quantum advantages in [18] and [27]—in which quantum devices may unconditionally outperform classical ones in the shallow-depth regime[1], the first step is to identify promising quantum candidates capable of effectively solving these problems.

Problem 1. *How can we synthesize constant-depth quantum circuits for a broad class of problems to harness potential quantum advantages?*



More specifically, circuit synthesis involves constructing a quantum circuit from available elementary gates that either performs a desired unitary transformation or, in the case of state generation, produces a target state from an initial fiducial quantum state [28, 29, 30]. This task is primarily driven by the goal of achieving the greatest possible quantum advantage while being aware of the limitations of current quantum devices. Its importance is highlighted in scenarios where quantum algorithms offer only polynomial advantages, requiring optimal or near-optimal quantum circuit implementations to fully realize their potential [31]. Consequently, achieving quantum advantage requires near-optimal quantum circuits to compare against optimal classical solutions, highlighting the critical role of circuit synthesis.

In the classical domain, circuit synthesis is quite a mature field due to extensive research over many decades. However, in the quantum domain, the problem is inherently more complex, and our understanding remains limited. This complexity arises in part because the search space for quantum circuits is significantly larger than that for classical circuits, making the development of efficient synthesis techniques even more crucial [29]. Therefore, while the quality of the final solution is important, the efficiency of the synthesis process itself is equally critical. Additionally, in practice, many of these techniques need to be fine-tuned to optimize their performance based on the specific hardware architecture and the quality of the quantum devices being used [32]. This is especially relevant for identifying quantum advantage in near-term devices, such as by limiting the depth of the synthesized quantum circuits.

[1] Throughout the text, we will refer to shallow-depth circuits as those of constant and at most logarithmic $\mathcal{O}(\log(n))$ depth.

After synthesizing compact quantum circuits for our large candidate set of problems and obtaining upper bounds on the quantum circuit sizes, the next step is to understand if these circuits are indeed more efficient—meaning they are less resource-intensive than any possible classical solution. In terms of computational complexity, and specifically circuit complexity, this involves identifying separations in power between classical and quantum circuit classes.

Problem 2. *What is the largest class of classical circuits that can be unconditionally separated from constant-depth quantum circuits?*

Notably, we aim to identify separations that can be proven without relying on computational assumptions. The ultimate goal is to identify fundamental differences in power between quantum and classical computational models. Groundbreaking progress in this area was made by Bravyi et al., whose seminal work laid the foundation for exploring such separations. Since then, these results have been extended in various directions, including comparisons with increasingly larger classical circuit classes [27], results on average-case hardness [33, 27, 34], incorporating noise effects in quantum circuits [35, 36, 37], and exploring sampling problems [38].

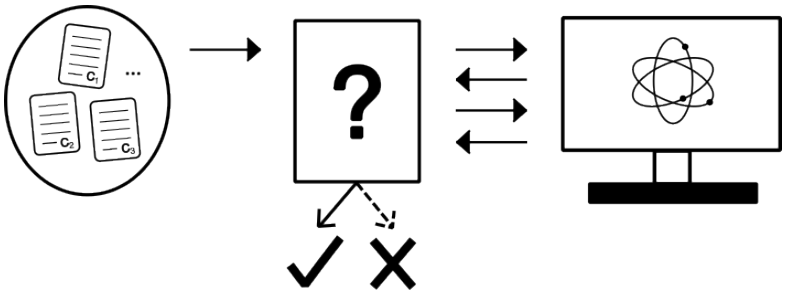
This line of research into the computational power of shallow-depth quantum circuits is especially significant, as these circuits are not only more practical—offering easier implementation on near-term quantum devices—but they also provide a powerful theoretical framework for exploring quantum advantages over classical systems. Moreover, they serve as an intriguing bridge between NISQ and early fault-tolerant quantum computers, leading toward large-scale fault-tolerant quantum systems.

After identifying computational problems for which quantum circuits can be constructed to solve them and formally proving that these circuits achieve advantages over classical ones, one might assume the final step is merely running these circuits on quantum hardware, announcing the newfound quantum advantages, and, if applicable, leveraging the results for practical use. However, there is a critical aspect of quantum advantage experiments often referred to as quantum verification or certification. For context, while the output of some quantum algorithms, like Shor's algorithm, are easy to verify (e.g., by multiplying the provided prime factors), others—such as those involving sampling problems—are significantly harder to verify. In these latter cases, the correctness of the outcome can be significantly compromised by noise, which might make the results easier to generate, thus undermining any possible quantum advantage. On the other hand, our general expectation is that classical devices may struggle to simulate these outcomes due to the inherent complexity or inefficiency of the simulation techniques, leaving us unable to assess the accuracy of the quantum implementation confidently.

One solution to this complex problem has emerged in the form of quantum verification protocols which can verify any quantum computation within BQP [39]. However, many important challenges persist. For instance, verifying quantum computations might require additional resources beyond the computations themselves [40] and comes with additional technical challenges for the current hardware providers [41]. Moreover, current verification processes work on a case-by-case basis, meaning we can only certify the correctness of a single quantum circuit through these tailor made protocols.

This raises an intriguing question: if we have identified a class of quantum circuits that demonstrate quantum advantage, could we certify a quantum device as capable of executing all circuits within this class, or at least certify it for running correctly multiple circuits of a particular form, rather than verifying each one individually? This consideration leads to the following problem.

Problem 3. *Can we certify the capability of a quantum computer to execute any instance from a fixed class of quantum circuits?*



This question falls within the realm of quantum benchmarks, as it seeks to formally characterize quantum devices based on their correctness across certain classes of quantum computations. In the current noisy regime of quantum devices, quantum benchmarks typically try to identify the computations that a device can reliably execute, while performance parameters such as time and space efficiency—primarily the focus in classical computations—are part of the benchmarking landscape but are of lesser relevance. This addresses a current gap, as most quantum benchmarks are heuristic in nature, unable to provide formal guarantees about the devices they characterize, and face challenges related to scalability and predictability.

1.2 Related work

In this subsection, we will briefly review the existing literature relevant to each of the research questions proposed. For this discussion, as well as for the remainder of this introductory chapter, we assume the reader has a general understanding of quantum computation (for a refresher on fundamental concepts, see [42]) and basic familiarity with computational complexity (for a refresher, see [43]). Additionally, for a deeper understanding of specific subtopics, such as Bell non-locality and non-local games, please consult [44]. For the classical and quantum shallow depth circuit classes, we will provide an informal description as these will be examined in greater detail and compared with each other in our results. For a formal perspective, please refer to [45] and [46, 47, 48].

In classical circuit complexity, the circuit classes characterizing classical parallel computations[2] are primarily described as classes of classical circuits with polynomial size and polylogarithmic depth. In particular, some of the main classes of study are the NC^i circuit classes, which are defined by circuits of depth $\log^i(n)$ and are composed of gates with bounded fan-in and unbounded fan-out. These classes form a natural hierarchy, where NC^i is contained within NC^{i+1}, capturing increasingly powerful parallel computations as i increases. For example, 2-SAT and the graph reachability problem reside in NC^2. However, it remains an open question whether this hierarchy is strict—meaning that $\mathsf{NC}^i \subsetneq \mathsf{NC}^{i+1}$ for all i, as generally conjectured—or if there exists an index i at which the hierarchy collapses, analogous to phenomena observed in arithmetic circuits [49].

Additionally, several important circuit classes have been defined between NC^0 and NC^1. For instance, AC^0 [3] which is defined as the class of constant-depth circuits with AND and OR gates of unbounded fan-in and fan-out, supports simple operations such as integer addition and subtraction. Notably, AC^0 strictly contains NC^0, which cannot compute AND or OR over the entire input. However, AC^0 is unable to evaluate modular operations over the Hamming weight of the full input, which motivates the definition of the $\mathsf{AC}^0[p]$[4] class. This extended class enhances AC^0 by incorporating modular gates (Mod_p) of unbounded fan-in and fan-out.

[2] We refer to computational models as parallel if they have at most polylogarithmic runtime.

[3] AC^0 has also been highly influential in theoretical computer science: for expample, tight lower bounds for fixed-depth AC^0 circuits have led to oracle constructions that separate different levels of the polynomial hierarchy [50].

[4] We refer to any complexity class C augmented with access to Boolean modular counting functions Mod_p, which compute whether the Hamming weight of the input string is congruent to 0 modulo p, as $\mathsf{C}[p]$.

Of growing practical interest, TC^0 is defined as the class of constant-depth circuits with unbounded fan-in and fan-out threshold gates, and capture the computational power of neural networks. On the one hand, TC^0 strictly contains $\mathsf{AC}^0[p]$, for every prime p, which cannot compute the majority function, while TC^0 can do so trivially. On the other hand, TC^0 is a subset of NC^1. Its strict containment, however, is only conjectured: as it is widely believed that iterated matrix multiplication does not lie in TC^0, even though it belongs to NC^1. Despite extensive efforts, proving the strict containment $\mathsf{TC}^0 \subsetneq \mathsf{NC}^1$ remains an open problem, underscoring its significance as a frontier for current lower-bound techniques. In summary, the known relationships between these classes are as follows (see Figure 1a for a visual representation),

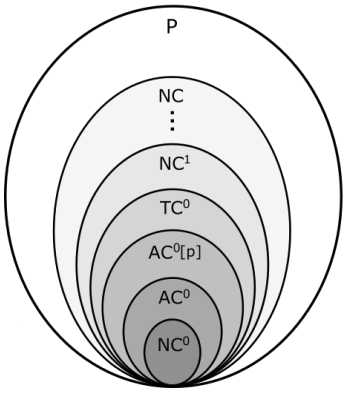
$$\mathsf{NC}^0 \subsetneq \mathsf{AC}^0 \subsetneq \mathsf{AC}^0[p] \subsetneq \mathsf{TC}^0 \subseteq \mathsf{NC}^1 \subseteq \mathsf{NC}^2 \ldots \subseteq \mathsf{NC} = \cup_i \mathsf{NC}^i \subseteq \ \ldots \ \subseteq \mathsf{P}. \tag{1.2.1}$$

As one might expect, the quantum parallel circuit classes, analogous to NC, are defined similarly by quantum circuits of polylogarithmic depth composed of gates of bounded fan-in and fan-out, referred to as QNC^i [46]. Notably, it is established that for decision problems QNC^0 is a strict subset of QNC^1, as QNC^0, akin to NC^0, is incapable of computing simple functions such as parity. However, QNC^1 encompasses several significant quantum algorithmic primitives, including the quantum Fourier transform, which serves as a fundamental subroutine in many quantum algorithms that are conjectured to archieve exponential quantum advantage. The higher levels of this hierarchy remain poorly understood, as only a limited number of problems have been identified within each level [51].

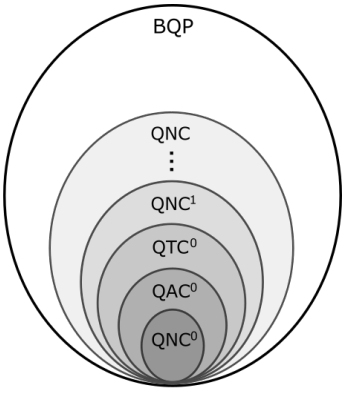
Furthermore, as in the classical case, the quantum analog classes of AC^0 and TC^0 have been defined and designated as QAC^0 and QTC^0. QAC^0 includes multi-qubit Toffoli gates [47], while QTC^0 further allows for multi-controlled quantum threshold gates [48]. In contrast to the classical case, these quantum circuit classes are poorly understood, as quantum lower bound techniques are significantly less developed than classical methods. For instance, it has been conjectured that QAC^0 cannot evaluate the parity function either [52]. However, no proof technique attempted over the ast 2 decades has been able to establish this result without imposing auxiliary constraints or structural assumptions on QAC^0 circuits. Even more surprisingly, until very recently [48], it was not known whether QTC^0 could evaluate the parity function either, a trivial task for the equivalent classical circuit class. The known relationships between these classes are as follows (see Figure 1b for a visual representation),

$$\mathsf{QNC}^0 \subsetneq \mathsf{QAC}^0 \subseteq \mathsf{QTC}^0 \subseteq \mathsf{QNC}^1 \subseteq \mathsf{QNC}^2 \subseteq \ \ldots \ \subseteq \mathsf{QNC} = \cup_i \mathsf{QNC}^i \subseteq \ \ldots \ \subseteq \mathsf{BQP}. \tag{1.2.2}$$

An essential consideration for quantum shallow-depth circuit classes is the inclusion or absence of special fanout gates, which must be implemented differently due to the no-cloning theorem. The aforementioned classes do not allow quantum fanout gates; as we will discuss later, its inclusion is conjectured to significantly extend their computational power [53].



(a) Illustration of the conjectured relationships among **classical** parallel circuit complexity classes.



(b) Illustration of the conjectured relationships among **quantum** parallel circuit complexity classes.

1.2.1 Synthesis of shallow-depth quantum circuits

The field of quantum circuit synthesis employs a wide range of approaches driven by the intrinsic computational complexity of the problem. This challenge has inspired the use of diverse techniques and heuristics, many adapted from classical computational methods, to achieve solutions that are practical and effective, even if not optimal. Beyond the diversity of techniques, the problem can be represented in multiple forms. While direct circuit representations are intuitive, other formulations, such as measurement patterns [54], directed acyclic graphs [55], phase polynomials [56], ZX diagrams [57], and tensor networks [58], have also gained traction. These varied representations allow for optimizations to be performed using different methodologies, each offering unique advantages. Additionally, machine learning techniques have increasingly been applied, showcasing a remarkable ability to generate high-quality solutions to computationally intensive problems, as seen also in other domains of computer science [30]. Together, these approaches highlight the field's interdisciplinary and evolving nature; for a broader overview, see [29].

Another key aspect of this problem is that it can be optimized based on different characteristics of the target circuit. Specifically, in line with our research question and the goal of creating hardware-friendly quantum circuits, considerable research has been devoted to developing depth-optimal quantum circuits. Efforts in this area have focused on reducing the depth of both general quantum circuits [59] and specific subclasses, such as Clifford circuits [60]. Additionally, computational problems that are typically resistant to parallelization have been studied to identify special subclasses that permit it. For instance, this approach has been applied to quantum simulation problems by analyzing specific Hamiltonians that allow for some degree of parallelization [61]. Furthermore, it is possible to map certain simulation problems onto the circuit architectures of parallel quantum models, such as those of boson samplers, enabling them to be solved with relatively small circuit depths [62].

Another important consideration is the effective co-design of quantum circuits alongside other aspects of quantum execution. For instance, a broader approach involves synthesizing circuits optimized not only for specific characteristics of the logical circuit but also for their fault-tolerant implementation using particular error correction codes [63, 56]. Notably, the most compact logical circuit may not always yield the most accurate results in a fault-tolerant setting due to the encoding constraints imposed by the error correction code. This co-design methodology can also be extended to the hardware level, incorporating factors such as connectivity, elementary operations, and noise models into the synthesis process [32]. For example, in the context of our research questions, parallel quantum models such as boson sampling circuits [14] can be optimized to reduce the number of required qubits by utilizing devices with all-to-all connectivity [64]. Similarly, the synthesis of IQP circuits has been explored across different hardware platforms, including configurations that support efficient MBQC computations [65] and extensions to continuous variable systems [66].

Finally, we emphasize that while the literature on quantum circuit synthesis and compilation is extensive, relatively little work focuses on synthesizing circuits for problems with provable, unconditional quantum advantage. This is partly due to our limited understanding of the types of problems that can enable quantum advantage, as well as the reliance of existing approaches on specific proof techniques to demonstrate computational advantage.

1.2.2 Quantum computational advantages in constant-depth

The computational advantage of quantum computers over classical ones has been theoretically studied across various computational problems, employing techniques such as query complexity and relativization, and identifying quantum advantages of varying magnitudes—ranging from polynomial to exponential—for different problems, with examples including Grover's and Shor's algorithms. However, most of these quantum algorithms, despite their significant advantages, remain far beyond the reach of near-term quantum devices. This limitation has shifted the focus toward shallow-depth quantum circuits, which can be imple-

mented on near-term hardware or, at the very least, require fewer resources, making them more feasible for practical realization. Moreover, significant quantum-classical separations with lower resource requirements have been demonstrated under reasonable computational assumptions for models such as IQP, boson sampling, and random circuit sampling, leading to notable demonstrations of quantum advantages on current hardware [67, 68]. Nevertheless, eliminating these assumptions—well-founded as they may be—to achieve unconditional separations remains a critical goal of ongoing research.

Towards this goal, starting with the simplest setting, we consider constant-depth quantum circuits with bounded fan-in gates. The first unconditional separation between constant-depth classical and quantum circuits, established by Bravyi et al. [18], has inspired various extensions over the last five years. These results are closely related, sharing similar quantum solutions based on non-local games while employing different techniques to prove classical hardness. Here, we provide a brief overview of the most relevant results, summarized concisely in Table 1.

Bounded fan-in and worst-case. We begin by noting that the result of [18], initiating this subfield, was achieved by using a non-local game previously described by [69]. However, unlike the usual definition of non-local games, this extended GHZ game allowed for some communication between the parties. This additional communication was utilized to construct a quantum circuit whose outputs implement the winning strategies for this game, while also linking the circuit's output to the solution of the 2D Hidden Linear Function (2D HLF) problem. The latter is a non-oracular version of the Bernstein-Vazirani problem [70], with the difference that the linear function $f(x) = \mathbf{a} \cdot \mathbf{x}$ to be learnt is not accessed directly via a quantum oracle, but is 'hidden' by a quadratic form. The authors use lightcone arguments that exploit bounded fan-in to show that, even with access to a random advice string, the efficacy of classical NC^0-type circuits in solving the 2D HLF problem is limited by the efficacy of classical strategies in winning the underlying non-local game. Thus, it is guaranteed that the worst-case success probability of such constant-depth classical circuits is unconditionally inferior to that of the constant-depth quantum circuit solving the problem.

Average-case and unbounded fan-in. Subsequent work extended this unconditional separation in two major directions: to average-case hardness or correlation bounds [33, 27], and to the class AC^0 of constant-depth circuits composed of AND and OR gates of unbounded fan-in [27]. The latter work achieved this through a task dubbed the Parity Halving Problem (PHP), which requires producing an output string with a Hamming weight whose parity matches the result of applying the Boolean Mod 4 operation to the input string. The authors applied a different lower bound technique to the PHP, namely (multi-)switching lemmas, to relate AC^0 circuits back to NC^0 circuits, which allowed locality arguments to be ported to the former. Switching lemmas describe how the function or relation computed by an AC^0 circuit simplifies when some of its input literals are randomly set to 0 or 1. The main result of [27] was that classical solutions to the PHP require subexponential-size AC^0 circuits. For the quantum upper bound, the authors utilized poor-man's cat states, which are local unitary (LU)-equivalent to n-party GHZ states, to efficiently solve the PHP with certainty on all inputs with polynomial-size quantum circuits. Furthermore, they showed an NC^0 reduction from the PHP to 2D HLF, proving that the new average-case separation also applies to the HLF problem.

Noise resilience. Traditional quantum fault-tolerant schemes have a significant depth overhead and might not be compatible with early quantum advantage proposals. Driven by this understanding, [35] showed that for the popular local stochastic noise model, errors occurring in any given layer of the circuit could be corrected with constant-depth circuits. They showed that transversal Clifford gates could be realized by a constant-depth circuit equipped with single-shot state preparation and readout, thereby proving that noise resilience could be achieved within a constant depth. For the Magic Square Problem (MSP) (inspired by the Mermin-Peres magic square [71, 72]) they exhibited a QNC^0 circuit in a 1D config-

uration of connectivity that has a quantum advantage against NC^0, and consequently a 3D configuration error-correcting code using the previous elements that offers robustness against noise while retaining the advantage against NC^0. Moreover, [37] recently proved the unconditional quantum advantage of noisy QNC^0 circuits over AC^0. Focusing on the task of possibilistic simulation of single-qubit Clifford gate teleportation (TELEP)—a task previously shown to be challenging for NC^0 [73]—the authors extended the previous separation to AC^0, while introducing a 1D-implementable quantum circuit, enabling a 3D fault-tolerant construction as shown in [35].

Finally, alternative approaches have considered different types of tasks and separations, including interactive protocols [74, 36] and sampling problems [38]. Some studies have also applied geometric restrictions to classical circuits [75] to obtain separations in the depth.

Problem	Advantage against	Geometry	Higher dimensions	Avg.-case hardness	Noise resilience
2D HLF [18]	$\mathsf{NC}^0/\mathrm{rpoly}$	2D	✗	✗	✗
PHP [27]	$\mathsf{AC}^0/\mathrm{rpoly}$	2D	✗	✓	✗
MSP [35]	$\mathsf{NC}^0/\mathrm{rpoly}$	1D	✗	✓	✓
TELEP [37]	$\mathsf{AC}^0/\mathrm{rpoly}$	1D	✗	✓	✓

Table 1: Summary of related work on unconditional separations of quantum and classical shallow-depth circuits. Please refer to the main text of Section 4.2.3 for problem abbreviations.

As a closing remark, we note that all these separations between quantum and classical shallow-depth circuit classes are unconditional, demonstrating exponential differences in success probabilities or circuit sizes. However, these gaps arise specifically in the context of the identified problems when the same constant-depth constraints are applied to their classical counterparts. Notably, these problems can still be solved classically with a polynomial overhead in computational resources, provided there are no depth restrictions. This underscores that the notion of separation here is highly context-dependent.

Multi-qubit quantum gates. Prior to the recent developments regarding unconditional quantum-classical separations concerning QNC^0, the unconditional superiority of constant-depth quantum circuits over their classical counterparts had already been demonstrated in specific cases involving circuit classes that utilize multi-qubit quantum gates. More specifically, QAC^0—the quantum analog of AC^0, which includes multi-qubit Toffoli gates (acting as quantum versions of AND gates) combined with quantum fanout gates, or equivalently $\mathsf{QAC}^0[2]$—has been shown to be able to compute Mod_q gates for arbitrary prime q [76], a capability that none of the classical $\mathsf{AC}^0[p]$ classes has. Consequently, unconditional separations between a quantum circuit class and its classical analog can be established by leveraging modular decision problems to distinguish them.

Later, the consideration of a single multi-qubit gate, the fanout gate, was shown to be sufficient to collapse the entire quantum constant-depth hierarchy, demonstrating that $\mathsf{QNC}^0[2] = \mathsf{QTC}^0$ [77]. For comparison, classical $\mathsf{NC}^0[2]$ is a much weaker circuit class than AC^0 and TC^0. The collapse of the same classes for their respective randomized circuit classes, where the circuits are requested uniquely to solve the problems with a constant positive success probability and not exactly, had been shown previously in [53], along with the additional demonstration that some special algorithmic procedures, such as the quantum Fourier transform, can be approximated in constant depth. The latter even provides strong evidence for a significant computational advantage in constant-depth with the use of multi-qubit quantum gates. Furthermore, the same quantum circuit class, $\mathsf{QNC}^0[2]$, has also been used to demonstrate an unconditional computational advantage over its classical analog, $\mathsf{NC}^0[2]$, for a more practical problem known as noisy decoding [78].

These circuits rely on more complex components, such as multi-qubit Toffoli and quantum fanout gates, compared to the simpler constant-depth quantum circuits that use only single- and two-qubit gate sets. However, recent advances and practical demonstrations indicate that certain quantum device platforms inherently support multi-qubit operations, including fanout gates [79, 80, 81, 82, 83]. Thus, as their practical feasibility improves with ongoing technical advancements, along with their already identified potential to enable computational advantages in parallel quantum computation, this underscores the importance of developing a deeper understanding of their computational capabilities, which directly aligns with our second research question.

1.2.3 Benchmarking quantum computers

To adress our question about certifying the capability of quantum devices to correctly execute a broad class of quantum circuits, we delve, as previously mentioned, into the field of quantum benchmarking. Within this field, various methodologies have been explored, each offering different perspectives on device performance and reliability.

Some benchmarks are highly specialized, focusing on certifying specific quantum processes or states rather than providing a comprehensive evaluation of the device [84]. In contrast, prominent metrics or benchmarks like quantum volume and mirrored quantum volume aim to assess a quantum device's performance by analyzing deviations when executing generic types of random unitaries defined by specific circuit widths and depths [85, 86]. More recently, fragmented approaches have emerged to evaluate larger quantum devices with limited connectivity. While these methods offer valuable insights into the quality of larger quantum circuits and the effects of noise on thereon—critical for ongoing technical advancements—their certification protocols/schemes remain less robust [87], particularly when it comes to the predictive power needed to estimate the fidelity of executing non previously tested unitaries.

Another class designated as application-based benchmarks gauge quantum device performance across various quantum algorithms and processes. They pinpoint the performance thresholds for these algorithms and highlight the deficiencies of metrics like quantum volume when applied to the former. As [88] elucidates, errors influence circuit compilation of fixed quantum routines differently to how they affect generic random unitary operations of similar dimensions. This distinction sometimes undermines predictions from quantum volume metrics. However, these benchmarks have their inadequacies, as they are not algorithm-oblivious, potentially leading to biased results for devices reacting differently to distinct algorithmic solutions.

To address the lack of predictive power in quantum benchmarks, the work of [89] combines error models and neural networks. Notably, by leveraging data from quantum devices, the error models excelled in estimating realistic gate fidelities, sometimes revealing discrepancies from manufacturer claims. Meanwhile, the neural network models demonstrated the ability to predict the success probabilities of specific unitaries affected by circuit noise, including those not present in the training data. However, both approaches raise unanswered questions regarding scalability and the need for concrete guarantees on the accuracy of their predictions.

Verification protocols, as discussed in [90, 91, 92], provide a means to certify specific quantum algorithms or protocols with minimal overhead, even in the presence of noise. These protocols leverage the potential of blind delegation of computations, incorporating embedded traps to ensure the integrity of the algorithm. Identifying deviations at these trap points offers valuable insights into the overall fidelity of the process. While these protocols have yet to be adapted for general benchmarking with broader predictive capabilities, they excel in providing rigorous certification of a quantum algorithm's validity.

Lastly, performance-driven benchmarks, such as CLOPS [93], emphasize the speed of quantum operations, shedding light on the time resources required for specific quantum tasks. This approach is motivated by classical benchmarking, where the main focus lies on the speed of operations, due to high confidence

in the correct execution of operations. However, on quantum hardware fast execution alone cannot be used as an indicator for computational performance as all available devices struggle with noise.

After analyzing existing benchmarks and their limitations, we conclude that an effective quantum benchmark should be unambiguous, platform-agnostic, and capable of inferring optimal global properties from limited data without underestimating a device's capabilities. Additionally, it must offer rigorously certifiable predictive insights into computational performance. However, many existing benchmarks fail to meet these criteria comprehensively [94]. Consequently, with our final research question, we aim to develop benchmarks that not only embody these essential attributes but also deliver predictive power specifically tailored to our objects of analysis: quantum circuits.

Chapter 2
Contributions and thesis structure

This thesis follows the compilation format, consisting of a series of papers by the author. Each paper is presented in total, with its online version, and forms an individual chapter. One modification will be made: in each chapter, the abstract will be replaced with an overview section that serves as an extended abstract, aligning with the language and background provided in the earlier parts of the thesis. All included papers are the original work of the author and coauthors, directly addressing the research questions posed. This section will provide a detailed explanation of each paper/chapter, discussing its contribution to the research questions, its novelty, and its connection to the current literature and concurrent work.

2.1 Synthesizing quantum circuits of constant-depth

Adressing our first research question, we introduce quantum circuit synthesis techniques and compact quantum circuits that can be executed either with or without unbounded fan-in multi-qubit gates. Accordingly, we will divide the following section describing the contributions with respect to this distinction. Furthermore, these results form the content of Chapter 3 and the second part of Chapter 5, and they are fully detailed in [95] and [96].

2.1.1 Bounded fan-in qubit quantum circuits

In Chapter 3, we primarily address Research Question 1, which explores the feasibility of synthesizing quantum circuits for specific candidate problems that might demonstrate a quantum computational advantage while maintaining a constant depth (i.e., without being resource-intensive). Our selected problem candidates are generalizations of the problems presented in the works of [18] and [27]. In particular, we consider the family of relation problems[1] defined as follows,

$$\mathcal{R}_{\langle f,p\rangle}(x) := \left\{ y \;\middle|\; y \in \mathbb{F}_2^m :\ |y| \bmod p = f(x),\ \forall x \in \mathbb{F}_2^n \right\}. \tag{2.1.1}$$

with p being an arbitrary prime, and f an arbitrary function of the type $\mathbb{F}_p^n \mapsto \mathbb{F}_p$, which maps an n-dimensional vector over the finite field $\mathbb{F}_p$ to a single element in $\mathbb{F}_p$[2]. For context, the relation problem used in [27] for an unconditional quantum separation is a specific instance of these relations, where $p = 2$ and f, consequently a Boolean function, is chosen as the second least significant bit function.

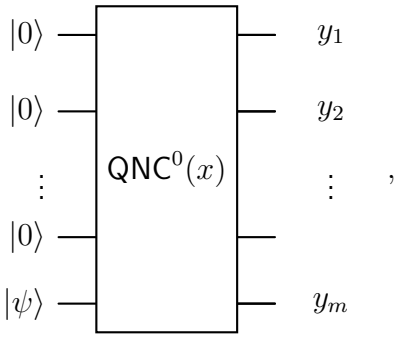
Attempting to solve an instance of this problem under the given conditions naturally leads to quantum computational models that are inherently parallel. For this reason, we utilized the measurement-based quantum computation (MBQC) model with constant-depth measurement patterns. In this model, state generation is generally of constant depth, as cluster states (universal resource states) can be created in constant depth. The effective depth of the circuit is measured by the number of sequential measurement

[1] A computational relation problem is defined by a relation of the form $R : \{0,1\}^n \mapsto \{0,1\}^m$ and is correctly solved by a circuit if, for an input x, it provides an output y such that $(x, y) \in R$.

[2] Note that subsets of the same family of relational problems are defined in chapters 3, 4, and 5 using different notations.

operations, which correspond to single-gate instructions in the standard circuit model. Our choice to analyze this model was based on three key considerations: first, it has been used for demonstrating constant-depth, unconditional quantum advantages while requiring minimal adaptive layers [18, 74, 97]. Second, previous studies have shown that the MBQC model provides a powerful framework for describing the parallelization of quantum circuits [98, 99]. Finally, these types of circuits have a formal description through phase polynomials [56], which is particularly relevant to our purposes due to their analytical properties. This feature suggests that these circuits may enable analytical studies of synthesis techniques and efficiency, which are not feasible with heuristic-based methods.

One of the main contributions we present in this chapter, as well as in the corresponding paper, is the ability to synthesize exact quantum circuits for all problems described in $Equation$ (2.1.1) for $p = 2$ and encompassing all possible Boolean functions $f : \{0,1\}^n \mapsto \{0,1\}$ that define the respective set of relational problems. These solutions exhibit two notable characteristics. First, they correspond to non-adaptive measurement-based quantum computations, utilizing a single measurement layer. Specifically, this approach aligns with the quantum circuits used in [18] to demonstrate quantum advantage while avoiding adaptive measurements, which pose significant technical challenges. However, our circuits rely on a specific input resource state: the GHZ state. While this state enables a formal description of these circuits as phase polynomials—facilitating a more rigorous analysis of computational efficiency, as we will discuss further—it cannot be generated in constant depth using bounded fan-in quantum circuits. Consequently, we have identified quantum circuits with input-independent advice states,



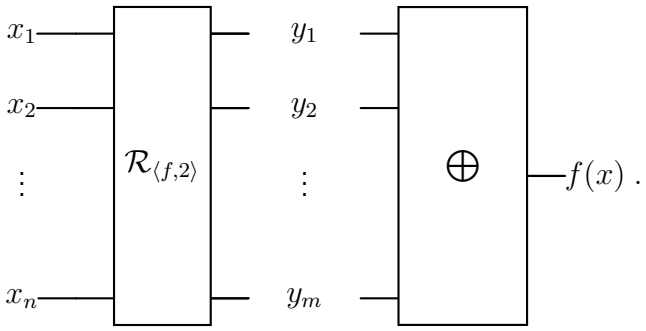
that solve the entire family of relational problems under consideration[3]. In terms of circuit complexity, these circuits belong to $\mathsf{QNC}^0/\mathsf{qpoly}$.

In exploring the elegant representation of quantum circuits as phase polynomials made possible by using the GHZ state as a resource state in the MBQC model, we identified particularly compact phase polynomials for all symmetric Boolean functions. We exhaustively tested all possible inputs up to a very large size and verified that these functions correspond to the optimal phase polynomials. This finding implies that they represent the smallest and most efficient quantum circuits in the MBQC model with a single measurement layer for evaluating the corresponding family of relation problems where the descriptive Boolean function f of the relation problems described in Equation (2.1.1) is a symmetric Boolean function. Notably, around the same time, a similar synthesis analysis for this model suggested polynomials of the same type in the context of studying contextuality within the same settings [1]. Although this work produced suboptimal solutions, it uncovered an interesting connection between the degree of the Boolean function and the necessity for contextuality. Furthermore, we have clarified that both the resource count and the complexity of the measurement operators, with respect to the Clifford hierarchy, are bounded by the degree of the Boolean function.

[3] In Chapter 4, we demonstrate that for certain subsets of the $\mathcal{R}_{\langle f,p\rangle}$ problems, the corresponding advice state can be eliminated, allowing for the construction of QNC^0 circuits. Also, although not part of the main focus of this thesis, we believe that the techniques introduced in [27] could allow for the removal of the advice state from the identified solutions for $p = 2$.

Finally, we note that the relational problems under consideration are computed within the framework of non-adaptive measurement-based quantum computations. As such, their solutions correspond to quantum strategies for XOR non-local games with biased input distributions—a class of multi-party correlations of significant theoretical interest [100, 101]. The connection between these relational problems and non-local games was first identified in [102]. It is, therefore, not surprising that the first demonstration of quantum advantage in [18] incorporated a non-local GHZ game into its framework. Similarly, the extension presented in [27], which demonstrates quantum advantage in the average-case hardness setting, is built upon a multi-party non-local game. Specifically, it involves generalized Mermin-type games, which are known to demonstrate exponentially higher correlations with the correct outcome when using quantum strategies compared to the best possible classical strategies. Beyond the capability of this model to synthesize quantum circuits that serve as candidate quantum strategies for non-local games, the corresponding classical-quantum deviations in winning probabilities could provide insights into the potential quantum advantages of the initial quantum circuits. Furthermore, our bounds on the complexity of the measurement operators and the number of necessary qubits with respect to the built-in Boolean functions establish a direct link between computational measures and resource requirements in the model. This connection extends equally to non-local games, where such relationships have been less understood and analyzed.

Evaluating Boolean Functions: Gate Count and Efficiency. In Chapter 3, we also address Research Question 2 regarding quantum advantages, adopting a slightly different approach compared to the analysis in the subsequent sections of this thesis. Specifically, we examine the efficiency of evaluating decision problems rather than relational problems using shallow-depth quantum circuits. For that, we consider that the solution circuits derived for computing the relational problems $\mathcal{R}_{\langle f,p\rangle}$ with $p = 2$ can be directly translated to circuits computing the corresponding Boolean function f that defines the relation. To accomplish this, it suffices to prepare the required resource state and compute the parity of the outcome string. By definition, the resulting bit corresponds to the output of the Boolean function f,



This approach—utilizing non-adaptive circuits to evaluate Boolean functions—was first explored in [102] to investigate the potential of near-term MBQC models for computing general Boolean functions. Here, the description of the model also incorporates unbounded fan-in parity gates frequently provided to the MBQC framework, enabling the construction of constant-depth quantum circuits within the $\mathsf{QNC}^0[2]$ class. In this setup, GHZ states can be prepared in constant depth, and the final parity operations can be implemented with a single gate. In the same work, it was demonstrated that even for simple functions like the AND of multiple variables, the non-adaptive MBQC model is highly inefficient, requiring an exponential number of elementary operations, $\Omega(2^n)$. However, for the pairwise AND function, defined by taking the AND operator between all the pairs of the input bits, the model required only a linear number of gates and qubits in the resource state, $\mathcal{O}(n)$, highlighting a significant gap in understanding the resources necessary for Boolean function evaluation.

Our contribution addresses this gap by reanalyzing the efficiency of Boolean function evaluation and determining the scaling for all symmetric Boolean functions, which are upper bounded by $\mathcal{O}(n^{\deg(f)/2})$ in their circuit size. Furthermore, we reanalyzed the resources required to evaluate these functions using only bounded fan-in gates, allowing for a direct comparison with classical circuits[4]. This analysis demonstrates that degree-2 symmetric Boolean functions exhibit an unconditional computational advantage in terms of gate count resorting to $\mathcal{O}(n)$ gates, while classically $\Omega(n\log(n))$ are required. Interestingly, these Boolean functions f correspond to the relations $\mathcal{R}_{\langle f,2\rangle}$, where an unconditional quantum advantage had previously been shown in [18, 27]. In contrast, higher-degree symmetric Boolean functions require more quantum resources than the best-known classical circuits.

Our findings suggest that the fixed depth constraints of non-adaptive MBQC limit the potential for creating more efficient quantum circuits for evaluating general Boolean functions, highlighting the model's limitations in efficiently parallelizing computations. Around the same time, another study examined the synthesis of quantum circuits for this model without depth constraints but with consideration for space constraints [103]. Although the solutions in that study involve technically more complex operations, such as adaptive measurements, they demonstrate that all symmetric Boolean functions can be efficiently evaluated using quantum circuits with linear depth and a single qubit[5].

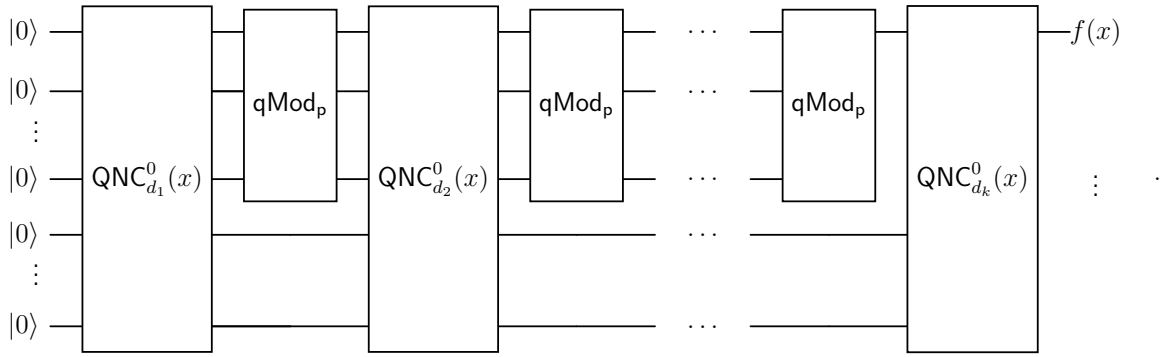
To summarize, in Chapter 3, we improve circuit synthesis techniques for non-adaptive measurement-based quantum computations. We generate and analyze circuits for a broad family of relations and corresponding decision problems, achieving model-dependent, efficient computations. This work also leverages an existing connection between quantum computations within the analyzed model and multi-party non-local games and presents new quantum strategies conjectured to be optimal for a large set of such games. Additionally, the identified quantum circuits, with the inclusion of unbounded fan-in parity gates, can be transformed into constant-depth evaluations of the analyzed relation problems and all Boolean functions, achieving constant and potentially minimal quantum depth within the $\mathsf{QNC}^0[2]$ circuit class. Although these circuits themselves face similar inefficiencies as previously analyzed, they have been effectively used as subprocesses within larger-depth quantum circuits to achieve highly efficient quantum evaluations of Boolean functions. For instance, the quantum circuit for the OR function was employed as a subprocess in an efficient emulation of the Threshold gate, demonstrating that the quantum circuit class $\mathsf{QNC}^0[2]$ collapses to QTC^0 [104]. This line of research has been further explored in a recent work [105], which achieves efficient, state-of-the-art constant-depth quantum solutions for Boolean function evaluations while employing additional quantum layers and relaxing temporal constraints.

2.1.2 Multi-qudit quantum gates

In the second part of Chapter 5, we further elaborate on the interest in marginally increasing the depth of quantum circuits to improve resource utilization. Additionally, we introduce multi-qubit gates that enable constant-depth realizations of Boolean functions. In particular, we demonstrate that for all prime p-dimensional qudit constant-depth quantum circuits with quantum p-dimensional unbounded fan-in modular gates, or p-dimensional unbounded quantum fanout gates, any symmetric Boolean function can be evaluated using a polynomial number of gates and in constant depth,

[4] This restriction results in circuits within the QNC^1 class, which, while still "shallow," are no longer of constant depth.

[5] While it is unsurprising that time is a valuable resource in quantum computation, one might consider Jozsa's conjecture, which posits that polylogarithmic-time quantum computations combined with polynomial-time classical computations could offer as much computational power as polynomial-time computations. Fortunately, as this thesis was being developed, this conjecture was disproven up to random oracles [51], reinforcing the importance of building quantum devices with extended coherence times to support polynomial-time quantum computations.



This includes the threshold function and implies that the collapse of $\mathsf{QNC}^0[p]$ to QTC^0 holds for all prime qudit dimensions, extending the collapse previously known for qubit quantum circuits [53, 77, 105, 52, 48, 106]. Specifically, for these circuits, we consider the relational problems $\mathcal{R}_{\langle f,p\rangle}$, where f defining the relation must correspond to a Boolean function, while p is a prime value. By evaluating the outcome string of these relational problems using a modular arithmetic operation, it becomes possible to compute the Boolean function f. The circuits are constructed using Fourier analytic techniques[6] and phase polynomials as in our qubit analysis.

Furthermore, we demonstrated that constant-depth quantum circuits implemented in higher prime dimensions can be realized using qubits with only a constant multiplicative increase in the gate count. Thus, providing any quantum unbounded fan-in prime modular arithmetic gate to qubit QNC^0 causes a collapse to QTC^0—a capability previously believed to require either binary quantum fanout or multi-qubit quantum parity gates. This finding contributes to the growing body of work aimed at understanding the potential of quantum circuits with multi-qudit quantum gates [77]. This result is achieved by mapping qudit operations onto tensor products of qubits. The qudit operations are constructed using exact unitary decomposition methods, as previously demonstrated in [107, 108]. Alternatively, we show that this collapse can also be achieved using any arbitrary classical modular prime gate combined with classical fanout[7]. Consequently, it follows that $\mathsf{QNC}^0_{\mathsf{cf}}[p]_{\mathsf{c}} = \mathsf{QACC}^0$ holds for all primes p [76]. Notably, this result contrasts sharply with classical analog circuit classes, where the choice of modular gates leads to different or incomparable complexity classes.

From a practical standpoint, although these results do not reveal significantly new quantum computational capabilities, they demonstrate that hardware manufacturing can be more flexible without compromising computational power. Moreover, we observe that employing qudits instead of qubits offers no inherent computational advantage for functions within these classes. However, qudits could enable simpler hardware implementations due to their potential for greater accuracy and efficiency in information storage and processing, as well as enhanced resilience to noise and improved error-correction schemes [109, 110]. For example, complex entangled states, such as multidimensional Greenberger-Horne-Zeilinger (GHZ) states and cluster states generated using our circuits, have demonstrated greater noise robustness than their qubit counterparts [111]. Additionally, these circuits, when augmented with arbitrary unbounded fan-in classical prime modular operations (MOD_p), can implement quantum subroutines of algorithms such as factoring, which is conjectured to offer exponential quantum advantages [53, 99], extending the interest of these classes of circuits beyond the use cases analyzed in this work.

[6] In these cases, the Fourier transforms are defined over Abelian groups, enabling the construction of circuits for functions over finite fields, of the form $f : \mathbb{F}_p^n \to \mathbb{F}_p$.

[7] In simple terms, classical fanout refers to the ability of quantum circuits to distribute classical information, such as measurement outcomes, to multiple destinations.

2.2 Extending the advantage of parallel quantum computation

After compiling the quantum circuits and presenting the previous results, one naturally questions whether these candidate circuits indeed hold the potential for quantum advantage and whether such advantages can be rigorously proven. Accordingly, we aim to use these quantum circuits to address our Research Question 2.

Push for relational problems. We note first that Chapter 4 and Chapter 5 focus solely on relational problems. Although a quantum advantage for decision problems is possible for instance, with respect to the gate count of the circuits —one of which we have previously identified—we emphasize that it was relatively small in scale. Moreover, for simple constant-depth quantum circuits in QNC^0, no computational advantage exists for decision problems compared to NC^0, which follows straightforwardly from lightcone arguments. However, exponential quantum advantages, with respect to circuit size/gate count, have been demonstrated for relational problems. Consequently, we focused on relational problems as the target computational problems of interest and the primary candidates for demonstrating significant quantum advantages.

Previously, it was conjectured that exponential quantum advantage was possible only for structured problems, such as hidden subgroup problems [112]. General quantum advantages for unstructured problems were thought to be limited to polynomial improvements, as in the quadratic speedup seen with Grover's algorithm [113]. However, though the conjecture stays untouched and is likely to hold, demonstrations showing that exponential quantum advantages for unstructured problems are possible appeared in a breakthrough result [114], with the important detail that it did not focus on decision problems. This insight introduces a triangular relationship among three properties: exponential quantum advantage, unstructured problems, and decision-type problems. It appears that only two out of these three properties can be achieved simultaneously. Thus, finding significant quantum advantages without requiring highly structured problems is more likely to be achieved with relational problems, making them particularly compelling in the quest for quantum advantage.

2.2.1 Constant-depth qupit quantum circuits

In Chapter 4, we broaden the scope of provable quantum advantages in parallel computation, extending them over larger and potentially practical classical parallel computation models, for all qudit constant-depth quantum circuits in prime dimensions. In addition, by incorporating realistic noise models into shallow quantum circuits, we demonstrate that these quantum advantages are robust to noise, preserving their feasibility under practical conditions. Finally, we quantify the resource costs for demonstrating quantum advantage, presenting a hierarchy of quantum advantage experiments and their associated resource requirements.

Higher dimensional qudit vs. biased polynomial threshold circuits. We compare shallow quantum circuits, as efficient and hardware-friendly as possible, with $\mathsf{bPTF}^0(k)$ constant depth classical circuits composed of unbounded fan-in gates that compute *biased polynomial threshold functions* (PTFs), which may be non-linear, but are constrained by a bias parameter k [115]. A polynomial threshold function[8] with bias k is defined as follows,

$$f(x) = \begin{cases} P(x), & \sum_{i=1}^{n} x_i \leq k \\ 1, & \sum_{i=1}^{n} x_i > k \end{cases} ; \quad , \tag{2.2.1}$$

[8] We note that the common definition in the literature takes $P : \mathbb{R}_2^n \to \mathbb{R}$, and sets $f(x) = \frac{1}{2}\left(1 + \mathrm{sgn}(P(x))\right)$. On the other hand here we are interested in polynomials of degree at most k over $\mathbb{F}_2^n$, with threshold behaviour determined by the bias parameter k.

with $P : \mathbb{F}_2^n \to \mathbb{F}_2$ being a polynomial over the binary field $\mathbb{F}_2 = \{0, 1\}$ and k restricting the maximum degree of P.

This parameter constrains their behavior to function as the Boolean OR when the input string has a Hamming weight of at least k. Conversely, the behavior can be inverted, in which case the bias parameter constrains it to the Boolean AND function, applied to input strings with a Hamming weight of at most $n-k$. For a constant, non-zero k (i.e., $k = \mathcal{O}(1)$), this class corresponds to Boolean circuits with unbounded fan-in AND, OR, and NOT gates of constant depth, that is AC^0 [43]. Notably, AC^0 represents the largest classical circuit class for which unconditional quantum advantages have been established in prior work [27]. When $k = \Theta(n)$, it encompasses the strictly larger TC^0 circuit class, which models neural networks. For intermediate scalings, this class includes access to unbounded fan-in threshold gates with biased activation regions (see Figure 2), such as majority gates with small fan-in, capturing some of the computational power of TC^0. Remarkably, a single threshold gate with $k = \text{polylog}(n)$ requires superpolynomial $\Omega(n^{\text{polylog}(n)})$-size AC^0 circuits to emulate. This leads to the inclusion sequence for classical parallel computational classes:

$$\mathsf{NC}^0 \subsetneq \mathsf{AC}^0 \subsetneq \mathsf{bPTF}^0(k) \text{ for } k = \omega(\log n). \tag{2.2.2}$$

Thus, we highlight that biased polynomial threshold circuits of constant depth provide a valuable framework for investigating quantum computational power, especially by analyzing the impact of varying the bias parameter and investigating their practical relevance (see Figure 3).

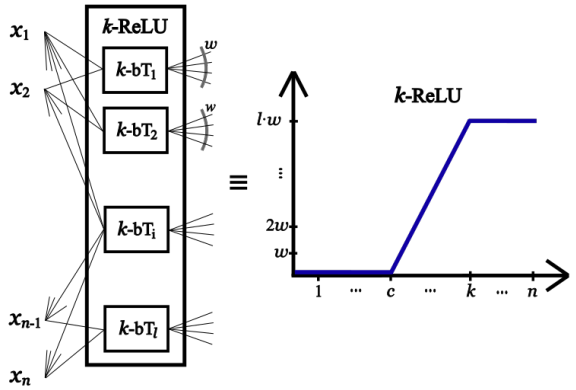


Figure 2: Representation of a k-ReLU gate within $\mathsf{bPTF}^0(k)$. This gate is equivalent to a ReLU gate, defined as $f(x) = \max\{0, x - c\}$ (where the center is shifted from 0 to c), up to an integer precision w, for any input string with a Hamming weight bounded by k. Our scheme considers ReLU : $\{0, 1\}^n \mapsto \{0, n - c\}$, which takes n-bit strings as input and interprets their Hamming weight as the input x.

In this context, we are able to show that for every prime p and large enough n, there is a constant-depth quantum circuit consisting of local gates over n 'qupits' consisting of $o(n^2)$ gates that has a constant correlation with the $\mathcal{R}_{\langle g,p\rangle}$ problem (Equation (2.1.1)), where $g(x) = -|x|/p \mod p$, which we name inverted strict modular relation problems (ISMRP). In contrast, any polynomial-size biased polynomial threshold circuit of constant depth and bias $k = n^{1/(5d)}$, with access to random bits, has exponentially small correlation $\exp\left(-\Omega\left(n^{3/5-\mathcal{O}(1)}\right)\right)$ with $\mathcal{R}_{\langle g,p\rangle}$. In other words, we show that biased constant-depth polynomial threshold circuits with a bias $k = \mathcal{O}(n^{1/d})$, require superpolynomial size to solve a relational problem that small (nearly linear sized) constant-depth qupit quantum circuits solve efficiently.

This result establishes even a stronger quantum advantage over classical models, surpassing those demonstrated in prior studies [27, 37, 18], by extending these advantages to the largest and most computationally powerful classical circuit classes studied to date (Equation (2.2.2)). Crucially, our bounds identify the maximum bias parameter for which $\mathsf{bPTF}^0(k)$ struggles to solve the relation problems considered,

as any larger bias k would enable these circuits to efficiently compute functions like parity and solve the problem with high probability using smaller circuits. Thus, the possible quantum advantage tied to this class of circuits, in terms of bias, reaches its theoretical limit.

Importantly, this classical circuit class models a wide range of neural networks and can, for instance, approximate activation functions like discretized ReLU, controlled by the bias parameter k (see Figure 2). It is also connected to transformer architectures, which are fundamental to modern large language models (LLMs). For instance, LLMs that use the transformer architecture can be simulated by AC^0 circuits when the attention mechanism is limited in certain natural ways [2, 3]. Meanwhile, TC^0, the standard for modeling discretized neural networks [4, 5], can simulate LLMs with realistic constraints on variable precision and autoregression [6, 7, 8]. Consequently, the biased polynomial threshold circuits we explore can simulate various neural network variants, supporting attention functions that extend beyond AC^0 and likely beyond TC^0. This opens the door to novel LLM architectures that still do not encompass the computational power and expressivity of shallow-depth quantum circuits.

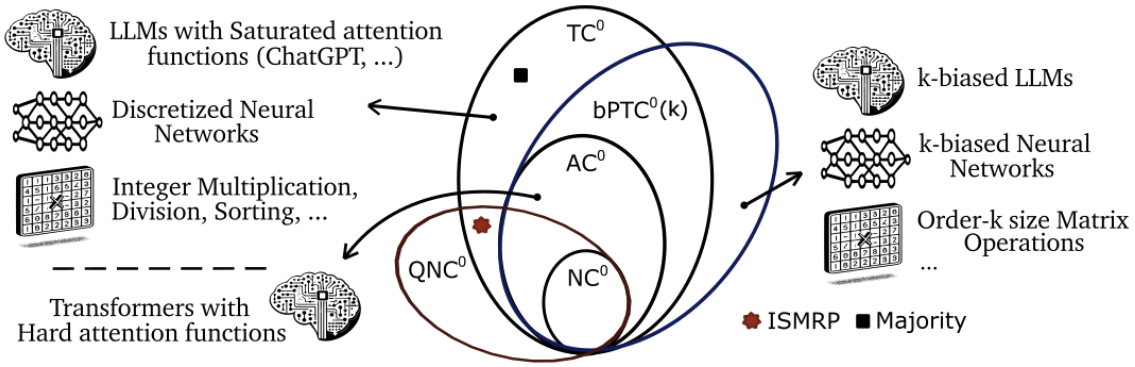


Figure 3: Relationships between the key classical and quantum circuit (complexity) classes considered. Notably, Majority is in TC^0 but not AC^0. The ISMR family of problems, introduced here, separates constant-depth quantum circuits from biased polynomial threshold circuits (bias $k = \omega(\log n)$). The latter class contains AC^0, and can solve NC^1-complete problems for super-logarithmic values of the bias and input lengths, suggesting non-trivial overlap with TC^0.

Moreover, this quantum advantage extends to an infinite family of problems, achieved using shallow-depth quantum circuits across all prime qudit dimensions with a standard, minimal finite gate set. This approach enables constant-depth implementations using the elementary operations available on standard quantum computing devices, including fault-tolerant architectures. Therefore, these explicit quantum advantages not only clarify the theoretical landscape but also have practical relevance, as many quantum computing platforms naturally operate in higher dimensions [110, 116].

While previous studies suggested that qudit non-local games with significant gaps in quantum versus classical winning probabilities could imply computational separations between classical and quantum circuit classes [117], such separations were not explicitly demonstrated. To address this, we introduced a family of non-local games that are called Modular XOR non-local games $\mathcal{G}_p$ (Figure 4), along with their corresponding relational problems. These games involve n parties, where each party P_i receives an input $x_i \in \mathbb{F}_p$ from a referee, forming a combined input string x with a Hamming weight $|x| = kp$, where $k \in \mathbb{N}$. Without further communication, each party responds with $y_i \in \mathbb{F}_p$, resulting in a collective output $y = (y_1, y_2, \ldots, y_n)$. The players win if the output y has Hamming weight equal to the additive inverse of k modulo p, i.e., $|y| = -k \mod p$[9]. Specifically, for all $\mathcal{G}_p$ with p prime, we constructively proved that classical strategies achieve, at best, exponentially lower correlation than quantum strategies as the number of parties increases.

[9] For context, $\mathcal{G}_2$ relates to the Mermin-Peres non-local game used in prior works [18, 27]

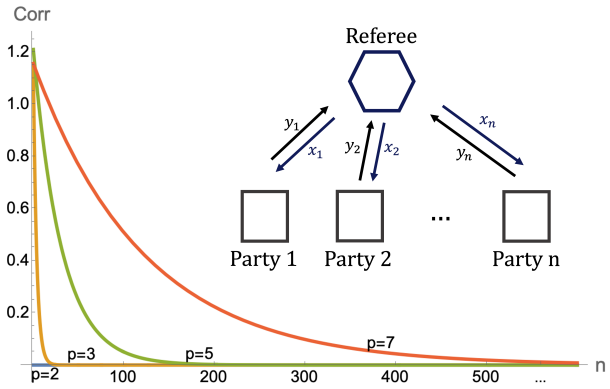


Figure 4: Representation of the Modular XOR games $\mathcal{G}_p$ involving n parties. The plot illustrates the upper bounds on the (winning) correlation of classical strategies for $\mathcal{G}_2$, $\mathcal{G}_3$, $\mathcal{G}_5$, and $\mathcal{G}_7$, whereas the optimal quantum strategy attains a maximum correlation of 1 for all $\mathcal{G}_p$.

We also emphasize that this separation between QNC^0 and $\mathsf{bPTF}^0(k)$ is achieved in the average-case hardness setting with uniform input distributions over binary inputs. To do this, our method elegantly manages *non-uniform input distributions* for qupit non-local games, handling the necessary dit-to-bit mappings to ensure that we use uniform (rather than biased) random restrictions for the biased polynomial threshold circuits. This approach allows us to obtain average-case $\mathsf{bPTF}^0(k)$ hardness with respect to the uniform distribution on binary inputs for the relation problems that demonstrate the separation. Additionally, it sidesteps the task of computing explicit success probability bounds for each game while still achieving comparable separations for all non-local games considered. This has been emphasised as an obstable in prior research, which predominantly relied on established quantum-classical distinctions in non-local games [18, 27]. Few developments beyond the conventional qubit setting were considered or achieved previously [118]. Finally, our relation problems $\mathcal{R}_{\langle g,p\rangle}$ remain binary, naturally suited for classical circuit classes, and do not impose hardness against Boolean circuits through arithmetic operations over other prime fields.

For the quantum solution, we generalized a method that translates optimal quantum strategies for these non-local games into small constant-depth quantum circuits that address the corresponding prime dimensional qupit relation problems. We specifically leveraged the use of generalized GHZ states in optimal strategies and the ability to generate multipartite entanglement between non-neighboring qupits in constant depth. We derived local unitary (LU)-equivalent generalized GHZ states and introduced a technique to efficiently describe the support of standard measurement outcomes of these states, factoring in phase dependencies. This enables us to use them in combination with an additional correction function that is computable in constant depth, relating to the outcome of the optimal quantum strategy. Moreover, it facilitates the determination of the success probability, which determines the quantum circuits' superiority over classical approaches.

Finally, in understanding the quantum advantages of parallel quantum computation, a requirement is a quantum candidate, but much of the challenge centered on defining the classical limitations of the corresponding parallel models. Our approach introduces a novel multi-switching technique that serves as a reduction tool, breaking down multi-output biased polynomial threshold circuits layer by layer into simpler classical computational models, such as decision trees, to which locality and light-cone arguments can be applied. This new result plays a pivotal role by simplifying these circuits into forms that reveal their locality, directly linking their performance to the optimal classical strategies for non-local games. More broadly, as these tools address not only binary decision problems but also search-type problems, this technique may hold independent significance, as neural networks—including large language models (LLMs)—are fundamentally sequence-to-sequence models, processing bit strings. Therefore, this approach

can be applied to relational problems beyond those examined here. Furthermore, parameterization by the bias parameter k provides a continuum of classical computational power levels, allowing for broader applicability. In particular, we use this framework to clarify the relationship between hardware requirements and the potential for quantum advantage.

Finite gate sets. Considering an infinite gate set, such as $\{\text{all single-qubit gates, CNOT}\}$ as assumed in Refs. [38, 78, 77] would certainly enable the realization of all our qudit circuits using qubit circuits in a noise-free setting [108, 107]. However, these solutions are not feasible in real hardware in constant depth [119] or in a fault-tolerant manner. Thus, in a more realistic context, any finite minimal gate set defined over a specific qudit dimension would require decomposition into its native gates if used to solve one of the $\mathcal{R}_{\langle g,p\rangle}$ problems from another prime dimension. In this regard, employing Solovay-Kitaev-type decompositions, these gates would likely necessitate log-depth decompositions to achieve sufficiently good approximations. Therefore, under this hardware-realistic definition of constant-depth qudit quantum circuits, we conjecture that for each prime p, there exists a relational problem $\mathcal{R}_{\langle g,p\rangle}$ that cannot be solved by any constant-depth quantum circuit using a minimal universal gate set for qudits of prime dimension q distinct from p, but can be solved by such a circuit using a gate set for dimension p. We propose that $\mathcal{R}_{\langle g,p\rangle}$ problems play a similar role for qudit constant-depth quantum circuit classes as Razborov-Smolensky-type modular problems do for classical $\mathsf{AC}^0[p]$ circuit classes. If validated, this conjecture could reveal important aspects of parallel quantum computations that depend on specific system dimensions and might also guide hardware manufacturers to consider quantum system dimensions beyond qubits.

Qubits Among all prime qudit dimensions, only two—2 and 3—enable us to translate quantum lower and classical upper bounds on correlations directly to corresponding bounds on success probabilities. Specifically, our results show that polynomial-size biased polynomial threshold circuits (of small bias) cannot solve the $\mathcal{R}_{\langle g,2\rangle}$ and $\mathcal{R}_{\langle g,3\rangle}$ problems with success probabilities appreciably larger than $1/2$ and $1/3$, which correspond to random guessing. In contrast, quantum circuits can solve the qubit case exactly, while in the qutrit case, the problem is solvable with a probability larger than and bounded away from $1/2$ by a constant.

In the qubit case we can in fact show that if classical circuits are required to solve $\mathcal{R}_{\langle g,2\rangle}$ exactly—that is, to produce a valid output string with certainty for all inputs—quantum advantages or separations of even greater magnitude are attained. This is achieved by showing that for large enough n, the size s of any depth-d k-biased polynomial threshold circuit, with access to random bits, that computes any valid y for $\mathcal{R}_{\langle g,2\rangle}$ is lower bounded as follows,

$$s = \Omega\left(\exp\left(\left(\frac{\sqrt{n}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)\right), \text{ for } k = \mathcal{O}(1). \tag{2.2.3}$$

We also recall that for $k = \mathcal{O}(1)$, constant-depth k-biased polynomial threshold circuits are equivalent to AC^0 circuits. Thus, our bounds in the exact-case augment upon prior work [27]. Furthermore, for the largest biased parameter considered, $k = n^{1/(5d)}$, we obtain the following size lower bound,

$$s = \Omega\left(\exp\left(\left(\frac{n^{3/10}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)\right). \tag{2.2.4}$$

We remark that this exact-case analysis is motivated by the fact that the quantum circuit in the qubit case is actually a deterministic solution to $\mathcal{R}_{\langle g,2\rangle}$, and therefore it is fair to compare it with classical deterministic circuits—to draw a complexity theoretic analogy, this is similar to comparing EQP and P, as

opposed to the more common comparison of BQP and BPP.

Thus, in the qubit case, we derived two bounds on the classical circuit size. The most robust demonstration of quantum advantage arises in the average-case hardness setting, where deviations can be directly accounted for. However, the exact-case hardness setting yields an even greater quantum advantage, establishing higher resource requirements for classical circuits and further amplifying the quantum advantage. This suggests that quantum advantage experiments could be achieved with fewer resources.

For this result, we developed a technique to capture the classical-quantum circuit separation using the standard representation for Boolean functions, specifically the algebraic normal form. More importantly, this approach avoids the need for non-local or contextual games [117], which have traditionally been essential in previous research, while showing that previous quantum advantages might benefit from new techniques to improve resource requirements and bring closer practical quantum demonstrations.

Noise-robustness. Finally, quantum systems are inevitably affected by noise, and error correction in the quantum realm is significantly more complex, raising questions about the *robustness* of computational advantage under noisy conditions. The third main result of Chapter 4 addresses this issue by proving that all our separations remain robust to noise, even when every step of the quantum computation is affected by local stochastic noise. Specifically, we show that for every prime p and sufficiently large n, there exists a relation problem $\mathfrak{R}_{\langle g,p\rangle}$ [10] such that, under qupit-based generalized local stochastic noise with a physical error rate below a constant threshold, there is a noise-resilient, constant-depth quantum circuit over qupits with local gates and all-to-all connectivity that computes it. This circuit, when equipped with logical $\left|T^{1/p}\right\rangle$-magic states[11] achieves a constant correlation with $\mathfrak{R}_{\langle g,p\rangle}$. In contrast, *any* (even noiseless) polynomial-size biased polynomial threshold circuit with large bias $k = n^{1/(5d)}$ has exponentially small correlation $\exp\left(-\Omega\left(n^{3/5-\mathcal{O}_d(1)}\right)\right)$ with $\mathfrak{R}_{\langle g,p\rangle}$.

It is worth noting that local stochastic noise is a standard model in quantum error correction, widely favored for its ability to account for gate-level noise, noisy input state preparation, and noisy measurements, while also accommodating (weakly) non-local errors. Locality in this context implies that the probability of a given Pauli error decreases exponentially with the number of sites it affects non-trivially. Furthermore, this model effectively captures fabrication faults and aligns with standard physical descriptions of noise, where the likelihood of errors diminishes exponentially as their weight increases [120].

More broadly, this result improves upon prior work in three main ways. First, it extends unconditional separations between noisy, shallow-depth quantum circuits and classical circuit classes. In particular, this is achieved for all prime qudit dimensions and against the largest classical circuit class to date. Notably, while most qudit dimensions specifically require logical T-type magic states as advice in their general formulation, we present a noise-resilient quantum advantage using Clifford qubit circuits that do not rely on such advice. Additionally, as a corollary, this establishes separations against NC^0 for each prime qudit dimension, potentially enabling near-term quantum advantage experiments due to the reduced resource requirements for this class and the favorable error-resilience properties of qudits.

Second, our approach extends the framework for noise-robust quantum advantages beyond the qubit Clifford model introduced in [35] by addressing qudit non-Clifford gates. Specifically, we demonstrate that for a particular CSS-type error correction code, this extension is achievable through the integration of logical T-type magic states, which can be affected by local stochastic noise, alongside qudit magic state injection protocols. The need for these more complex quantum circuits arises from the inability to violate Bell inequalities with stabilizer states for any qupit dimension beyond qubits, as shown in [121, 122, 123], which also suggests that the same limitation extends to quantum-classical separations in circuit complexity. Additionally, as part of our noise-resilient separation, we designed new quantum circuits in the form of

[10] Each $\mathfrak{R}_{\langle g,p\rangle}$ problem is related to our $\mathcal{R}_{\langle g,p\rangle}$ as the former is defined by the solution obtained from a quantum circuit encoded with an error correction code to solve the latter.

[11] These states are defined as $\left|T^{1/p}\right\rangle = \frac{1}{\sqrt{p}}\sum_{j=0}^{p-1} e^{i\frac{2\pi\cdot j}{p^2}}|j\rangle$ and related to magic states.

non-adaptive Clifford circuits with input-independent advice states to solve ISMR problems and achieve fault-tolerant quantum solutions.

Third, our work extends shallow-depth computational separations and error-correction mechanisms across arbitrary prime qudit dimensions, demonstrating their robustness. Specifically, we illustrate how the qupit surface code, when combined with the hard decision renormalization decoder, supports fault-tolerant implementation of these quantum circuits. This advancement includes the development of single-shot logical state preparation for qupits. We extend the 3D block construction from [124] to higher dimensions, showing that a particular measurement pattern yields a reduced state corresponding to a logical GHZ_2 state, up to local Clifford corrections. Furthermore, previous work used the minimum weight perfect matching decoder, which performs poorly in higher dimensions. By using a different decoder, we show that corrections can still be performed and recover the desired states with exponentially high confidence.

Resource estimates. When testing computational advantages with physical implementations, it is essential to pinpoint the circuit depth d and input size n (i.e., the number of input qubits) at which a transition occurs in the ratio between classical and quantum circuit sizes. Specifically, at what depths and input sizes does quantum advantage emerge? To address this, we estimated such values by solving for the points where the new asymptotic lower bounds on the size of the best classical circuits match the quantum upper bounds. We determine in Table 2, the smallest input size n for a given depth d, corresponding to the shallowest quantum circuits with all-to-all connectivity, where quantum circuits of different qudit dimensions outperform classical ones in solving the respective $\mathcal{R}_{\langle g,p\rangle}$ and $\mathfrak{R}_{\langle g,p\rangle}$ problems.

	NC^0 Regime				$\mathsf{bPTF}^0(k)$ Regime			
Quantum circuit type	Min F^*	$F=2$	$\cdots$	$F=8$	$k=1$	$k=2$	$\cdots$	$k=n^{1/(5d)}$
Qubits	2540	9364	$\cdots$	1.6×10^{8}	4.3×10^{13}	5.5×10^{16}	$\cdots$	5.0×10^{26}
Noisy Qubits	-	7.0×10^{10}	$\cdots$	1.5×10^{23}	1.8×10^{38}	8.0×10^{44}	$\cdots$	1.1×10^{75}
Qutrits	125162	1952660	$\cdots$	1.3×10^{10}	5.5×10^{14}	1.1×10^{18}	$\cdots$	2.3×10^{30}
Noisy Qutrits	-	3.5×10^{11}	$\cdots$	9.0×10^{30}	3.5×10^{52}	9.0×10^{60}	$\cdots$	3.4×10^{101}
Ququints	1.0×10^{8}	1.5×10^{9}	$\cdots$	8.0×10^{12}	9.5×10^{14}	1.7×10^{18}	$\cdots$	5.0×10^{30}
Noisy Ququints	-	3.0×10^{14}	$\cdots$	7.0×10^{33}	5.5×10^{52}	1.4×10^{61}	$\cdots$	7.0×10^{101}

Table 2: Estimates of the input sizes n required to demonstrate quantum advantage using the previously identified constant-depth quantum circuits, with all-to-all connectivity, in both noise-free and noise-affected settings, based on the quantum upper and classical lower bounds determined. We also examine the minimal fan-in (F) scenario by comparing quantum circuits with classical circuits of equal locality—keeping the same depth and fan-in at each level as the quantum circuits. Additionally, we analyze exact-case hardness bounds for the qubit setting to establish lower resource estimates. For the remaining qudit dimensions, we rely on average-case hardness to derive comparable estimates.

For context, the transition point for Shor's factoring algorithm is estimated to be $\sim$1,700 qubits, 10^{36} Toffoli gates, and a circuit depth of 10^{25} [125], while for the HHL quantum matrix inversion algorithm it is roughly 10^{8} qubits and a depth of 10^{29} [11].

In comparison, recent advances in quantum hardware have prioritized scaling up the number of qubits over extending coherence times, leading to a greater emphasis on shallow quantum circuits [126, 127]. Notably, the quantum advantages over constant-depth classical circuits with bounded fan-in gates, as studied in this thesis, require only thousands of qudits. These setups demonstrate classical intractability in tasks such as Bell violations [128, 129]. Classical circuits computing the considered relation problems with equal efficiency must have a depth of at least $d=\Omega(\log n)$, enabling a clear quantum advantage

when a quantum circuit solves the same problem at a strictly smaller depth. In practical scenarios with noise, the quantum circuit depth may increase by a constant factor due to error correction overheads, while the minimal classical circuit depth remains at least $d = \Omega(\frac{\log n}{\log\log n})$. While noise-driven adjustments raise the input size required to observe quantum advantage, these requirements are still significantly smaller, in terms of total resource counts, than those needed for other quantum advantage demonstrations.

Progressing up the hierarchy of computational power toward demonstrations of unconditional quantum advantage, the challenge lies in outperforming larger classical constant-depth circuit classes, such as biased polynomial threshold circuits. Achieving such quantum advantages would require significantly greater quantum resources, yet could still compare favorably to other *conditional* quantum advantage demonstrations in noise-free settings. These quantum circuits maintain the advantage of shallow depth, potentially making them more practical for near-term quantum hardware. However, comparisons between noise-affected quantum circuits and noise-free classical circuits often demand unrealistic resources, exposing a fundamental imbalance. Classical error-correcting codes, which have already benefited from decades of refinement, may also be essential for achieving the exponentially high-efficiency rates to those expected of error-corrected quantum circuits [130]. Although noise levels in quantum and classical systems are unlikely to converge [131, 132, 133], resource estimates that account for noise on both sides—combined with advances in quantum hardware—are anticipated to bring these comparisons closer to those of noise-free models.

Beyond the estimates in Table 2, our bounds extend to all prime qudit dimensions and facilitate the exploration of quantum circuits with varying hardware connectivities. In noise-resilient scenarios, all-to-all connectivity is required, while in noise-free comparisons, architectures can range from $p+1$-dimensional configurations for qudit dimension p, to fully all-to-all connectivities. Among these, our estimates represent the lowest obtained for equivalent unconditional quantum separations, with the qubit case achieving the smallest resource requirements due to its derivation from exact-case hardness bounds. Although this setting reflects a less robust form of quantum advantage, it significantly reduces resource requirements, bringing theoretical predictions closer to the capabilities of current quantum devices. These estimates could be improved for all qudit dimensions by requiring classical circuits to better approximate quantum solutions and by adding connectivity restrictions to classical circuits. While these constraints are not formally part of the NC^0 definition, they can be chosen to reflect the limitations of realistic classical hardware and can lower the input size needed to demonstrate quantum advantage [75]. More generally, tighter lower-bound techniques and the discovery of computational problems with greater quantum advantages could further refine these estimates.

Finally, our estimates indicate that these quantum advantage experiments could serve as powerful quantum benchmarks, providing a systematic framework for evaluating and comparing computational capabilities. To explore larger quantum computational advantages, one could test classical circuits with larger fan-in gates, establishing new benchmarks and creating a structured "ladder of quantum advantages" to assess increasingly stronger computational separations. This hierarchy can be expanded by examining the ability of shallow quantum circuits to outperform more advanced classical circuit classes, such as the biased polynomial threshold circuits analyzed in this work. Adjusting the bias parameter within these circuits further amplifies potential quantum advantages. Notably, the described levels based on the bias are separated by small multiplicative factors, making them an effective tool for benchmarking small scale hardware improvements.

2.2.2 Constant depth multi-qubit Toffoli quantum circuits

In the first half of Chapter 5 we report our final set of contributions related with extending quantum advantages in parallel quantum computation. We first demonstrate that constant-depth quantum circuits with quantum advice can solve relation problems that are intractable for any polynomial-size $\mathsf{AC}^0[p]$ circuits.

Specifically, for all primes p, we show that $\mathsf{QNC}^0/\mathsf{qpoly} \not\subseteq \mathsf{AC}^0[p]$. For context, this result extends the $\mathsf{AC}^0[2]$ separation established in [27] to general primes. To prove this, we once again resort to relation problems of the $\mathcal{R}_{\langle f,p\rangle}$ type. For each separation involving a specific $\mathsf{AC}^0[p]$ class, the corresponding prime p serves as the parameter for the relation problem, while the function f represents another distinct Boolean modular prime operation: for instance, Mod_q with $q \neq p$. This allows us to extend the classical lower bound to all $\mathsf{AC}^0[p]$ classes for prime p. Concurrently, we show that by exploiting higher-dimensional gate sets, these problems can be solved in constant-depth by quantum circuits equipped with a copy of a high-dimensional GHZ state.

This separation achieves only a quantum/classical distinction based on error probabilities, differing by a constant value. To strengthen these separations, we consider the parallel repetition of these modular problems as proposed in [27, 34]. In our results, we simplify the proof in [27] by focusing on the one-sided error of the quantum circuits and leveraging foundational results in circuit complexity, such as the Vazirani XOR lemma [134] and the Razborov-Smolensky separations [19].

We proceed by demonstrating that for $\mathsf{QAC}^0_{\mathsf{cf}}$ circuits—characterized by the use of multi-qubit controlled Toffoli gates, which enable classical but not quantum fanout—we achieved an unconditional separation between this complexity class and all $\mathsf{AC}^0[p]$ classes for any prime p. This result advances the exploration of minimal resources required for quantum-classical separations, a question recently posed in [135]. By excluding quantum fanout and introducing only *classical fanout*, our result may arguably represent the weakest resource introduced to date to achieve such a separation [136, 76, 99]. Additionally, it is worth emphasizing that this result is consistent with the standard definition of QAC^0, as outlined in the literature [137, 135, 138, 139], while incorporating classical fanout. Classical fanout refers to the ability to copy classical information obtained from mid-circuit measurements an arbitrary number of times—a relatively simple operation from a technical perspective that does not conflict with the no-cloning theorem.

To achieve this result, we show that QAC^0 circuits can generate higher-dimensional GHZ states. This is accomplished using optimal depth and size circuits for GHZ state preparation, leveraging poor-man's cat states and balanced binary trees as the entanglement structure of the resource states. Combined with the quantum circuits previously analyzed within the $\mathsf{QNC}^0/\mathsf{qpoly}$ class, this enables us to prove that for every prime p, there exists a constant-depth quantum circuit employing multi-qudit Toffoli gates that computes a relation problem that cannot be computed by any polynomial-size $\mathsf{AC}^0[p]$ circuit. Consequently, when qubit Toffoli gates are extended to allow control over arbitrary single-qubit operations, we achieve an unconditional separation between this complexity class and all $\mathsf{AC}^0[p]$ classes simultaneously.

It is important to note that achieving this result with our techniques requires considering qudits of higher dimensions. Furthermore, we observe that $\mathsf{QAC}^0_{\mathsf{cf}}$ circuits can prepare a GHZ state without the need for unbounded parity gates, as proposed in [135]. Instead, our $\mathsf{QAC}^0_{\mathsf{cf}}$ circuits utilize quantum measurements and adaptive operations to generate these states. Interestingly, the capability to produce this state does not imply the ability to compute the parity function. We conjecture that generating the GHZ state through a unitary procedure (i.e., without measurements) would require quantum fanout. Thus, the question of whether the parity function can be computed in QAC^0 without fanout gates, or even with only classical fanout gates, remains an open problem for further investigation [52].

Adding to the interest of these results, recall that although further research is needed to explore the feasibility and potential technical improvements, recent advances and practical demonstrations suggest that some quantum device platforms inherently support multi-qubit operations. In particular, some of these platforms facilitate the construction of multi-qubit Toffoli gates [79, 80, 81, 82].

Finally, we outline certain limitations of the specific instances of $\mathcal{R}_{\langle f,p\rangle}$ problems that have been employed to demonstrate quantum-classical separations, particularly those for which there are efficient quantum upper bounds. Notably, for the previously considered $\mathcal{R}_{\langle f,p\rangle}$ problems—where p is a prime, and f is another modular prime function—there exist $\mathsf{NC}^0[q]$ circuits capable of solving them [12]. This finding

[12] These circuits are composed of bounded fan-in gates, as in NC^0, but additionally include *a single* unbounded fan-in MOD_q gate.

characterizes $\mathsf{AC}^0[p]$ as the largest constant-depth classical circuit class for which a quantum-classical separation can be demonstrated in the plain model using the aforementioned relational problems. Indeed, we conjecture that this specific set of relational problems might suffice to achieve unconditional separations between the aforementioned $\mathsf{AC}^0[p]$ classical circuit classes and less powerful quantum circuit classes, such as QNC^0. However, separations against larger classes like TC^0 (or beyond) are unattainable with these problems.

Nonetheless, more complex versions of the $\mathcal{R}_{\langle f,p\rangle}$ type relation problems remain viable, as the function f permits a broad spectrum of options. However, separations for larger classes have so far been demonstrated only in the interactive setting, utilizing entirely different types of computational tasks [74].

2.3 Verification-inspired quantum benchmarking

In Chapter 6, we address our Research Question 3, which asks whether we can formally certify the capability of a quantum device to execute a particular set of quantum circuits. Fortunately, we were able to answer this question positively by proposing a new quantum benchmarking protocol. In short, this protocol can certify a quantum device's ability to execute all elements within certain tested class of quantum circuits correctly. Furthermore, repeated use of the protocol enables a robust characterization of the device and provides a more precise description of the general class of quantum circuits it can reliably execute.

Recall that quantum benchmarking aims to assess and compare the performance of various quantum devices, with one of the main objectives being to help end-users to determine the suitability of a specific device for their applications. Currently, there is no universally accepted standard for benchmarking. Although the implicit goal is to predict the computational performance of quantum devices, most benchmarking approaches to date have been largely heuristic. Recognizing the need for a more formal foundation for these assessments, recent efforts [89] have sought to theoretically reinforce the validity of insights derived from heuristic quantum benchmarks. A natural and desirable step forward is to transition to a more formal and predictive benchmarking framework, which aligns directly with the focus of our research question. With this goal in mind, we first propose new paradigms that extend the properties previously sought in benchmarking schemes.

For a fair assessment of quantum devices, we propose that benchmarks should adhere to several key principles. They should be *noise-model agnostic*, avoiding assumptions about specific noise effects to allow a broad characterization of interference, and *hardware agnostic*, ensuring performance is evaluated impartially across different hardware types. Benchmarks should also be *algorithm agnostic*, so they are not limited to evaluating devices based solely on existing algorithms, particularly given the evolution of quantum algorithms and sporadic emerging breakthroughs. Instead, they should provide insight into performance across a wide range of potential quantum processes. Additionally, benchmarks require *predictive capacity* with results that are *certifiable*, offering reliable predictions about device performance. Finally, benchmarks should be *practical*, *efficient*, and *scalable*, with clear instructions and minimal computational overhead to ensure their viability as quantum devices grow in size and complexity. In summary, this new paradigm aims to establish a more rigorous framework for making predictive and reliable statements about quantum device performance.

A crucial step toward advancing quantum benchmarking is recognizing its connection to the closely related field of quantum verification, as highlighted in [140], since both focus on certifying quantum computations. Our main contribution is to establish a strong connection between these fields by presenting a generic method for compiling blind verification schemes into benchmarking protocols. This compiler leverages a reduction from worst-case to average-case computations, an inherent feature of blind verification schemes with strong security. We further prove that benchmarking schemes obtained through this verification-inspired approach achieve new, stronger benchmarking standards. Applying this compiler to the robust quantum verification scheme of [91], we derive a concrete benchmarking protocol that

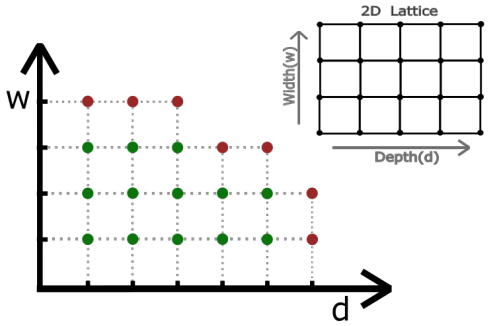
inherits favorable properties, including efficiency and noise robustness. The new scheme is free from heuristics and assumptions about the noise structure, and it is largely platform-agnostic, requiring only that the platform supports adaptive measurements. It provides results that serve as certificates of computational power rather than heuristic performance indicators. Moreover, it is efficient, as the integrated cryptographic techniques do not impose additional hardware overhead.

A distinctive feature of this quantum benchmarking method is its departure from prior approaches that relied on heuristics or validated only specific algorithmic instances. Instead, we leverage cryptographic tools—blindness and verification—to achieve broad quantum device characterization, resulting in a robust benchmarking method that addresses existing methodological gaps. Formulated within the MBQC model, our benchmark can be translated into standard circuit models with negligible overhead. Specifically, it verifies the dimensions of graph states, and consequently, the class of circuits derived from translating our initial computation or unitary operation into the MBQC model using a resource state based on these graphs or their subgraphs.

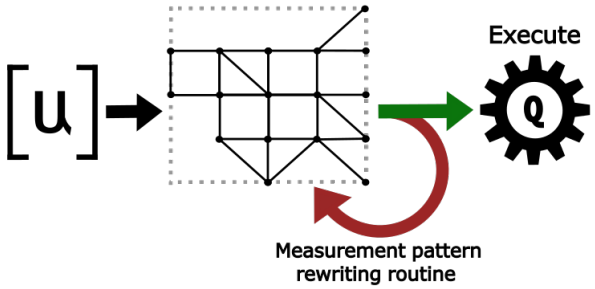
Deploying our benchmark can be precisely understood as a process designed to characterize quantum devices in terms of the computational classes they can successfully implement. This process also includes cataloging these classes to provide valuable insights for their future users. Specifically, since classes in our framework are represented as graphs, the benchmark could be implemented by testing graphs of progressively larger sizes. However, this approach faces the challenge of an exponentially expanding search space. To address this, we propose a more practical and informative characterization using 2D cluster states parameterized by two values: width and depth. This choice is motivated by the fact that 2D cluster states serve as universal resource states within the MBQC model, and their graph structure resembles that used in previous benchmarks, such as quantum volume, while offering formal guarantees. Additionally, it provides a straightforward analogy with 2D circuits. In particular, this benchmark would yield a figure of merit, as illustrated in Figure 5a.

With such a characterization, a device user could determine whether their computation will execute correctly on the device. This requires only translating the unitary operation into a planar MBQC measurement pattern and verifying whether the resulting graph matches or is a subgraph of a certified graph. If the initial graph does not align with any certified graph, the unitary can be rewritten according to specific rules, potentially as a either deeper/sequential or more parallel computation, which might then correspond to a verified graph (see Figure 5b). This approach seamlessly integrates the benchmark into a practical workflow, enabling optimal resource utilization and extending the computational applications of quantum devices. Overall, it facilitates the design of circuits tailored to the specific capabilities of quantum devices, offering valuable insights into their operational potential.

Lastly, verification-inspired quantum benchmarking protocols tend to reject noisy hardware compared to more flexible and less rigid benchmarking techniques, which might be a critique of their use in near-term devices. This feature reflects the inability of such hardware to meet the standards for guaranteed computational power. Nevertheless, recent advances in the development of quantum hardware have demonstrated that our approach to quantum benchmarking is becoming practical. First, an academic experiment conducted on ion-trap devices demonstrated another successful implementation of verification protocols [141], which "indirectly" implements the proposed benchmark. Later, experiments by a commercial hardware provider using the same hardware platform showed the successful implementation of the quantum benchmark here proposed [142]. Moreover, this experiment successfully certified the device's ability to execute any measurement pattern in the MBQC setting with a depth of five and a width of six, achieving a fairly high success probability. Additionally, it verified the execution of a measurement pattern involving 52 vertices.



(a) Depiction of a device characterization using tests based on graph states on 2D lattices.



(b) Illustration of the graph state optimization process to assess and enhance quantum device accuracy.

Figure 5: Once a quantum device is certified for various graph states, one can ask which computations, i.e., unitaries, can be certifiably run on the device. Unitaries can be compiled onto a graph state for the equivalent MBQC pattern. However, the same unitary can be mapped onto different graph states. Some of these might be certified on the device; others might not. Therefore, one can attempt to rewrite the computation to find a graph state that is certified on the device so it can implement the computation.

2.4 List of publications

List of publications of the author that directly form a part of this thesis,

- Michael de Oliveira, Luís S. Barbosa, and Ernesto F. Galvão, "Quantum advantage in temporally flat measurement-based quantum computation", Quantum 8, 1312 (2024),
  DOI: https://doi.org/10.22331/q-2024-04-09-1312, content of Chapter 3 and Ref. [95].

- Min-Hsiu Hsieh, Leandro Mendes, Michael de Oliveira, and Sathyawageeswar Subramanian, "Unconditionally separating noisy QNC0 from biased threshold circuits of constant depth", arxiv pre-prints,
  DOI: https://arxiv.org/abs/2408.16378, under review at Nature Communications, content of Chapter 4 and Ref. [143].

- Alex Bredariol Grilo, Elham Kashefi, Damian Markham, and Michael de Oliveira,"The power of shallow depth Toffoli and qudit quantum circuits", arxiv pre-prints,
  DOI: https://arxiv.org/abs/2404.18104, under review at ACM Transactions of Quantum Computation, content of Chapter 5 and Ref. [96].

- Johannes Frank, Elham Kashefi, Dominik Leichtle, and Michael de Oliveira, "Heuristic-free verification inspired quantum benchmark" , arxiv pre-prints,
  DOI: https://arxiv.org/abs/2404.10739, under review at Quantum Science and Technology, content of Chapter 6 and Ref. [144].

Part II
Results

Chapter 3

"There is a way to do it better - find it."

— Thomas A. Edison

Synthesizing quantum circuits of constant-depth

Overview

Several classes of quantum circuits have been shown to provide a quantum computational advantage under certain assumptions. The study of ever more restricted classes of quantum circuits capable of quantum advantage is motivated by possible simplifications in experimental demonstrations. In this section, we study the efficiency of non-adaptive measurement-based quantum computation. We propose new constructions for the deterministic computation of $\mathcal{R}_{\langle f,p\rangle}$ type relation problems with $p = 2$, and arbitrary Boolean functions, drawing on correlations present in multi-qubit Greenberger, Horne, and Zeilinger (GHZ) states. We characterize the necessary measurement complexity using the Clifford hierarchy and also generally decrease the number of qubits needed with respect to previous constructions. In particular, we identify a family of Boolean functions for which deterministic evaluation using non-adaptive MBQC is possible, featuring quantum advantage in width and number of gates with respect to classical circuits. The paper providing the content for this section can be found in [95].

3.2 Introduction

One of the primary motivations for studying quantum information and computation, along with the search for practical advantage, is to clarify the longstanding question of what generates the classical-quantum separation in information processing power [145]. Phenomena such as non-locality and contextuality have been identified as possible sources of quantum computational advantage [146, 147, 148], among others [149, 150, 151]. It appears that no single phenomenon can be associated with all forms of quantum information processing advantage, so it is important to identify and study quantum advantage in different models and regimes.

In this pursuit, the measurement-based quantum computation (MBQC) model, first presented by Raussendorf and Briegel in [152] as a sequence of adaptively selected single-qubit measurements on a highly entangled quantum state, is a natural setting to study quantum-to-classical separations. In particular, it allows for demonstrations that specific structures of non-local correlations between qubits are necessary for universal quantum computation [153]. Nevertheless, a temporal structure for the measurements is simultaneously imposed to achieve universality [154, 155, 156]. This means statically selected measurements are not computationally expressive enough to implement arbitrary quantum algorithms, even with access to highly correlated resource states. A judicious choice of side classical computation and control is essential for universality in the MBQC model.

The importance of time structure in measurement-based quantum computation is not fully understood. For instance, it is conjectured that classical computers cannot efficiently simulate instantaneous quantum polynomial (IQP) circuits [157], despite the fact that these circuits of commuting gates have no temporal structure, and can be implemented with no adaptivity in MBQC [158]. These results have motivated the study of temporally flat computation for demonstrations of quantum advantage [159, 160, 161]. Despite having an output distribution that is hard to simulate classically, no practical application for IQP circuits has been found. Even for quantum circuits with temporal order, constant improvements in classical simulation techniques contribute to enlarging the classes of circuits that are classically efficiently simulable.

This paper studies non-adaptive MBQC computations which have no temporal structure for the measurement operators. In this setting, there exist circuits with quantum advantage, defined as those that cannot be efficiently simulated classically [162, 163]. Here we will be interested in investigating how this model can be applied to decision problems. Previously, in references [154, 102, 164, 1, 165] this model was studied with the goal of evaluating Boolean functions. In the next section we briefly describe the main results we obtain here, relating to the use of non-adaptive MBQC for the evaluation of Boolean functions.

3.2.1 Our results

For temporally unstructured MBQC computations, we propose new constructions that synthesize quantum circuits for the deterministic evaluation of Boolean functions. These constructions reduce the number of qubits required in the GHZ resource states used, especially for the case of symmetric Boolean functions. We also improve the circuit synthesis process, removing an exponential scaling of previous constructions on the degree of the Boolean function [164, 1]. Regarding the complexity of the single-qubit measurements, we also characterize the maximum level of the Clifford hierarchy required for deterministic Boolean function evaluation.

Theorem 3.2.1. *(Informal) Any Boolean function f can be evaluated deterministically in the non-adaptive MBQC model using measurement operators of the* $\deg(f)$*-level of the Clifford hierarchy.*

This result describes a measure of complexity for the evaluation of Boolean functions in the model while defining the type of measurement operators necessary to maximally violate a multipartite Bell inequality with dichotomic observables and outcomes [102]. Furthermore, this strengthens the idea that the

degree of a Boolean function that can be computed with certain given quantum resources introduces a hierarchy for quantum correlations, connecting measures of computational complexity with measures of non-classicality beyond the binary characterization, as suggested in [1].
We translate the abstract description of the quantum circuits obtained from the constructions into specific circuits based on a fixed gate set. Also, we characterize how the solutions produced by the constructions are related to various circuit complexity measures. This enabled us to prove a quantum over classical separation on the circuit level for the computations of a specific class of Boolean functions with degree two.

Theorem 3.2.2. *(Informal) Any classical circuit with unary and binary Boolean operators with single fan out computes symmetric Boolean functions f with* $\deg(f) = 2$ *with* $\Theta(n * \log_2(n))$ *gates and circuit width. In contrast, realizations of the non-adaptive MBQC model compute these functions with* $\Theta(n)$ *gates and circuit width.*

The proof compares these quantum circuits with the length of the classical Boolean formulas, which describe classical circuits without a memory. The comparison is indeed very strong in the sense that the quantum circuit also does not use any memory (see Figure 6 for a pictorial representation of this subclass of circuits). Interestingly, this result also has an interpretation in terms of Bell inequalities. In particular, the evaluation of the specific family of Boolean functions for which we identify quantum advantage corresponds to maximal violations of generalized Svetlichny inequalities [166]. Therefore, a maximal violation of these inequalities starting from a specific size implies a corresponding circuit separation [1]. More precisely, any quantum state that can be prepared and measured with linearly bounded quantum circuits that violate these inequalities maximally implies a corresponding circuit separation.

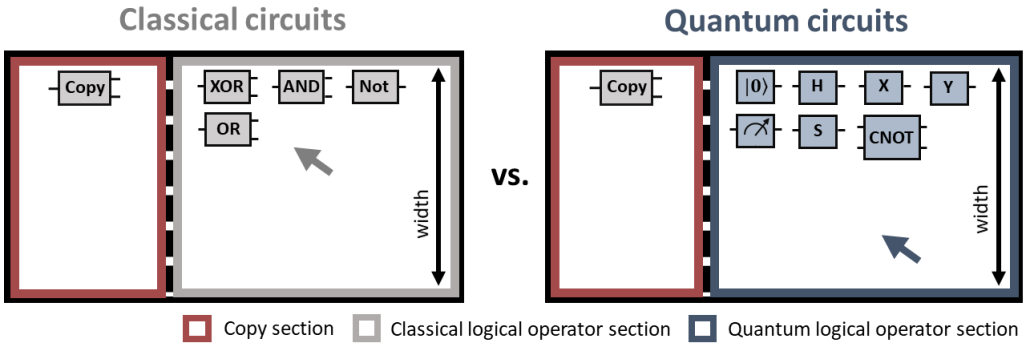


Figure 6: Illustration of the circuit classes for which we prove a separation in Theorem 3.2.2. For the classical class, the input can be copied only in the copy section, which precedes the logical operator section. For the quantum circuits, the same occurs. This prevents the circuits from employing the copy operation at any point in time, which would replicate the effect of memory in the circuits.

Finally, we extend the previous analysis to higher degree symmetric Boolean functions, and conjecture that these have no advantage with respect to classical circuits.

Conjecture 1. *(Informal) Symmetric Boolean functions f with* $\deg(f) \geq 3$ *evaluated within the non-adaptive MBQC model do not entail circuits with better scalings than classical circuits with unary and binary Boolean operators with single fan out.*

This conjecture is based on several independent results. The first is the exponential lower bound proven in [102] for the general AND function which is symmetric and has an instance for any possible degree.

[1] Recall that the separation in Theorem 3.2.2 is proven to hold asymptotically. Therefore, there is a minimum problem size above which all the quantum circuits are more efficient than the best classical circuits.

Then, on the same level, we conjecture and provide numerical support that symmetric Boolean functions require quantum states whose size scales at a greater rate than the corresponding number of classical bits required in the optimal classical circuits for the same computations. Furthermore, we show that with the Clifford+T gate set the necessary measurement operators cannot be synthesized exactly, and always need to be approximated. In contrast, there is no difficulty in computing the same function deterministically with classical circuits. This illustrates some of the restrictions resulting from the imposed flat temporal order.

3.2.2 Related work

Non-local games. Computations within the non-adaptive MBQC model were proven to have a one-to-one correspondence to multi-party Bell inequalities with dichotomic observables and outcomes [102]. Therefore, the classical and the quantum bounds determined in previous works for this type of Bell inequalities immediately translate to computational efficiencies of the respective functions in the non-adaptive MBQC model [167]. For the quantum bounds, we obtain the maximal efficiency that can be obtained from quantum resources to compute the respective Boolean functions, and equivalently, the same happens for classical bounds. From this relation, we can compare our work with a large spectrum of results and techniques on the optimal strategies for non-local games [167, 168, 169].
For instance, in [170, 171], the authors consider arbitrary symmetric non-local games, which correspond to the symmetric Boolean functions we study here. Their results complement our own in the probabilistic setting. We determined the size of the quantum resource to obtain deterministic computations. At the same time, they considered the efficiency obtained if a quantum resource with a size equal to the input string is evaluated, with the additional comparison to a classical resource of the same size. Their main results show a mean advantage of $\sqrt{\log(n)}$ between the use of classical or quantum resources. Moreover, the magnitude of this separation appears to be similar to the separation demonstrated in Theorem 3.2.2 due to the $\log(n)$ term. Still, it is not comparable to our results, given that they consider the classical limit based on measurement-based computations with classical resources, while we allow for arbitrary classical circuits without memory.
Furthermore, in the study of communication complexity, the non-local boxes usually employed can be interpreted as non-adaptive MBQC computations [172]. This observation relates this model to results obtained in the study of non-local boxes. The study of this relationship is two-fold. First, in references [173, 174, 175], the hypothesis of axiomatizing quantum mechanics with information-theoretical principles intends to limit the quantum correlations allowed with these principles. Therefore, the resulting limits to the non-local boxes' efficiencies also relate to the maximal efficiency of the corresponding non-adaptive MBQC computations [176, 177, 178]. On the other hand, non-local games defined on a non-binary set of observables and outcomes can be interpreted as a generalization of the non-adaptive MBQC model to qudits , together with the respective classical linear side processors [179, 180].

Quantum circuits. A breakthrough result by Bravyi, Gosset, and König [181] shows that a class of relation problems can be solved in constant depth with a quantum circuit, while requiring logarithmic depth in a classical computer [181]. This result has been extended in various manners [74, 73]. In [182], the analysis is extended from worst-case difficulty to average-case difficulty. In [183], the probability with which a classical circuit correctly computes the relation problem was exponentially reduced with the size of the input string. Dealing with the same separation, the works of [184, 185] have made it more robust against noise.
The first connection of our work to these results is that the proof relies on non-local games that are fundamental to the separation. These are naturally described with the non-adaptive MBQC model, which could be used to extend the relation problems with different non-local games, as done in [183]. Nevertheless, a direct connection to these circuit separations can be made with the work of [27]. In this paper, the

authors present the parity halving problem PHP : $\{0,1\}^n \mapsto \{0,1\}^n$, defined for all $x \in \{0,1\}^n$ with even Hamming weight and $y \in \{0,1\}^{n+1}$ as

$$|x|/2 \equiv |y| \mod 2 \,, \tag{3.2.1}$$

with $|x|$ and $|y|$ describing the Hamming weight of the respective bitstrings. The authors prove that this relation problem can be computed in QNC^0 while not being computable with any polynomial size circuit of the AC^0 class. Furthermore, this is precisely a sub-problem of the relation problem built in the quantum circuit that we use to prove a quantum advantage for a decision problem. In particular, the Hamming weight of the input string and the output string from the measurement layer performed on the GHZ state describes the full parity halving problem FPHP : $\{0,1\}^n \mapsto \{0,1\}^{n+1}$, defined for all $x \in \{0,1\}^n$ and $|y| \in \{0,1\}^n$ as

$$\begin{cases} |x|/2 \equiv |y| \; mod\, 2, & |x| \mod 2 = 0 \\ |x-1|/2 \equiv |y| \; mod\, 2, & |x| \mod 2 = 1 \end{cases} . \tag{3.2.2}$$

We translate this into a decision problem by taking the parity of the output string $|y|$ of this relation (Figure 7). Interestingly, using a sub-circuit with a quantum advantage respective to the local relation problem does not imply an advantage for the equivalent decision problem. Only the inverse implication would be unconditionally true.

Another related work concerning shallow quantum circuits is that of [186]. Here constant-depth quantum circuits that solve a sampling problem related to the class of symmetric Boolean functions we have explored are proposed. In addition, an advantage to the best constant depth circuits has been identified for these problems.

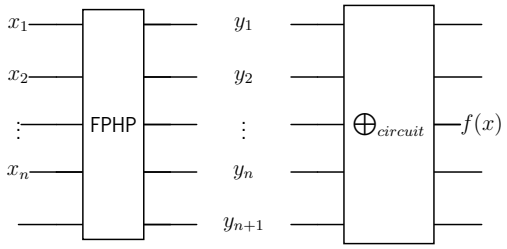


Figure 7: Circuit for evaluating the symmetric Boolean functions with degree two built from the FPHP relation by the composition of a classical circuit that computes the parity.

From another perspective, the quantum advantage for the decision problem we identified is described in terms of the number of gates and the width of the circuits. Various related results have been obtained for circuit width. For instance, in [187], the authors prove a quantum advantage using the same symmetric Boolean function of degree two that we adress here. The advantage regarding the equivalent classical model is proven by fixing a width size for the computational space and showing that these computations can be computed with greater efficiency in an actual quantum device, IBM's quantum computer. Therefore, our result can be interpreted as the computation of the same function without any width restriction but with a temporal restriction on the time ordering for the quantum circuit (all the measurements have to commute).

Finally, quantum width advantages for a class of symmetric Boolean functions in the setting of branching programs were demonstrated in [188]. In particular, these functions separate optimally quantum and classical branching programs and are generalizations of the symmetric Boolean function we used in the proof of Theorem 3.2.2.

3.2.3 Chapter organization

The remainder of the chapter is organized as follows. Section 3.3 provides the background for the non-adaptive measurement-based quantum computation model. Section 3.4 introduces our new constructions to select the measurement instructions for Boolean function evaluation. Subsequently, Section 3.5 analyses theoretical resource bounds for the model, providing the technical proof of Theorem 3.2.1 and support for Conjecture 1. Section 3.6 translates the model into realizable computations in the form of circuits, providing the technical proof of Theorem 3.2.2 and additional support for Conjecture 1. Finally, section 3.7 concludes with a brief discussion of the results.

3.3 Background

3.3.1 The non-adaptive measurement-based quantum computation model

The definition of the non-adaptive MBQC (NMQC$_{\oplus}$) model builds upon that of the MBQC model [102]. In general terms, the MBQC model is a universal quantum computing model which performs computations by a sequence of measurements on a highly entangled quantum state [152, 153]. The measurement operators applied during the computation must be selected adaptively, depending on previous measurement outcomes. The NMQC$_{\oplus}$ model differs from the general MBQC in that it does not allow for measurement adaptivity. Therefore, it cannot select subsequent measurement operators assuming knowledge about the outcome of previous measurements. Without the capacity to adapt, the order of single-qubit measurements performed on the resource state is not defined, and the temporal structure becomes flat [157, 189]. Furthermore, in the usual MBQC model [152, 153], each measurement is chosen using a limited classical side computer capable only of linear computations. Thus, in NMQC$_{\oplus}$ this linear side computer can immediately determine all the measurement bases required for the computation based on the input string. Then, the selected measurement operators are applied to the resource, producing a set of outcome values. These outcomes are fed again into a linear side computer, which delivers the final output of the entire computational process. A detailed description of the model can be given as a procedure consisting of the following three main stages:

Stage 1. A classical input string $x \in \{0,1\}^n$ is pre-processed by a linear side computer generating a binary output $s_i \in \{0,1\}$ for each qubit i of the resource state, with k qubits. Without loss of generality, this computation for each qubit is given by the general linear expression,

$$s_i = L_i(x) = a_{i_0} \oplus \bigoplus_{j \in S_i} \left(a_{i_j} \oplus x_j\right) \tag{3.3.1}$$

for all $i \in \{0,1,...,k-1\}$, with $a_{i_j} \in \{0,1\}$, and $S_i \subseteq [n]$.

Stage 2. To each bit-value of $s_i \in \{0,1\}$ we associate a single-qubit measurement operator on the equator of the Bloch sphere as,

$$M_i(s_i) = \cos(\theta_i + \phi_i * s_i)\sigma_x + \sin(\theta_i + \phi_i * s_i)\sigma_y \tag{3.3.2}$$

for all $i \in \{0,1,...,k-1\}$, with σ_x and σ_y standing for the Pauli X and Y operator, respectively, and angles θ_i and ϕ_i defining the specific measurement operators. In the measurement of,

$$M_0(s_0) \otimes M_1(s_1) \otimes ... \otimes M_i(s_i) \otimes ... \otimes M_{k-1}(s_{k-1}) \ , \tag{3.3.3}$$

on resource state $|\Psi_k\rangle$, each one of these measurements yields a binary output $m_i \in \{-1, 1\}$, with an associated probability for all $i \in \{0, 1, ..., k-1\}$. The result of this process is a string x_m of symbols ± 1 with length k composed of eigenvalues of the measurement operators $\otimes_{i=0}^{k-1} M_i(s_i)$,

$$x_m = m_0, m_1, \ldots, m_i, \ldots, m_{k-1} \text{ with } m_i \in \{-1, 1\} . \tag{3.3.4}$$

Stage 3. In the final stage, the auxiliary linear side processor is employed again to compute the binary output based on the previously obtained string x_m,

$$L_f(x_m) = a_0 \oplus \bigoplus_{i=1}^{k} \big(a_i \oplus g(m_i)\big) \tag{3.3.5}$$

where $g(x) = \dfrac{1-x}{2}$ acts as a renormalization function to map the ± 1 outcomes to the more usual bit values 0 and 1. The complete process is illustrated in Figure 8 (a), summarizing the different stages of the NMQC$_\oplus$ model.

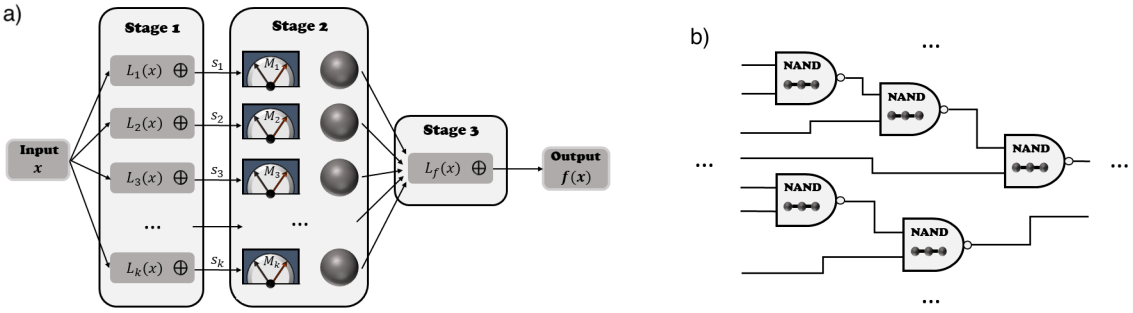


Figure 8: (a) The complete NMQC$_\oplus$ procedure. The input string $x \in \{0, 1\}^n$ is shared with each linear function of the pre-processing stage. Then the binary values s_i are forwarded to the measurement process, determining a specific measurement operator on all qubits. In the end, the measurement results are fed into a linear function to compute the result of the NMQC$_\oplus$ computation. (b) Representation of a digital circuit composed of NAND gates that are implemented using individual NMQC$_\oplus$ computations.

3.3.2 Measurement assignments

The lack of adaptivity prevents NMQC$_\oplus$ from being a universal model for quantum computation, as MBQC is. Nevertheless, Anders and Browne demonstrated that it could evaluate the NAND gate deterministically with a 3-qubit GHZ state and measurements of its stabilizer group. This particular gate is special because it describes a complete gate set for a universal classical computer. Consequently, any Boolean function can be computed by composing layers of NMQC$_\oplus$ evaluations (Figure 8 (b)). However, by composing different NMQC$_\oplus$ computations, the whole computational process becomes adaptive. This implication can be easily understood as the result of a specific NAND gate will define the measurement operators of the NAND gates to its outputs are fed into. Therefore, we will focus mainly on the computation capabilities of the NMQC$_\oplus$ model without any compositions.

Subsequent to the work of Anders and Browne, which introduced the main ideas, the first detailed description of the NMQC$_\oplus$ model was presented by Hoban *et al.* in [102]. This work examines the set of instructions necessary for each stage of the model, the measurement assignments, in order to deterministically evaluate Boolean functions. Formally,

Definition 1. (Measurement assignment). *A measurement assignment for a* NMQC$_\oplus$ *computation*

is composed by a set of linear functions L_i, a set of dichotomic measurement operators $M_i(s_i)$, for all $i \in \{0, 1, ..., k-1\}$, a final linear function L_f and a resource state Ψ_k, which takes the form of a k-qubit quantum state.

In the same work, the authors proved that there is a measurement assignment for any Boolean function, $f : \{0,1\}^n \mapsto \{0,1\}$, which deterministically evaluates the function in the NMQC$_\oplus$ model. These measurement assignments require general GHZ states as a resource state [190, 191]. Moreover, such highly entangled quantum states were proven to be the most efficient resource state for these computations, concerning the number of required qubits [192, 102], as stated in the following theorem.

Theorem 3.3.1. *[102] In* NMQC$_\oplus$ *with quantum resources, any Boolean function $f(x)$, on an arbitrary n-bit string x, can be computed with unit probability. This can be achieved with the resource state $\frac{1}{\sqrt{2}}\big(|0\rangle^{\otimes(2^n-1)} + |1\rangle^{\otimes(2^n-1)}\big)$ i.e. a (2^n-1) qubit generalised GHZ state.*

This result proves that classical universality can be achieved within the NMQC$_\oplus$ model, without the need to compose various layers of NMQC$_\oplus$ computations. Furthermore, it motivates the following research question

"What are the measurement assignments for an arbitrary Boolean function?"

Hoban *et al.* state that their proof for Theorem 3.3.1 also provides a constructive method to discover valid measurement assignments. We argue that this process is computationally too expensive, as it searches over all possible combinations of linear functions for the pre-processing stage and over a continuum for all the angles of the measurement stage. This assertion will be supported by the complexity analysis of the constructions proposed in [164, 1], and which are much more computationally efficient than the original proposal in [102].
In the remainder of this section, we will review two different constructions that can be used to derive specific measurement assignments for a Boolean function. The two methods were introduced in [164], while the second construction was reintroduced independently in [1].

Construction 1: The Fourier representation ($\mathcal{FR}$) ([164]) The first construction merges two ideas to derive a method that yields a measurement assignment for any Boolean function, $f : \{0,1\}^n \mapsto \{0,1\}$. It starts by verifying that the expected value of a measurement operator composed of single-qubit measurements on a general GHZ state, Ψ_{GHZ}, can be described with a cosine function,

$$\big\langle \Psi^k_{GHZ} \big| \otimes_{i=1}^k M_i(s_i) \big| \Psi^k_{GHZ} \big\rangle = \cos\big(\sum_{i=1}^k \theta_i + \phi_i s_i\big) = \cos\big(\underbrace{\sum_{i=1}^k \theta_i + \phi_i L_i(x)}_{\text{poly}_\text{f}(\text{x})}\big)\,, \tag{3.3.6}$$

whose assignment in a multi-linear polynomial, poly$_\text{f} : \{0,1\}^n \mapsto \pi(\mathbb{R})$ [192] [2].

Definition 2. (Multi-linear polynomial). *A multi-linear polynomial is a function* poly$_f : \{1,-1\}^n \mapsto \mathbb{C}$ *that can be written as,*

$$\text{poly}_f(x) = \sum_{S \subseteq [n]} c^*_S \prod_{i \in S} x_i \tag{3.3.7}$$

where each c^*_S is a complex number.

[2] We use the notation poly$_\text{f}$ throughout the text to represent either polynomial of the form poly$_\text{f} : \{-1,1\}^n \mapsto \mathbb{C}$ or poly$_\text{f} : \{0,1\}^n \mapsto \mathbb{C}$. Both of these mathematical objects serve as representations for a Boolean function of the form $f : \{0,1\}^n \mapsto \{0,1\}$, and the type of poly$_\text{f}$ used will depend on the convenience given the context where it is used.

This polynomial has real coefficients θ_i and ϕ_i, which relate to the parameters of the measurement operators from Equation (3.3.2), and the binary variables s_i related to the binary values from Equation (3.3.1) that select the measurement bases. Thus, finding measurement assignments that compute a specific Boolean function is equivalent to finding the corresponding multi-linear polynomials. In particular, a valid polynomial has to map each input $x \in \{0,1\}^n$ to a multiple of π that is congruent modulo two to the output of the Boolean function. Then, the total measurement operator maps to the eigenvalue corresponding to the output of the Boolean function, for all $f(x) : \{0,1\}^n \mapsto \{0,1\}$ and $x \in \{0,1\}^n$,

$$\left\langle \Psi^k_{GHZ} \right| \otimes_{i=1}^k M_i(L_i(x)) \left| \Psi^k_{GHZ} \right\rangle = \cos\Big(\underbrace{\sum_{i=1}^k \theta_i + \phi_i L_i(x)}_{\mathsf{poly_f(x)}} \Big) = (-1)^{f(x)} \ . \tag{3.3.8}$$

The existence of such a multi-linear polynomial is guaranteed for every semi-Boolean function ($f_* : \{0,1\}^n \mapsto \mathbb{R}$) [193]. A specific instance can be computed through the Walsh-Hadamard transform $\mathcal{F}$, which is the Fourier transform on the Boolean group $(\mathbb{Z}/2\mathbb{Z})^n$ [3]. This transformation yields the Fourier coefficients ($\widehat{f}(S)$) given an ordered vector with all the possible input strings evaluated by the Boolean function,

$$\frac{1}{2^n}\mathcal{F}^T \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{2^n-1}) \\ f(x_{2^n}) \end{pmatrix} = \begin{bmatrix} \widehat{f}(S_0) \\ \widehat{f}(S_1) \\ \vdots \\ \widehat{f}(S_{2^n-1}) \\ \widehat{f}(S_{2^n}) \end{bmatrix} . \tag{3.3.9}$$

These are the coefficients of the corresponding multi-linear polynomial, $c^*_S = \widehat{f}(S)$. Therefore, for all $f : \{0,1\}^n \mapsto \{0,1\}$ and $x \in \{0,1\}^n$,

$$\mathsf{poly_f}(x) = \pi * \Big(\sum_{S \subseteq [n]} \widehat{f}(S) * g^{-1}\Big(\bigoplus_{i \in S} x_i \Big) \Big) = \pi * f(x) \ , \tag{3.3.10}$$

with $g^{-1}(x) = 1 - 2x$.
The multi-linear polynomial $\mathsf{poly_f}$ obtained through this process can be used to construct a measurement assignment that computes the Boolean function deterministically in the NMQC$_\oplus$ model. The resource state will be a general GHZ state of k qubits,

$$\left| \Psi^k_{GHZ} \right\rangle = \frac{1}{\sqrt{2}} \big(|0\rangle^{\otimes k} + |1\rangle^{\otimes k} \big) \ , \tag{3.3.11}$$

with k standing for the sparsity of the multi-linear polynomial $\mathsf{poly_f}$, defined as follows.

Definition 3. (Sparsity). *The sparsity of a multi-linear polynomial* $\mathsf{poly_f} : \{0,1\}^n \mapsto \mathbb{R}$ *is designated as the number of non-zero Fourier coefficients,*

$$\mathsf{sparsity}(\mathsf{poly_f}) = \mathsf{sparsity}\big(\mathcal{F}(f)\big) = \sum_{S \subseteq [n] \setminus [0]} \delta\big(\widehat{f}(S)\big), \ \ \delta(x) = \begin{cases} 1 & x \neq 0 \\ 0 & otherwise \end{cases} \ , \tag{3.3.12}$$

which results from applying the Fourier transform $\mathcal{F}$ to the semi-Boolean function, $f_* : \{0,1\}^n \mapsto \mathbb{R}$, such that, for all $x \in \{0,1\}$, $f_*(x) = \mathsf{poly}_{f_*}(x)$.

[3] A description for the Walsh-Hadamard transform is given in Appendix A.1.

In the pre-processing stage (Stage 1), the linear Boolean functions which select the measurement operators, are all the parity bases ($\bigoplus_{l\subseteq S} x_l$), corresponding to non-zero Fourier coefficient ($\widehat{f}(S_i) \neq 0$) of poly_f,

$$s_i = L_i(x) = \bigoplus_{l\subseteq S_i} x_l \ , \tag{3.3.13}$$

with $i \in \{0, 1, ..., k-1\}$. The angles of the measurement operators, from Equation (3.3.2) in Stage 2, are built from the non-zero Fourier coefficients that constructed the multi-linear polynomial resorting to Equation (3.3.10), such that, for all $i \in \{0, 1, ..., k-1\}$,

$$\theta_i = \frac{\pi}{k} * \sum_{S\subseteq[n]} \widehat{f}(S) \ . \tag{3.3.14}$$

Notice that each θ_i always has the same value for all measurement operators. Only the values of ϕ_i change, and these angles are defined by the corresponding non-zero coefficients of the polynomial,

$$\phi_i = -2\pi * \widehat{f}(S_i) \ , \tag{3.3.15}$$

for all $i \in \{0, 1, ..., k-1\}$. Therefore, each non-zero Fourier coefficient $\widehat{f}(S_i)$ links a specific parity basis $L_i(x) = \bigoplus_{l\in S_i} x_l$ to the corresponding measurement operator $M_i(s_i)$, which will be applied to one of the qubits of the resource state.
Finally, the linear Boolean function of the post-processing stage, Stage 3, will be the linear sum of all the measurement outcomes (m_i),

$$L_f(x_m) = \bigoplus_{i=1}^{k} g(m_i) \ , \tag{3.3.16}$$

for all Boolean functions, with a measurement assignment determined directly by the equivalent multi-linear polynomial.

Example 1. *For the* AND *function with input strings of size two,* $\text{AND}(x_1, x_2) = x_1 * x_2$*, the discrete Fourier transform applied to the corresponding truth table computes the Fourier coefficients,*

$$\frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ -1/4 \\ -1/4 \\ 1/4 \end{bmatrix} \ . \tag{3.3.17}$$

These are used to build the corresponding multi-linear polynomial resorting to Equation (3.3.10)*,*

$$\text{poly}_{\text{AND}}(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1 \oplus x_2 \ . \tag{3.3.18}$$

Then, the coefficients of the polynomial are used to create the measurement assignment. In this case, the polynomial translates to a 3-qubit GHZ state, $|\Psi^3_{GHZ}\rangle = 1/\sqrt{2}\big(\,|0\rangle^{\otimes 3} + |1\rangle^{\otimes 3}\,\big)$*, the linear functions*

$$L_1(x) = x_1, \ L_2(x) = x_2, \ and \ L_3(x) = x_1 \oplus x_2 \ , \tag{3.3.19}$$

and the corresponding measurement operators

$$M_1(s_1) = (\neg s_1)\sigma_x + s_1\sigma_y, \ M_2(s_2) = (\neg s_2)\sigma_x + s_2\sigma_y \ and \ M_3(s_3) = (\neg s_3)\sigma_x - s_3\sigma_y \ . \tag{3.3.20}$$

Moreover, the AND *function evaluated on all possible inputs, for* $x \in \{0, 1\}^2$*, in the* $\text{NMQC}_{\oplus}$ *model is*

equal to,

$$\sigma_x \otimes \sigma_x \otimes \ \ \sigma_x \left|\Psi^3_{GHZ}\right\rangle = + \left|\Psi^3_{GHZ}\right\rangle,\ \sigma_x \otimes \sigma_y \otimes -\sigma_y \left|\Psi^3_{GHZ}\right\rangle = + \left|\Psi^3_{GHZ}\right\rangle$$
$$\sigma_y \otimes \sigma_x \otimes -\sigma_y \left|\Psi^3_{GHZ}\right\rangle = + \left|\Psi^3_{GHZ}\right\rangle,\ \sigma_y \otimes \sigma_y \otimes \ \ \sigma_x \left|\Psi^3_{GHZ}\right\rangle = - \left|\Psi^3_{GHZ}\right\rangle\ .$$

Yielding the correct correspondence between the resulting eigenvalues and the input strings, such that the final state is equal to $(-1)^{\mathsf{AND}(x_1,x_2)} \left|\Psi^3_{GHZ}\right\rangle$.

This first construction, taken from [164] was presented with a detailed translation from a valid multi-linear polynomial, for a specific Boolean function, to a correct and deterministic measurement assignment. The construction is characterized by a process that finds a multi-linear polynomial, which in this case is the Walsh-Hadamard transform. The subsequent constructions differ precisely in the process of finding an equivalent multi-linear polynomial for the function. Thus, it is this process alone that distinguishes between them, as the conversion from the multi-linear polynomial to the measurement assignment will be always the same in every case. Therefore, this process will not be repeated for the next constructions.

Construction 2: The extended Fourier representation ($\mathcal{EF}$) ([164, 1]) The second construction presented in [164], equivalent to the construction introduced in [1], applies the discrete Fourier transform using a different representation of the Boolean function, the Algebraic Normal Form (ANF).

Definition 4. (Algebraic Normal Form). *Any Boolean function* $f : \{0,1\}^n \mapsto \{0,1\}$ *can be expressed uniquely as,*

$$f(x) = \bigoplus_{S\subseteq[n]} c_S \prod_{i\in S} x_i \tag{3.3.21}$$

with $c_S \in \{0,1\}$.

Given the ANF of a Boolean function, this construction applies the Walsh-Hadamard transform to each monomial ($\prod_{i\in S} x_i$). The resulting multi-linear polynomials are summed to assemble a single multi-linear polynomial representing the desired Boolean function. This process constructs a correct multi-linear polynomial and the corresponding correct measurement assignment due to the additivity property of the Walsh-Hadamard transform,

$$\mathsf{poly_f}(x) = \pi * \frac{1}{2^n}\mathcal{F}^T\left(\bigoplus_{S\subseteq[n]} c_S \prod_{i\in S} x_i\right) \tag{3.3.22}$$

$$\equiv \pi * \sum_{S\subseteq[n]} c_S * \frac{1}{2^n}\mathcal{F}^T\left(\prod_{i\in S} x_i\right). \tag{3.3.23}$$

In summary, we described two constructions that produce correct measurement assignments for any Boolean function. The measurement assignments obtained from them may differ. In particular, the sparsity of the multi-linear polynomials is the main difference and one of the most significant properties related to the efficiency of the NMQC$_\oplus$ model. Generally, the $\mathcal{EF}$ construction produces sparser polynomials than the $\mathcal{FR}$ construction. This separation can be exponentially large for some families of Boolean functions [4] [164, 194] . Thus, asymptotically the $\mathcal{EF}$ construction is identified as the construction that produces the most efficient measurement assignments.

[4] We designate as families of Boolean functions a set of Boolean functions with a precise description and a single instance for each input string size. The general AND is an example of a family of Boolean functions.

3.4 Evaluation of Boolean functions in the $\mathsf{NMQC}_{\oplus}$ model

This section presents new constructions for the deterministic evaluation of Boolean function in the NMQC$\oplus$ model. First, we introduce a new construction for symmetric Boolean functions (SBF). Then, we demonstrate how this new method can be integrated into another construction, yielding more efficient NMQC$\oplus$ computations for arbitrary Boolean functions.

3.4.1 Symmetric Boolean functions

In Boolean functions analysis [193], the set of symmetric Boolean functions i.e., those which are invariant under permutations of the input strings, was studied for various properties and applications in computer science [195, 196, 197, 198]. As a result of this property, the ANF representation has the following form, for all $x \in \{0,1\}^n$,

$$f(x) = c_0 \oplus c_1 * C^1(x) \oplus c_2 * C^2(x) \oplus c_3 * C^3(x) \oplus ... \oplus c_n * C^n(x) = \bigoplus_{k=0}^{n} c_k * C^k(x), \quad (3.4.1)$$

where each $c_k \in \{0,1\}$ and the $C^k(x)$ terms represent the complete symmetric functions (CSF), defined as

Definition 5. (Complete symmetric function). *The complete symmetric function (CSF) of dimension k, $C^k(x) : \{0,1\}^n \mapsto \{0,1\}$ is defined as the sum, modulo 2, of all terms with degree equal to k, i.e.*

$$C^k(x) = \bigoplus_{i_1=1}^{n-k+1} x_{i_1} \Big(\bigoplus_{i_2=i_1+1}^{n-k+2} x_{i_2} \big(\ ... \ \big(\bigoplus_{i_N=i_{k-1}+1}^{n} x_{i_N} \big) \big) \Big) . \quad (3.4.2)$$

These functions are the building blocks used to generate the entire set of symmetric Boolean functions [199], according to Equation (3.4.1). Therefore, a construction that renders multi-linear polynomials for the CSF functions will be universal for the set of symmetric Boolean functions.

Construction 3: CSF Representation ($\mathcal{CSF}$) We propose a new construction that builds a multi-linear polynomial for an arbitrary symmetric Boolean function. This construction takes advantage of the decomposition established in the following lemma.

Lemma 1. *All complete symmetric functions C^k, for an arbitrary dimension k, can be decomposed as the product of complete symmetric functions with a dimension equal to power of two, i.e. for all $x \in \{0,1\}^n$*

$$C^k(x) = \bigwedge_{r \in R_k} C^{2^r}(x), \quad (3.4.3)$$

with $R_k \subseteq [\lceil \log(k) \rceil]$, such that $\sum_{r \in R_k} 2^r = k$.

Proof. We obtain the following lemma for single conjunction.

Lemma 1.1. *The composition of two complete symmetric functions with degrees p and q equals the complete symmetric function with degree $p+q$,*

$$C^p \wedge C^q = C^{p+q} \quad (3.4.4)$$

if q is a power of two and p is a number composed of powers two larger than q, i.e. $p = \sum_{i \in N} 2^{n_i}$, $q = 2^m, \forall i \in N \ (n_i > m)$

Proof. We proceed by strong induction.

Base case:

$$C^p \wedge C^1 = C^{p+1}, with\ p = \sum_{i \in N} 2^{n_i}, \forall_{i \in N}\ (n_i > 0) \tag{3.4.5}$$

To prove this equality, we resort to its ANFs,

$$\underbrace{\Big(\bigoplus_{1 \leq i_1 < i_2 < \ldots \leq k} x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_p} \Big)}_{Exp_1} \wedge \underbrace{\bigoplus_{i=0}^{k} x_i}_{Exp_2} = \bigoplus_{1 \leq i_1 < i_2 < \ldots \leq k} x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_{p+1}} \tag{3.4.6}$$

and apply the AND operator to verify that LHS equals the RHS. However, the expression will contain terms that are not only simple linear sums of AND operators composed of single bits. The following rule will reduce the terms with linear sums inside the AND operators possible

$$(a \oplus b) \wedge c = (a \wedge c) \oplus (b \wedge c) \tag{R1}$$

By applying repeatedly Rule R1 to the LHS of Equation (3.4.6) we obtain an expression with a term for each element of Exp_2 composed with Exp_1

$$C^p \wedge C^1 = \Big(\underbrace{\Big(\bigoplus_{1 \leq i_1 < i_2 < \ldots \leq k} x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_p} \Big)}_{Exp1} \wedge x_1 \Big)$$
$$\oplus \Big(\underbrace{\Big(\bigoplus_{1 \leq i_1 < i_2 < \ldots \leq k} x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_p} \Big)}_{Exp1} \wedge x_2 \Big) \oplus \cdots \oplus \Big(\underbrace{\Big(\bigoplus_{1 \leq i_1 < i_2 < \ldots \leq k} x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_p} \Big)}_{Exp1} \wedge x_k \Big) \tag{3.4.7}$$

The same rule is applied to each of the terms in Exp_1 with the single term of Exp_2. At the end, each term of Exp_1 is composed with all elements contained in Exp_2

$$\begin{aligned} &C^p(x) \wedge C^1(x) = \\ &\quad \mathbf{x_1} \wedge x_2 \wedge \ldots \wedge x_p \wedge \mathbf{x_1} \ \oplus \quad x_1 \wedge \mathbf{x_2} \wedge \ldots \wedge x_p \wedge \mathbf{x_2} \quad \oplus \ldots \oplus x_1 \wedge x_2 \wedge \ldots \wedge x_p \wedge \mathbf{x_k} \oplus \\ &\mathbf{x_1} \wedge \ldots \wedge x_{p-1} \wedge x_{p+1} \wedge \mathbf{x_1} \oplus x_1 \wedge \ldots \wedge x_{p-1} \wedge x_{p+1} \wedge \mathbf{x_2} \oplus \ldots \oplus x_1 \wedge \ldots \wedge x_{p+1} \wedge \mathbf{x_k} \oplus \\ &\qquad\qquad\qquad\qquad \vdots \\ &x_{k-p} \wedge \ldots \wedge x_{k-1} \wedge x_k \wedge \mathbf{x_1} \oplus x_{k-p} \wedge \ldots \wedge x_k \wedge \mathbf{x_2} \oplus \ldots \oplus x_{k-p} \wedge \ldots \wedge x_{k-1} \wedge \mathbf{x_k} \wedge \mathbf{x_k} \end{aligned} \tag{3.4.8}$$

In order to reduce this expression to a unique ANF, it will be necessary first to remove repetitions of bits in each term. For instance, the first term of Exp_1 composed with the first term of Exp_2 repeats the x_1 bit. Afterward, it will be necessary to count the number of times each term does repeat itself. If a term occurs an odd number of times, it remains and has an exclusive existence in the ANF; if it occurs an even number of times, it disappears. In particular, there will be only terms of size p, for which the unique element introduced from Exp_2 was already part of the term, and the terms of size $p + 1$ for which this term was

new. The terms of size p occur once each time that the new bit is equal to an existing one

$$\#\big(\big|x_i \wedge x_j \wedge ... \wedge x_n\big| = p\big) = \binom{p}{1} = p \tag{3.4.9}$$

and as p is by definition an even number. This implies that no term of size p exists in its final ANF, which is consistent with the RHS of Equation (3.4.6).
Finally, to verify the equality, it has to be guaranteed that each element of RHD of the equation exists and appears an odd number of times in the LHS of (3.4.6). The first condition is verified because all terms of size p exist in Exp_1, and the product with Exp_2 produces all possible combinations of terms with size $p+1$. It remains to prove that the terms with $p+1$ arguments appear an odd number of times. For that, let us consider a unique term,

$$x_1 \wedge x_2 \wedge ... \wedge x_p \wedge \mathbf{x_i},\ i \notin \{1, 2, ..., p\} \tag{3.4.10}$$

with $p+1$ different indexes can occur repeatedly due to the composition of another product term that initially shared p indexes and was composed with the missing bit. Therefore each term with $p+1$ different bits is created as

$$\#\big(\big|x_i \wedge x_j \wedge ... \wedge x_n\big| = p+1\big) = \binom{p+1}{p} = p+1 \tag{3.4.11}$$

which is an odd number by definition, proving the base case.

Induction step:

$$\forall_{y<m+1}\, C^p \wedge C^{2^y} = C^{p+2^y} \implies C^p \wedge C^{2^{m+1}} = C^{p+2^{m+1}},\ with\, p = \sum_{i \in N} 2^{n_i},\ \forall_{i \in N}\, m+1 > n_i \tag{3.4.12}$$

To prove that the RHS of the consequent equality is equal to its LHS, we rewrite all these functions in the ANF. The procedure is the one used in the base case resorting to Rule R1 repeatedly. Subsequently, it is necessary to count the number of times this process creates each term to verify its inclusion in the ANF. However, this process is more complex for these cases than for the base case because there are significantly more repetitions, and the decomposition generates product terms with a mixture of lengths. Once more, we resort to the symmetry argument to analyze only one product term for each size. Therefore, we obtain potential terms from the size 2^{m+1} to $2^{m+1}+p$, given that they could repeat from p to 0 times in the expression. Furthermore, the number of times each of these terms repeats is equal to

$$\#\big(\big|x_{i_1} \wedge x_{i_2} \wedge ... \wedge x_{i_p}\big| = 2^{m+1}\big) = \binom{2^{m+1}}{p}\ repetitions \tag{3.4.13}$$

$$\#\big(\big|x_{i_1} \wedge x_{i_2} \wedge ... \wedge x_{i_p} \wedge \mathbf{x_{j_1}}\big| = 2^{m+1}+1\big) = \binom{2^{m+1}}{p-1} \cdot \binom{1}{1}\ repetitions \tag{3.4.14}$$

$$\vdots$$

$$\#\big(\big|x_{i_1} \wedge ... \wedge x_{i_p} \wedge \mathbf{x_{j_1}} \wedge ... \wedge \mathbf{x_{j_{q-1}}}\big| = 2^{m+1}+p-1\big) = \binom{2^{m+1}}{1} \cdot \binom{p-1}{p-1}\ repetitions \tag{3.4.15}$$

$$\#\big(\big|x_{i_1} \wedge x_{i_2} \wedge ... \wedge x_{i_p} \wedge \mathbf{x_{j_1}} \wedge ... \wedge \mathbf{x_{j_q}}\big| = 2^{m+1}+p\big) = \binom{2^{m+1}}{0} \cdot \binom{p}{0}\ repetitions. \tag{3.4.16}$$

(3.4.17)

It is easy to verify that terms of size $2^{m+1}+p$ exist only once. However, to verify the equality, it is necessary to prove that all other terms of smaller dimensions exist an even number of times. In order to achieve that, it helps to write the modular division by two of these number of combinations with two additional description presented in [164],

$$C^{2^r}(x) = \binom{|x|}{2^r} \, mod \, 2 = Bit^r\Big(|x|\Big) \tag{3.4.18}$$

showing that the CSF with a degree equal to a base of two has two equivalent representations. One as a modular division by two of a combinatorial term involving the Hamming weight of the input string, and another as a specific bit r of the binary representation of the Hamming weight $|x|$ which will be designated $Bit^r(|x|)$. With these descriptions and the hypothesis for the induction step, we obtain that

$$\binom{2^{m+1}}{p} \, mod \, 2 = \bigwedge_{i \in N} Bit^{n_i}(2^{m+1}) = 0 \tag{3.4.19}$$

with $p = \sum_{i \in N} 2^{n_i}$ and as the binary representation of 2^{m+1} has only the $(m+1)$th bit equal to 1, and all other bits and combinations of bits of the binary representation, smaller than $m+1$, are all equal to 0. This proves that the expression in Equation (3.4.19) is always zero by the definition of the number p, and therefore the terms of length 2^{m+1} vanish. The same argument can be used for all terms of length $2^{m+1}-1$ to $2^{m+1}+p-1$, as the the modular division by two of their number of repetitions will be always an even number. This statement holds because

$$\binom{2^{m+1}}{\{1,2,...,p\}} \, mod \, 2 = \bigwedge_{i \in S \subseteq N} Bit^{n_i}(2^{m+1}) = 0 \tag{3.4.20}$$

which concludes the proof that all terms of length smaller then $2^{m+1}+p$ vanish, verifying the induction step. □

Finally to prove Lemma 1, it is only necessary to obtain the binary representation of k and, with Lemma 1.1 compose all the base of two components of k sequentially in increasing size order.

$$\Big(\Big(\big(C^{2^{i_1}} \wedge C^{2^{i_2}}\big) \wedge C^{2^{i_3}}\Big)... \wedge C^{2^{i_n}}\Big) = C^{2^{i_1}+2^{i_2}+2^{i_3}+...+2^{i_n}} = C^k, \; with \; i_1 < i_2 < ... < i_n \tag{3.4.21}$$

obtaining exactly the decomposition stated by Lemma 1. □

This decomposition demonstrates that a subset of CSFs is sufficient to build the entire set of symmetric Boolean functions using the logical operators AND ($\wedge$) and XOR ($\oplus$).

Proposition 1. *For any symmetric Boolean function f, there is an equivalent multi-linear polynomial* poly_f *that can be built solely by composing multi-linear polynomials from the subset of complete symmetric functions which have a degree equal to a power of two, such that, for all $x \in \{-1,1\}^n$*

$$\text{poly}_f(x) = \sum_{i=0}^{n} c_i \Big(\prod_{r \in R_i} \mathcal{GC}\Big(C^{2^r}(x)\Big)\Big), \tag{3.4.22}$$

with $R_i \subseteq [\lceil \log(i) \rceil]$, such that $\sum_{r \in R_i} 2^r = i$ and $\mathcal{GC}$ any general construction that generates multi-linear polynomial for CSFs.

Proof. The representation of SBF used in Equation (3.4.1) can be adapted to depend uniquely on the CSF set with a degree equal to a power of two. Recall that this translation is possible by Lemma 1. Then, any

general construction $\mathcal{GC}$ produces valid multi-linear polynomials for these two representations,

$$\text{poly}_f(x) = \mathcal{GC}\Big(\bigoplus_{i=0}^{n} c_i * C^i\Big) = \mathcal{GC}\Big(\bigoplus_{i=0}^{n} c_i\Big(\bigwedge_{r\in R_i} C^{2^r}(x)\Big)\Big), \tag{3.4.23}$$

with $R_i \subseteq [\lceil \log(i) \rceil]$, such that $\sum_{r\in R_i} 2^r = i$. Then, additivity can be used to push the application of the general construction from the entire expression to the CSFs, given that the linear sum $\oplus$ and the Real sum $+$ coincide for Boolean functions and their multi-linear polynomials,

$$\mathcal{GC}\Big(\bigoplus_{i=0}^{n} c_i\Big(\bigwedge_{r\in R_i} C^{2^r}(x)\Big)\Big) \equiv \sum_{i=0}^{n} c_i * \mathcal{GC}\Big(\bigwedge_{r\in R_i} C^{2^r}(x)\Big), \tag{3.4.24}$$

Subsequently, we will resort to a lemma from [164] that defines the translation from a conjunction operation on the function to the equivalent operation on multi-linear polynomials. In particular, this lemma shows that, for a set of Boolean function $f_1, f_2, ..., f_k$ with common input variables $x_1, x_2, ..., x_n$, their composition by logical AND $(\wedge)$ is equivalent to the product of the corresponding multi-linear polynomials,

$$\bigwedge_{j=1}^{k} f_j(x) = \prod_{j=1}^{k} poly_{f_j}(x) \ . \tag{3.4.25}$$

Consequently, a valid multi-linear polynomial can be built from the polynomials with the CSFs of dimension equal to a power of two alone,

$$\text{poly}_f(x) \equiv \sum_{i=0}^{n} c_i * \mathcal{GC}\Big(\bigwedge_{r\in R_i} C^r(x)\Big) \equiv \sum_{i=0}^{n} c_i\Big(\prod_{r\in R_i} \mathcal{GC}\Big(C^r(x)\Big)\Big). \tag{3.4.26}$$

Finally, note that the construction used to build these polynomials has only to be general in respect to any CSF with a degree equal to a power of two. □

The same decomposition holds for the equivalent multi-linear polynomials as proven by Proposition 5.5.26 using the multiplication and addition over the Reals. Thus, it describes a valid method to generate a multi-linear polynomial for an arbitrary symmetric Boolean function, defining our new construction.
This construction introduces modularity into the process of determining valid multi-linear polynomials. Consequently, to build these mathematical objects for symmetric functions, one can use any valid construction or known multi-linear polynomials, which are equivalent to the CSF in the decomposition. Therefore, the sparsest representations could be combined to attain the most efficient $\text{NMQC}_{\oplus}$ computations. For instance, if one uses the $\mathcal{EF}$ construction from section 3.3.2 to compute multi-linear polynomials for symmetric Boolean functions, the same polynomials would result from the $\mathcal{CSF}$ construction. However, suppose one uses the multi-linear polynomial presented in [102] for the C^2 function whenever it appears in the decomposition. In that case, the final multi-linear polynomials will always be sparser than any polynomials determined by either the $\mathcal{FR}$ or the $\mathcal{EF}$ constructions. This construction motivates the search for sparser multi-linear polynomials for this specific set of CSF, with a degree equal to a power of two.

CSF polynomials The $\mathcal{CSF}$ construction stresses the importance of determining efficient representations for the subset of CSF on the entire set of symmetric Boolean functions. Therefore, we identified a general form for the multi-linear polynomials corresponding to the CSFs, which have a degree equal to a power of two. We also proved that these are correct for any input size considering the CSFs with a degree lower or equal to 64, as proved in the lemma below. Therefore, to the authors' knowledge, these are the

most compact multi-linear polynomial known for those CSFs.

Lemma 2. *For $k = 2, 4, 8, 16, 32$ and 64, the general form of multi-linear polynomials represented by Equation* (3.4.27) *correctly computes the corresponding complete symmetric function $C^k : \{0,1\}^n \mapsto \{0,1\}$ for all $x \in \{0,1\}^n$, using a general GHZ state as a resource, with the* NMQC$_\oplus$ *model.*

$$\text{poly}_{C^k}(x) = \frac{\pi}{2^{k-1}}\left(\sum_{j=1}^{k/2}\binom{n-\frac{k}{2}-j}{\frac{k}{2}-j}(-1)^j\left(\sum_{S_i\subseteq[n],|S_i|=j}\bigoplus_{i\in S_i}x_i - \sum_{S_i\subseteq[n],|S_i|=n-j+1}\bigoplus_{i\in S_i}x_i\right)\right) \tag{3.4.27}$$

Proof. In order to verify that the first instances of the general form presented in Proposition 5.5.26 compute the corresponding CSF for any possible input, it helps to manipulate the general expression given by Equation (3.4.27). The goal is to rewrite this expression such that it can be expressed in function of the Hamming weight $\sum_{i=1}^{n} x_i = |x|$, which is sufficient to determine the result of the function, and the size of the input string n that is necessary to determine the specific polynomial. In particular, the polynomials are defined, for all $x \in \{0,1\}^n$, as

$$\text{poly}_{C^k}(x) = \frac{\pi}{2^{k-1}}\left(\sum_{j=1}^{k/2}\binom{n-\frac{k}{2}-j}{\frac{k}{2}-j}(-1)^j\left(\sum_{S_i\subseteq[n],|S_i|=j}\bigoplus_{i\in S_i}x_i - \sum_{S_i\subseteq[n],|S_i|=n-j+1}\bigoplus_{i\in S_i}x_i\right)\right). \tag{3.4.28}$$

For this transformation, the coefficients remain equal, but the transformed parity bases ($\bigoplus_{i\in S} x_i$) change to a different expression that evaluates to the same values based on the newly selected parameters,

$$\sum_{S_i\subseteq[n]\wedge|S_i|=j}\bigoplus_{i\in S_i}x_i = \sum_{k=0}^{\frac{2j+(-1)^j+1}{2}}\binom{|x|}{2k+1} * \binom{n-|x|}{j-2k-1}. \tag{3.4.29}$$

It will also be useful to distinguish the sum of the transformed parity bases described with sets of size $n-j+1$, which are defined as

$$\sum_{S_i\subseteq[n]\wedge|S_i|=n-j+1}\bigoplus_{i\in S_i}x_i = \begin{cases} \sum_{l=0}^{\frac{2j+(-1)^j+1}{2}}\binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-1}, & |x|\ mod\ 2 = 0 \\ -\sum_{l=0}^{\frac{2j+(-1)^j+1}{2}}\binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-2}, & |x|\ mod\ 2 = 1 \end{cases}. \tag{3.4.30}$$

Using this transformation and considering the case of $|x|\ mod\ 2 = 0$, the general expression in Equation (3.4.27) becomes, for all $|x| \in \{0, 2, 4, ..., n\}$ and $n \in \mathbb{N}$,

$$\text{poly}_{C^k}(n,|x|) = \frac{\pi}{2^{k-1}}\left(\sum_{j=0}^{k/2}\binom{n-k/2-j}{k/2-j}(-1)^j\left(\sum_{l=0}^{\frac{2j+(-1)^j+1}{2}}\binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-1} - \sum_{l=0}^{\frac{2j+(-1)^{j-1}-1}{2}}\binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-2}\right)\right). \tag{3.4.31}$$

For $|x|\ mod\ 2 = 1$ it can be rewritten for all $|x| \in \{1, 3, 5, ..., n\}$ and $n \in \mathbb{N}$, as

$$\text{poly}_{C^k}(n, |x|) = \frac{\pi}{2^{k-1}} \left(\sum_{j=0}^{k/2} \binom{n - k/2 - j}{k/2 - j} (-1)^j \left(\sum_{l=0}^{\frac{2j+(-1)^j+1}{2}} \binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-1} + \sum_{l=0}^{\frac{2j+(-1)^{j-1}-1}{2}} \binom{|x|}{2l+1} * \binom{n-|x|}{j-2l-2} \right) \right) . \tag{3.4.32}$$

Fortunately, with the use of Wolfram Mathematica code we could reduce the expressions. In particular, it was possible to show that the dependency on n disappears, and the resulting expression is equivalent to

$$\text{poly}_{C^k}(n, |x|) = \begin{cases} \frac{\pi}{2^{\frac{k}{2}} * \frac{k}{2}!} \prod_{i=0}^{\frac{k}{2}} (|x| - 2i) = \pi * \binom{\frac{|x|}{2}}{\frac{k}{2}}, & |x|\ mod\ 2 = 0 \\ \frac{\pi}{2^{\frac{k}{2}} * \frac{k}{2}!} \prod_{i=0}^{\frac{k}{2}} (|x| - 1 - 2i) = \pi * \binom{\frac{|x|-1}{2}}{\frac{k}{2}}, & |x|\ mod\ 2 = 1 \end{cases} , \tag{3.4.33}$$

with $|x| \in \mathbb{N}$, $n \in \mathbb{N}$, $k = 2, 4, 8, 16, 32$ and 64. Now it remains to show that this simpler expression computes the intended functions, and for that, we resort again the Bit^l notation

$$\text{poly}_{C^k}(n, |x|) = \begin{cases} \binom{\frac{|x|}{2}}{\frac{k}{2}}\ mod\ 2 = Bit^{log(k/2)}\left(\frac{|x|}{2}\right), & |x|\ mod\ 2 = 0 \\ \binom{\frac{|x|-1}{2}}{\frac{k}{2}}\ mod\ 2 = Bit^{log(k/2)}\left(\frac{|x|-1}{2}\right), & |x|\ mod\ 2 = 1 \end{cases} \tag{3.4.34}$$

Knowing that

$$Bit^l(x) = Bit^{l-1}(x/2) \tag{3.4.35}$$

we obtain

$$\text{poly}_{C^k}(n, h) = \begin{cases} Bit^k\big(|x|\big), & |x|\ mod\ 2 = 0 \\ Bit^k\big(|x| - 1\big), & |x|\ mod\ 2 = 1 \end{cases} \tag{3.4.36}$$

Moreover, for any $k > 1$ and $|x|\ mod\ 2 = 1$

$$Bit^k(|x| - 1) = Bit^k(|x|) \tag{3.4.37}$$

such that

$$\text{poly}_{C^k}(n, |x|) = Bit^k(|x|) = \binom{|x|}{k} mod\ 2 = C^k(n, |x|) \tag{3.4.38}$$

$$\equiv C^k(x) \tag{3.4.39}$$

as intended. □

Example 2. *The complete symmetric function of degree* 4 *is defined for all* $x \in \{0, 1\}^n$ *as*

$$C^4(x) = \bigoplus_{i_1=1}^{n-3} x_{i_1} \left(\bigoplus_{i_2=i_1+1}^{n-2} x_{i_2} \left(\bigoplus_{i_3=i_2+1}^{n-1} x_{i_3} \left(\bigoplus_{i_4=i_3+1}^{n} x_{i_4} \right) \right) \right) , \tag{3.4.40}$$

and the corresponding multi-linear polynomials have the following form

$$\mathsf{poly}_{C^4}(x) = \frac{\pi}{8}\Big((n-3) * \Big(\sum_{i=1}^{n} x_i - (\bigoplus_{j=1}^{n} x_j)\Big) + \sum_{i=1}^{n}\Big((\bigoplus_{j=1}^{n} x_j) \oplus x_i\Big) - \Big(\sum_{i=1}^{n}\sum_{j=i}^{n} x_i \oplus x_j\Big)\Big) \tag{3.4.41}$$

for all $x \in \{0,1\}^n$. Moreover, we compute their sparsity by summing the number of Fourier coefficients from different terms. Specifically, the single-bit terms contribute n coefficients, as given by $\sum_{i=1}^{n} x_i$. The parities that involve all bits except one, expressed as $\sum_{i=1}^{n}\Big((\bigoplus_{j=1}^{n} x_j) \oplus x_i\Big)$, contribute another n coefficients. Additionally, there are $\binom{n}{2}$ coefficients from the pairwise linear sums represented by $\sum_{i=1}^{n}\sum_{j=i+1}^{n} x_i \oplus x_j$. Finally, there is a single term from the parity of all the bits, given by $\bigoplus_{j=1}^{n} x_j$.

$$\mathsf{sparsity}\big(\mathsf{poly}_{C^4}(x)\big) = \binom{n}{2} + 2 * \binom{n}{1} + 1 = \frac{n^2}{2} + \frac{3n}{2} + 1 = \mathcal{O}(n^2)\,, \tag{3.4.42}$$

Therefore, this reduces the number of qubits required for the NMQC$\oplus$ *computation, compared to the previously result in [1] that scales asymptotically with $\Theta(n^3)$.*

The combination of the $\mathcal{CSF}$ construction with the CSF polynomials translates to a new set of sparse multi-linear polynomials, which are equivalent to symmetric Boolean functions. Consequently, these polynomials reduce the number of qubits required for the deterministic evaluation of symmetric Boolean functions in the NMQC$_\oplus$ model. As a concrete example, in Example 2 we show that with this construction, the C^5 function taking input strings of size $|n| = 6$ demands a 43-qubit GHZ state, whereas the previous best solution resorts to a 63-qubit GHZ state [1]. Additionally, in subsection 3.5.3 we will use the construction outlined in Proposition 5.5.26, in conjunction with the multi-linear polynomials from Lemma 2, to establish new upper bounds on the number of qubits needed for evaluating symmetric Boolean functions.

3.4.2 General Boolean functions

Could we some up with a procedure to find measurement assignments for general Boolean functions, and not just for the symmetric ones? Ideally, this procedure should find the sparsest multi-linear polynomials while requiring a small computational effort. These conditions guarantee that the classical pre-processing is not excessively time-consuming and that the NMQC$_\oplus$ computations are as efficient as possible. In this subsection, we propose a new procedure that reduces the cost of determining the measurement assignments, compared to the $\mathcal{EF}$ construction from section 3.3.2. This reduction is achieved without compromising the efficiency of the measurement assignments, given that the generated multi-linear polynomials have equal or even smaller sparsities.

Construction 4: The Krawtchouk representation ($\mathcal{KR}$) This solution uses the ANF of the Boolean functions to translate each monomial ($\prod_{i\in S} x_i$), to a multi-linear polynomial. However, in contrast to previous techniques, this translation resorts to the symmetric properties of these monomials. First, one can notice that all these terms are symmetric Boolean functions if evaluated independently. Therefore, it is possible to translate each of them to their simplified value vector [195] using a function $\nu : (\{0,1\}^n \mapsto \{0,1\}) \mapsto \{0,1\}^{n+1}$, defined as

$$\nu\Big(\prod_{i\in S} x_i\Big) = \delta_{|S|}\Big(\sum_{i\in S} x_i\Big) \tag{3.4.43}$$

$$= (\underbrace{0}_{0}, \underbrace{0}_{1}, \ldots, \underbrace{0}_{|S|-1}, \underbrace{1}_{|S|})\,, \tag{3.4.44}$$

with $|S| \subseteq [n]$ and $\delta_{|S|}$ the Kronecker delta function at $|S|$.

Definition 6. (Simplified value vector). *The simplified value vector of a symmetric Boolean function ($v_f : \mathbb{N} \mapsto \{0,1\}$) is defined as a function mapping the Hamming weights of the input string to the corresponding Boolean values*

$$v_f(|x|) = f(x), \tag{3.4.45}$$

and recall that $|x| = \sum_{i=1}^{n} x_i$ stands for the Hamming weight of x. Thus, the bit-string v_f of length $n+1$ unambiguously determines the symmetric function,

$$v_f = \{\underbrace{v_0}_{|x|=0}, \underbrace{v_1}_{|x|=1}, \underbrace{v_2}_{|x|=2}, \ \ldots\ , \underbrace{v_n}_{|x|=n}\}\ . \tag{3.4.46}$$

Given this compact representation for each monomial, one can translate such vectors to a multi-linear polynomial using the Krawtchouk transform ($\mathcal{K} : \mathbb{Z}^n \mapsto \mathbb{R}^n$), described in Appendix A.2. This transform computes the $n+1$ different Fourier coefficients of an arbitrary symmetric semi-Boolean function $f_* : \{0,1\}^n \mapsto \mathbb{R}$, such that,

$$\frac{1}{2^n}\mathcal{K}^T \begin{pmatrix} v_f(0) \\ v_f(1) \\ \vdots \\ v_f(n) \end{pmatrix} = \begin{bmatrix} \widehat{f}(|S|=0) \\ \widehat{f}(|S|=1) \\ \vdots \\ \widehat{f}(|S|=n) \end{bmatrix}\ . \tag{3.4.47}$$

In a subsequent step, the multi-linear polynomial equivalent to the monomials, which are symmetric Boolean functions, will be constructed using the computed Fourier coefficients as follows,

$$\mathsf{poly}_f(x) = \pi * \Big(\sum_{i=0}^{n} \widehat{f}(|S|=i) * \sum_{S_j \subseteq [n], |S_j|=i} g^{-1}\Big(\bigoplus_{l \subseteq S_j} x_l \Big) \Big)\ . \tag{3.4.48}$$

In the end, each of the resulting polynomials is summed, or subtracted, from the others to obtain the final multi-linear polynomial that represents the queried function,

$$\mathsf{poly_f}(x) = \pi * \sum_{S \subseteq [n]} (-1)^{t_S} * c_S * \frac{1}{2^n}\mathcal{K}^T\Big(\nu\Big(\prod_{i \in S} x_i \Big)\Big) \tag{3.4.49}$$

$$= \pm \mathsf{poly_1}(x) \pm \mathsf{poly_2}(x) \pm \ldots \pm \mathsf{poly_n}(x) \equiv f(x)\ . \tag{3.4.50}$$

with $t_S \in \{0,1\}$.
Additionally, it will be necessary to prove that valid measurement assignments are generated for the intended Boolean functions. To do so, we show that the multi-linear polynomials obtained for each monomial using Equation (3.4.48) are correct, and these can be summed or subtracted to build the final multi-linear polynomial. These two proofs are sufficient for our purpose because these are the only differences we have introduced with respect to the $\mathcal{EF}$ construction.

Lemma 3. *Any symmetric semi-Boolean function, $f_* : \{0,1\}^n \mapsto \mathbb{R}$, has at most $n+1$ different Fourier coefficients that can be computed using the Krawtchouk transformation of dimension $n+1$, with $\mathcal{O}(n^5)$ operations.*

Proof. In order to prove the lemma, we will simplify the Fourier transform sequentially with conversions that do not impact on the computation of the Fourier coefficients of symmetric semi-Boolean functions until we obtain the Krawtchouk transform.
First, we will reduce the number of lines in the transform by exploiting a symmetry between the Fourier coefficients, as stated below.

Lemma 3.1. *All symmetric semi-Boolean functions, $f_* : \{0,1\}^n \mapsto \mathbb{R}$, have equal Fourier coefficients $\widehat{f}(S)$ for all the parity bases ($\bigoplus_{l \subseteq S} x_l$) defined by sets with the same size*

$$S_i, S_j \in [n],\ |S_i| = |S_j| \implies \widehat{f}(S_i) = \widehat{f}(S_j) \tag{3.4.51}$$

Proof. The equation used to compute the Fourier coefficient of symmetric semi-Boolean function can be rewritten using the simplified value vector of the function,

$$\widehat{f}(S) = \sum_{x_i \in \{0,1\}^n} f(x_i) X_S(x_i) \tag{3.4.52}$$

$$= \sum_{i=0}^{n} v_f(i) \sum_{x \in \{0,1\}^n \wedge |x| = i} X_S(x)\ . \tag{3.4.53}$$

From this equation, it can be verified that the values of the v_f term are always equal for the same function. Therefore, it is sufficient to show that for all parity bases defined by sets $S_i, S_j \subseteq [n]$, of the same size $|S_i| = |S_j|$, the following equality holds

$$\sum_{x \in \{0,1\}^n \wedge |x| = i} X_{S_i}(x) = \sum_{x \in \{0,1\}^n \wedge |x| = i} X_{S_j}(x)\ . \tag{3.4.54}$$

This equality is very simple to verify as all the input strings that evaluate to 1, one of the parity basis will evaluate to 1 the other for a specific permutation (p). This permutation has only to exchange the bits of the input string selected by one of the sets with the bits selected by the other set. Then for all $x \in \{0,1\}^n$ and $|x| = i$,

$$X_{S_i}(x) = X_{S_j}(p(x))\ . \tag{3.4.55}$$

Furthermore, as no permutation p applied to the input strings changes the number of strings that are evaluated to 1 by the parity basis, the sum of before and after the permutation is exactly equal,

$$\sum_{x \in \{0,1\}^n \wedge |x| = i} X_{S_j}(p(x)) = \sum_{x \in \{0,1\}^n \wedge |x| = i} X_{S_j}(x)\ . \tag{3.4.56}$$

Finally, this proves that equality of Equation (3.4.54) provided one of the permutations described previously. □

This property shows that at most $n+1$ different Fourier coefficients exist, one for each possible size of the parity bases. Thus, it is sufficient to use $n+1$ lines of the Fourier transform to determine exactly those values,

$$\mathcal{F}' = \begin{bmatrix} X_{|S_i|=0}(x_1) & X_{|S_i|=0}(x_2) & X_{|S_i|=0}(x_3) & \dots & X_{|S_i|=0}(x_{2^n}) \\ X_{|S_i|=1}(x_1) & X_{|S_i|=1}(x_2) & X_{|S_i|=1}(x_3) & \dots & X_{|S_i|=1}(x_{2^n}) \\ X_{|S_i|=2}(x_1) & X_{|S_i|=2}(x_2) & X_{|S_i|=2}(x_3) & \dots & X_{|S_i|=2}(x_{2^n}) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ X_{|S_i|=n}(x_1) & X_{|S_i|=n}(x_2) & X_{|S_i|=n}(x_3) & \dots & X_{|S_i|=n}(x_{2^n}) \end{bmatrix}\ . \tag{3.4.57}$$

For the second reduction of the transformation matrix, it is necessary to prove that the Fourier coefficients can be determined based on the simplified value vector of the function v_f. Hence, we start with the equation computing Fourier coefficients using a general line of the Fourier transform and apply the necessary

transformation with some additional simplifications,

$$\widehat{f}(S) = \sum_{x_i \in \{0,1\}^n} f(x_i) X_S(x_i) \tag{3.4.58}$$

$$= \sum_{i \in \{0,1\}^n} v_f(|x_i|) * X_S(x_i) \tag{3.4.59}$$

$$= \sum_{x' \in \{0,1\}^{|S|}} \sum_{y' \in \{0,1\}^{n-|S|}} v_f(|x'| + |y'|) * X_S(x') \tag{3.4.60}$$

$$= \sum_{x' \in \{0,1\}^{|S|}} \sum_{y' \in \{0,1\}^{n-|S|}} v_f(|x'| + |y'|) * (-1)^{|x'|} * \binom{|S|}{|x'|} \tag{3.4.61}$$

$$= \sum_{i=0}^{|S|} \sum_{k=0}^{n-|S|} v_f(i+k) * \binom{n-|S|}{k} * (-1)^i * \binom{|S|}{i} . \tag{3.4.62}$$

The resulting equation is already written based on the simplified value vector and the size of the Fourier basis. However, it is necessary to isolate the term corresponding to the simplified value vector of the function so that one can define the elements of a matrix that computes the Fourier coefficients if applied to that object. For that we resort to the following change of variables $j = i + k$, yielding

$$\widehat{f}(S) = \sum_{i=0}^{|S|} \sum_{j=0}^{n} v_f(j) * (-1)^i * \binom{|S|}{i} * \binom{n-|S|}{j-i} \tag{3.4.63}$$

$$= \sum_{j=0}^{n} v_f(j) * \sum_{i=0}^{|S|} (-1)^i * \binom{|S|}{i} * \binom{n-|S|}{j-i} . \tag{3.4.64}$$

The last equation allows us to define the lines of a matrix that computes the Fourier coefficients when applied to the simplified value vector,

$$\begin{bmatrix} K_{1,1} & K_{1,2} & K_{1,3} & \dots & K_{1,n+1} \\ K_{2,1} & K_{2,2} & K_{2,3} & \dots & K_{2,n+1} \\ K_{3,1} & K_{3,2} & K_{3,3} & \dots & K_{3,n+1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ K_{n+1,1} & K_{n+1,2} & K_{n+1,3} & \dots & K_{n+1,n+1} \end{bmatrix} * \begin{bmatrix} v_f(0) \\ v_f(1) \\ v_f(2) \\ \vdots \\ v_f(n) \end{bmatrix} = \begin{bmatrix} \widehat{f}(|S| = 0) \\ \widehat{f}(|S| = 1) \\ \widehat{f}(|S| = 2) \\ \vdots \\ \widehat{f}(|S| = n) \end{bmatrix} , \tag{3.4.65}$$

with the following rule defining the entries of the matrix for the resulting transformation,

$$K_{i,j} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{n-j}{i-k} . \tag{3.4.66}$$

The rule obtained for each matrix element is precisely the rule used to build the Krawtchouk transformation, proving that one can compute the Fourier coefficient with this transformation.

In the end, to prove the bound on the number of operations, it is necessary to determine the number of operations necessary to build the matrix. In particular, each element results from the sum of combinatorial terms, which compute factorial terms. These can be exponentially bigger than n. Nevertheless, they have at most n^2 digits, which are numbers that are simple to operate with for efficient number multiplication algorithms. The latter are bounded by $\mathcal{O}\big(n\ log(n)\big)$. Therefore, the operations necessary to compute each element of the Krawtchok transform are bounded by $\mathcal{O}(n^3)$. Then, these are computed for all matrix

elements, which pushes the upper bound to $\mathcal{O}(n^5)$.

□

Thus, Lemma 3 proves that the multi-linear polynomials created for each monomial are correct. In addition, it provides proof that Construction 4 can translate each monomial in the ANF exponentially more efficiently to a multi-linear polynomial compared to the use of the Fourier transform for the same process, as employed in the previous constructions.
Finally, the fact that one can use the computed polynomials, multiplied or not by a factor of -1, relates to the simple observation that they are equivalent to the modular two operations, $\mathsf{poly}_{f_*}(x)\ mod\ 2 = -\mathsf{poly}_{f_*}(x)\ mod\ 2$. Therefore, being positive or negative does not change the parity of the angle, which determines the result of the NMQC$_\oplus$ computation (Equation (3.3.8)). Hence, one can use this new degree of freedom to reduce the number of qubits necessary to evaluate the selected Boolean function. These reductions can be achieved by attributing signs to the polynomials, $(-1)^{t_S}$, such that sets containing equal parity bases have their coefficients summing to zero. Consequently, this would remove the necessity to use these specific qubits that would otherwise be part of the resource and required to be measured. A simple demonstration of this new possibility is provided in Example 3 where we allocated the signs such that one qubit could be removed from that specific NMQC$_\oplus$ computation. Furthermore, one could now define heuristics that attempt to cancel out as many coefficients as possible. We leave, however, the search for such heuristics for future work.

Example 3. *For the function $h : \{0,1\}^3 \mapsto \{0,1\}$ with input strings of size three,*

$$h(x_1, x_2, x_3) = x_1 * x_2 \oplus x_2 * x_3\ , \tag{3.4.67}$$

the Krawtchouk transform applied to the simplified value vector to each monomial computes the Fourier coefficients,

$$\frac{1}{4}\begin{bmatrix}1 & 2 & 1\\ 1 & 0 & -1\\ 1 & -2 & 1\end{bmatrix} * \underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix}}_{\nu(x_1*x_2)} = \begin{bmatrix}1/4\\-1/4\\1/4\end{bmatrix}\ ,\ \frac{1}{4}\begin{bmatrix}1 & 2 & 1\\ 1 & 0 & -1\\ 1 & -2 & 1\end{bmatrix} * \underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix}}_{\nu(x_2*x_3)} = \begin{bmatrix}1/4\\-1/4\\1/4\end{bmatrix}\ . \tag{3.4.68}$$

These are used to build the respective multi-linear polynomials using Equation (3.4.48)*,*

$$\mathsf{poly}_{x_1x_2}(x) = \pm\Big(\frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1 \oplus x_2\Big)\ and\ \mathsf{poly}_{x_2x_3}(x) = \pm\Big(\frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_2 \oplus x_3\Big)\ . \tag{3.4.69}$$

Then, the resulting polynomials are combined to obtain the multi-linear polynomial equivalent to $h(x)$,

$$\begin{aligned}\mathsf{poly}_h(x) &= +\Big(\frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1 \oplus x_2\Big) - \Big(\frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_2 \oplus x_3\Big) && (3.4.70)\\ &= \frac{1}{2}x_1 + \frac{1}{2}\cancel{x_2} - \frac{1}{2}\cancel{x_2} - \frac{1}{2}x_3 - \frac{1}{2}x_1 \oplus x_2 + \frac{1}{2}x_2 \oplus x_3 && (3.4.71)\\ &= \frac{1}{2}x_1 - \frac{1}{2}x_3 - \frac{1}{2}x_1 \oplus x_2 + \frac{1}{2}x_2 \oplus x_3\ . && (3.4.72)\end{aligned}$$

This polynomial is then used to create the corresponding measurement assignment. In this case, it translates to a 4-qubit GHZ state, $|\Psi^4_{GHZ}\rangle = 1/\sqrt{2}\big(\,|0\rangle^{\otimes 4} + |1\rangle^{\otimes 4}\,\big)$, the linear functions

$$L_1(x) = x_1,\ L_2(x) = x_3,\ L_3(x) = x_1 \oplus x_2,\ and\ L_4(x) = x_2 \oplus x_3\ , \tag{3.4.73}$$

and the corresponding measurement operators

$$M_1(s_1) = (\neg s_1)\sigma_x + s_1\sigma_y,\ M_2(s_2) = (\neg s_2)\sigma_x - s_2\sigma_y, \tag{3.4.74}$$

$$M_3(s_3) = (\neg s_3)\sigma_x - s_3\sigma_y,\ and\ M_4(s_4) = (\neg s_4)\sigma_x + s_4\sigma_y\ . \tag{3.4.75}$$

Moreover, the function evaluated on all possible inputs, for $x \in \{0,1\}^3$, in the NMQC$_\oplus$ *model is equal to $(-1)^{h(x_1,x_2,x_3)}\,|\Psi^5_{GHZ}\rangle$ as intended. Notice also that assigned the positive sign to one of the polynomials and the negative sign to the other in Equation* (3.4.70)*. This specific attribution removed a term (x_2 in Equation* (3.4.71)*) that would otherwise demand an extra qubit of the resource state.*

Concerning the computational cost of determining the measurement assignments, this construction avoids resorting to the Fourier transform, which computes a number of Fourier coefficients that is exponential in n, the size of the input string. Furthermore, as proven in Lemma 3, there are only $n+1$ different Fourier coefficients that characterize symmetric Boolean functions. Thus, it is only necessary to select $n+1$ lines of the Fourier matrix to determine these coefficients. That is precisely one of the reductions applied to obtain the Krawtchouk transform and explains one of the computational advantages of this solution. Additionally, the transformation does not use these $n+1$ lines of the Fourier matrix to determine the coefficients. Alternatively, each line of the Krawtchouk transformation computes the demanded Fourier coefficients based on the simplified value vector of the SBF. Therefore, it reduces the determination of each Fourier coefficient to a counting problem that can be solved in polynomial time. This observation entails another exponential advantage with respect to the approach bared on the Fourier transform.
In order to quantify the advantage behind the use of the Krawtchouk transform, we need to characterize the cost of the approach that uses the Fourier transform. This cost is defined by the degree of the Boolean function.

Definition 7. (Degree). *The degree of a Boolean function f is equal to the cardinality of the largest monomial in its ANF,*

$$\deg(f) = max\{|S| \mid c_S \neq 0\} \tag{3.4.76}$$

In particular, one needs a Fourier matrix of size $2^{\deg(f)} \times 2^{\deg(f)}$ to translate the largest monomial into a multi-linear polynomial. This process enforces the existence of at least $2^{2\deg(f)}$ operations to determine the coefficient of this singular term, imposing a run-time complexity lower bound of $\Omega(2^{2\deg(f)})$ to this method. In addition, one must consider that it will be necessary to compute the multi-linear polynomials of all the other monomials ($m = \sum_{s\in[n]} c_S$) and then sum them all to form the resulting polynomial. Accounting for these additional operations, we roughly estimate the computational cost of the $\mathcal{EF}$ construction to be $\mathcal{O}\big(poly(m) * 2^{2\deg(f)}\big)$[5]. In contrast, for the Krawtchouk transform, the computational effort equivalent to determining the Fourier coefficients of the largest monomial is defined by applying a matrix of size $(\deg(f)+1) \times (\deg(f)+1)$. Furthermore, the entries of this matrix are determined in polynomial time, as proven by Lemma 3. Thus, the process used to determine all the necessary Fourier coefficients decreases from an exponential to a polynomial magnitude.
Although we have found a significant reduction in one of the subprocesses necessary to determine the measurement assignments, there is still an unavoidable exponential scaling with respect to the degree of the Boolean function.

Proposition 2. *For any Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$, a measurement assignment for the* NMQC$_\oplus$ *model can be determined with a computational complexity that scales exponentially with the degree of the Boolean function and polynomially with the number of monomials, $m = \sum_{s\in[n]} c_S$, in its ANF representation, i.e. $\mathcal{O}\Big(m * \Big((\deg(f)+1)^5 + 2^{\deg(f)}\Big)\Big)$.*

[5] The term $poly(m)$ represent a function of the type $f(x) = x^c$ with $c \in \mathbb{N}$.

Proof. The Krawtchouk transform gives the number of operations required to determine all the necessary Fourier coefficients. This matrix has $(\deg(f)+1)^2$ entries, which can be computed with at most $(\deg(f)+1)^3$ operations. Therefore, in total, there are at most $m*(\deg(f)+1)^5$ operations, as proven by Lemma 3 to determine all the necessary Fourier coefficients. These coefficients build the multi-linear polynomials equivalent to the monomials m in the ANF of the Boolean function, using Equation (3.4.48). Subsequently, there is an additional cost due to the need to sum the coefficients related to the multi-linear polynomials resulting from the various monomials. This specific process assembles the final multi-linear polynomial and has at most $m*2^{\deg(f)}$ operations. The exponential term $2^{\deg(f)}$ results from the minimal sparsity for the largest monomial in the Boolean function. That this is the minimal sparsity is proven by a theorem in [164] showing that the sparsity of any Boolean function $f:\{0,1\}^n\mapsto\{0,1\}$ is at least $2^{\deg(f)}$. Combining the two terms, we arrive at the cost value described in the proposition. □

The bound proven in Proposition 2 shows that the $\mathcal{KR}$ construction is almost as efficient as it could be to determine the measurement assignments that result from this technique. This statement is justified by demonstrating that the computational effort of determining the multi-linear polynomials regarding their sparsity is at most polynomial. Therefore, exponential scalings only exist because the object we are constructing is of exponential size. In contrast, the Fourier transform introduces an additional exponential effort to determine each one of the Fourier coefficients. Thus, determining valid measurement assignments with this transformation implies a computational effort of exponential size with respect to the sparsity of the multi-linear polynomials. Furthermore, this result highlights the importance of finding and employing constructions that produce sparse multi-linear polynomials. These reduces the number of qubits for the NMQC$_\oplus$ computation and, in principle, facilitates the process of determining their measurement assignments.

Parameterized construction 1: The symmetric complements ($\mathcal{SC}(\zeta)$) This construction merges all the contributions we have presented so far. In particular, it incorporates the advantage of the $\mathcal{CSF}$ construction to determine sparse-multi-linear polynomials for symmetric Boolean functions, with the computational reductions obtained with the $\mathcal{KR}$ construction.
The construction starts from the ANF of a Boolean function $f:\{0,1\}^n\mapsto\{0,1\}$ and creates two additional Boolean functions based on a previously selected symmetrization function ζ.

Definition 8. (Symmetrization function). *A symmetrization function is a function* $\zeta:\big(\{0,1\}^n\mapsto\{0,1\}\big)\mapsto\big(\{0,1\}^n\mapsto\{0,1\}\big)$ *that receives an arbitrary Boolean function,* $f:\{0,1\}^n\mapsto\{0,1\}$*, and maps it to a symmetric Boolean function,* $f_{\text{sym}}:\{0,1\}^n\mapsto\{0,1\}$*, such that* $\zeta(f(x))=f_{\text{sym}}(x)$*, for all* x*.*

The first Boolean function we need to define is the symmetrized function of f respective to ζ. This function is computed by applying the symmetrization function ζ to f in order to obtain f_{sym}. Then, the complement of f with respect to its symmetrized function f_{sym}, designated as $\tilde{f}$, has to be computed.

Definition 9. ($\tilde{f}(x)$). *The complement of* f *with respect to the symmetrization function* ζ *is defined as,*

$$\tilde{f}(x)=f(x)\oplus\zeta(f(x))=f(x)\oplus f_{\text{sym}}(x)\ . \tag{3.4.77}$$

The construction translates all three Boolean functions $f(x)$, $\zeta(f(x))$, and $f(x)\oplus\zeta(f(x))$ to multi-linear polynomials. The $\zeta(f)$ function is translated using the $\mathcal{CSF}$ construction, given that it is a symmetric Boolean function, while for f and $\tilde{f}$ the $\mathcal{KR}$ construction is employed to generate the following polynomials

$$f(x),\ \tilde{f}(x)\ \xmapsto{\mathcal{KR}}\ \text{poly}_{f_1}(x),\ \text{poly}_{\tilde{f}}(x);\quad \zeta(f)(x)\ \xmapsto{\mathcal{CSF}}\ \text{poly}_{\zeta(f)}(x)\ . \tag{3.4.78}$$

From the resulting multi-linear polynomials, two valid solutions can be built. The first is poly_{f_1} which derives directly from f using the $\mathcal{KR}$ construction. The second solution results from the linear sum of $\text{poly}_{\tilde{f}}(x)$ and $\text{poly}_{\zeta(f)}(x)$, provided that $\tilde{f}(x) \oplus \zeta(f)(x) = f(x)$. In the end, the sparsest of the two multi-linear polynomials will be used to generate the measurement assignment for the $\text{NMQC}_{\oplus}$ computation,

$$\text{poly}_f = min\big\{\text{sparsity}\big(\text{poly}_{f_1}\big),\ \text{sparsity}\big(\text{poly}_{\tilde{f}}(x) + \text{poly}_{\zeta(f)}(x)\big)\big\}. \tag{3.4.79}$$

In order to demonstrate how advantageous this construction could be, we will select a specific symmetrization function ζ_C. This symmetrization function takes the Boolean function f and introduces all the missing terms such that if there is a term of degree k in its ANF, then the corresponding CSF C^k will occur in its symmetrized version.

Definition 10. ($\zeta_C(x)$). *The symmetrization function $\zeta_C(x) : \big(\{0,1\}^n \mapsto \{0,1\}\big) \mapsto \big(\{0,1\}^n \mapsto \{0,1\}\big)$ is defined as*

$$\zeta_C(f(x)) = \bigoplus_{k \in T} c_k * C^k(x),\ with\ T = \{|S| \mid c_S \neq 0\}. \tag{3.4.80}$$

We can, then, use this function f to generate a multi-linear polynomial for an arbitrary Boolean function. In particular, we shall consider a class of Boolean functions for which there is an asymptotical advantage of this new constriction over all the previously described.

Definition 11. ($[aC^k]$). *The class of almost complete symmetric functions $[aC^k]$ is composed by families of Boolean functions f such that,*

$$f(x) \in [aC^k] \implies f(x) = C^k(x) \bigoplus_{i=0}^{t} \left(\prod_{j \in S_i} x_j \right) \tag{3.4.81}$$

with $|S_i| = k$ for all $i \in [0,t]$, $x \in \{0,1\}^n$ and $k, t \in \mathbb{N}$.

Function in this class miss only a fixed set of terms is missing to be precisely equal to a complete symmetric function C^k.

Lemma 4. *All families of Boolean functions in the $[aC^k]$ class have multi-linear polynomials with sparsities that scale at most with the same asymptotical rate as its symmetrized function C^k.*

Proof. Suppose the $\mathcal{SC}$ construction is used to compute the multi-linear polynomials for a family of Boolean functions f contained in the class of $[aC^k]$. In this case, these will result from combining two other multi-linear polynomials: the multi-linear polynomials corresponding to the symmetrized Boolean functions $\zeta_C(f)$ and multi-linear polynomials corresponding to the complement of f with regard to $\zeta_C(f)$, $\tilde{f}$. These particular functions will be by the definitions of $[aC^k]$ and ζ_C equal to some CSF C^k for the former, and equal to some constant number terms ($\bigoplus_{i=0}^{t}(\prod_{j \in S_i} x_j)$) for latter. Hence, the multi-linear polynomials resulting from the $\zeta_C(f)$ have sparsities equal to C^k, and the polynomials corresponding to the complement $\tilde{f}$ have a constant sparsity. Therefore, assymptotically, only the multi-linear polynomials corresponding to the symmetrized functions contribute to their sparsity. □

The first observation that follows from Lemma 4 is that for specific elements in the $[aC^k]$ class, one can take direct advantage of the sparse CSF polynomials described in Section 3.4.1 to lessen their upper bound. In Example, 4, a specific family of Boolean functions was defined for which the best qubit scaling went from $\Theta(n^{64})$ to $\Theta(n^{32})$ simply with the use of the $\mathcal{SC}$ construction with the ζ_C function and a CSF

polynomial. Furthermore, similar reductions are feasible for the whole $[aC^k]$ class. In contrast, the $\mathcal{EF}$ construction and similar techniques were utterly blind to the possibility of inheriting the sparsity of SBFs.

Example 4. *Consider function $aC^{64}(x) = C^{64}(x) \oplus x_1x_2....x_{64}$, which is built using a CSF function and a constant term added to it. If one uses the $\mathcal{EF}$ construction, the resulting multi-linear polynomials have the following form,*

$$poly_{aC^{64}(x)} = \left(\sum_{i=0}^{64} c_i \left(\sum_{S\subseteq[n],|S|=i} \bigoplus_{j\in S} x_j \right) \right) - poly_{x_1x_2...x_{64}}(x) \tag{3.4.82}$$

with a sparsity of sparsity$(poly_{aC^{64}(x)}) = \sum_{i=0}^{64} \binom{n}{i} - 1$*, that scales asymptotically with $\Theta(n^{64})$. Suppose the $\mathcal{SC}$ construction is used on the same function. In this case, it builds different multi-linear polynomials using the symmetrized function $\zeta_C(aC^{64})$, which is exactly equal to C^{64} and the complement $\widetilde{aC^{64}}$, which is equal to the term $x_1x_2...x_{64}$. Therefore the resulting multi-linear polynomials will have the following form*

$$poly_{aC^{64}(x)} = \left(\sum_{i=0}^{32} c_i \left(\sum_{S\subseteq[n],|S|=i} \bigoplus_{j\in S} x_j - \sum_{S\subseteq[n],|S|=n-i+1} \bigoplus_{j\in S} x_j \right) \right) + poly_{x_1x_2...x_{64}}(x)\ , \tag{3.4.83}$$

which has a sparsity of sparsity$(poly_{aC^{64}(x)}) = 2\sum_{i=1}^{31} \binom{n}{i} + \binom{n}{32} + \sum_{j=33}^{64} \binom{64}{j}$*, that scales asymptotically with $\Theta(n^{32})$, and which determines the number of qubits necessary to evaluate the Boolean function deterministically in the* NMQC$_\oplus$ *model.*

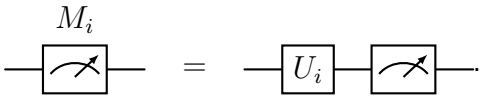
In summary, the construction proposed entails the possibility of exploring SBF to generate more efficient NMQC$_\oplus$ computations, not only for symmetric but also for general Boolean functions. From a practical point of view, we introduced a specific symmetrization function ζ_C to demonstrate the impact the construction could have on a whole class of Boolean families $[aC^k]$ by reducing the computational requirements of all their elements regarding an evaluation in the NMQC$_\oplus$ model. However, there is plenty of space for improvement in selecting and creating other symmetrization functions. One could define more elaborated functions that symmetrize the arbritary Boolean functions so that classes larger than $[aC^k]$ are reduced. Consequently, further study of this construction is in order to illustrate its potential advantages.

3.5 Bounding the resources for the NMQC$_\oplus$ model

This section discusses bounds on the resources needed for the deterministic evaluation of Boolean functions in the NMQC$_\oplus$ model. For that, we will analyze the Clifford hierarchy of the measurement operators, and the sparsity of the resource state. Both measures follow naturally from the definition of the model, which performs computations via measurements applied to a resource state with auxiliary linear computations. In particular, these measures will be examined in light of the new constructions proposed in the previous section, producing mainly new upper bounds for the NMQC$_\oplus$ model.

3.5.1 Clifford hierarchy of the measurement operators

The measurement operators required for deterministic computations in the NMQC$_\oplus$ model are described by Equation (3.3.2) as single qubit rotation around the Bloch sphere. In particular, they can be decomposed into a unitary operation U_i and a measurement on the computational basis,



Therefore, the corresponding unitaries U_i capture the complexity of the measurement operations imposed by the Boolean function in the NMQC$_\oplus$ model [6]. Intuitively, operations are considered simple if they require a small number of gates and low-depth circuits to be executed. In contrast, other more complex unitaries require more computational resources [204, 203, 205]. In order to obtain a better characterization, Gottesman and Chuang proposed the Clifford hierarchy to define specific levels of complexity for unitary operations [206].

Definition 12. (Clifford hierarchy). *The Clifford hierarchy is defined recursively as*

$$Cl_k := \{\, U \mid UPU^\dagger \in Cl_{k-1},\ \forall P \in \mathcal{P}\} \tag{3.5.1}$$

for $\mathcal{P}$ the Pauli group.

This specific hierarchy has been connected to several fundamental results. For instance, the second level of the hierarchy, Cl_2, defines the Clifford group, which was proven to be efficiently simulable by classical computers, a result known as the Gottesman-Knill theorem [207]. It is also known that any element of the third level, Cl_3, together with the Clifford group generates a universal gate set for quantum computation. For instance, the gate set Clifford+T is often used as a universal gate set, denoting by T the unitary from the third level of the hierarchy [208, 203].
For the NMQC$_\oplus$ model, it was proven that the level of the measurement operators in the Clifford hierarchy directly bounds the complexity of the Boolean functions that can be evaluated deterministically. More precisely, in [1], the authors prove that any measurement assignment with measurement operators belonging to at most the kth level of Clifford hierarchy cannot deterministically compute Boolean functions of degree greater than k. This result defines a lower bound on the hierarchy level required for deterministic computations.
Reciprocal to this finding, we determine an upper bound for the highest level of the Clifford hierarchy required for a particular Boolean function. For that, we define the granularity of semi-Boolean functions, given that this quantity determines the type of measurement operators used in a measurement assignment (Equation (3.3.15) and Equation (3.3.2)).

Definition 13. (Granularity). *The granularity of a semi-Boolean function $f_* : \{0,1\}^n \mapsto \mathbb{R}$ is given by*

$$\mathsf{gran}(f) = \max_{S\subseteq[n]} \mathsf{gran}\big(\widehat{f}(S)\big) \tag{3.5.2}$$

with the granularity of each Fourier coeficient or any rational number being equal to the smallest k such that $\widehat{f}(S) * 2^k \in \mathbb{N}$.

Then, we prove that the $\mathcal{KR}$ construction always produces measurement assignments based on semi-Boolean functions that have a granularity bounded by the degree of the Boolean function.

Lemma 5. *All Fourier coefficients of a multi-linear polynomial resulting from the $\mathcal{KR}$ construction applied to a Boolean function f have their granularity* $\mathsf{gran}\big(\widehat{f}(S)\big)$ *bounded by* $\mathsf{deg}(f)$.

Proof. The Fourier transform applied to each monomial ($\prod_{k\in S_j} x_k$) in the ANF of the Boolean function generates a multi-linear polynomial poly_i with Fourier coefficients $\widehat{f}(S_{j,1})$ of the type $\frac{l_{S_{j,k}}}{2^i}$, with $S_{j,k} \subseteq$

[6] Interestingly, determining the complexity of unitary operations has been an important research topic in quantum computing, with extensive applications from circuit synthesis to quantum error correcting codes [200, 201, 202, 203].

S_j, $k \in [2^i]$, and $i = \deg\big(\prod_{k\in S_j} x_k\big)$ by definition of the Fourier transformation [193]. Therefore, the same is true for the Krawtchouk transform, which computes the same Fourier coefficients.
Then, all such individual polynomials are summed or subtracted based on the coefficients $c_i \in \{0, 1\}$ to generate the final multi-linear polynomial,

$$\underbrace{(-1)^{c_1}\Bigg(\overbrace{\frac{l_{S_{1,1}}}{2^{i_1}}}^{\widehat{f}(S_{1,1})} \bigoplus_{k\in S_{1,1}} x_k + \frac{l_{S_{1,2}}}{2^{i_1}}\ldots + \ldots + \frac{l_{S_1,i_1}}{2^{i_1}}\ldots\Bigg)}_{\mathsf{poly}_1} + \tag{3.5.3}$$

$$\underbrace{(-1)^{c_2}\Bigg(\frac{l_{S_{2,1}}}{2^{i_2}}\ldots + \frac{l_{S_{2,2}}}{2^{i_2}}\ldots + \ldots + \frac{l_{S_2,i_2}}{2^{i_2}}\ldots\Bigg)}_{\mathsf{poly}_2} + \ldots = \mathsf{poly}_f\ . \tag{3.5.4}$$

Therefore the resulting Fourier coefficients ($\widehat{f}(S)$) for each of the parity bases ($\bigoplus_{k\subseteq S} x_k$) in the final multi-linear polynomial are equal to the sum of all non-zero Fourier coefficients that have the same parity base in the multi-linear polynomials resulting from each monomial, $\frac{l_S}{t} = \sum_{S_{j,k}=S} \frac{(-1)^{c_j} * l_{S_{j,k}}}{2^{i_j}}$. Consequently, the denominator t of the resulting value is equal to the least common multiple between all denominators, given that we are summing and subtracting rational numbers. Additionally, all the denominators are of the form $\frac{1}{2^i}$, for which the $l.c.m(\frac{1}{2^{i_1}}, \frac{1}{2^{i_2}}, ..., \frac{1}{2^{i_n}}) = max(\frac{1}{2^{i_1}}, \frac{1}{2^{i_2}}, ..., \frac{1}{2^{i_n}})$. This proves that, for any Fourier coefficient in the final multi-linear polynomial, the corresponding denominator is at most equal to the largest denominator present in the sum. Thus, the largest denominator is at most equal to the denominators of the monomials defining the degree of the Boolean function, which implies that the resulting Fourier coefficients have a denominator that is at most equal to $\frac{1}{2^{\deg(f)}}$, proving the lemma. □

In addition, we relate the level of the Clifford hierarchy of measurement operators with the angles' granularity using the semi-Clifford hierarchy.

Definition 14. (Semi-Clifford hierarchy). *A gate U is in the semi-Clifford hierarchy of level-k if $U = A_1 D A_2$, such that $A_1, A_2 \in Cl_2^N$ are Clifford gates and $D \in Cl_k^N$ is the diagonal. The set of k-th level Clifford gates is denoted by SCl_k^N.*

Lemma 6. *All measurement operators of type $M = \cos\big(\phi\big)\sigma_x + \sin\big(\phi\big)\sigma_y$ are in the level-$\mathsf{gran}\big(\phi\big)+1$ of the Clifford hierarchy.*

Proof. This type of operator can be characterized by the semi-Clifford hierarchy, given that it can be decomposed as,

$$\begin{aligned} I * M_i * \sigma_x &= I\big(\cos\big(\phi\big)\sigma_x + \sin\big(\phi\big)\sigma_y\big)\sigma_x && (3.5.5)\\ &= \cos\big(\phi\big)I - i\sin\big(\phi\big)\sigma_z) && (3.5.6)\\ &= R_Z\big(\phi\big)\ , && (3.5.7) \end{aligned}$$

noting that $R_Z\left(\frac{\pi l}{2^j}\right)$ is a diagonal unitary gate of level-2^{j+1} in the Clifford hierarchy [204], while $I, \sigma_x \in Cl_2$. We conclude that gates of the form M are in the $\mathsf{gran}\big(\phi_i\big)$-level of the semi-Clifford hierarchy. Additionally, in [203] it was proven that $Cl_k^1 = SCl_k^1$, such that all $M \in SCl_{2^{i+1}}^1 = Cl_{2^{i+1}}^1$. Therefore we have proven that a measurement operator of the type of M is in the $\mathsf{gran}\big(\phi\big)$-level of the Clifford hierarchy. □

Finally, it is possible to define an upper bound on the level of the Clifford hierarchy of the measurement operators necessary for deterministic evaluations of Boolean functions.

Theorem 3.5.1. *For any Boolean function f, the $\mathcal{KR}$ construction generates deterministic* $\mathrm{NMQC}_{\oplus}$ *evaluations that use measurement operators belonging to at most the* $\deg(f)$*-level of the Clifford hierarchy*

Proof. With Lemma 5 it is possible to bound the granularity of the Fourier coefficients determined by the Krawtchouk transform for an arbitrary Boolean function. These coefficients then define the angle of the measurement operator based on Equation (3.3.15). Therefore, it is equally possible to bound the granularity of all the angles used in the measurement assigned by the degree of the Boolean function. In addition, the granularity of θ_i from Equation (3.3.2) can be made equal to zero with a linear shift of the Boolean function (a free operation). Consequently, the granularity $\mathsf{gran}(\theta_i + \phi_i)$ of the total angle defining the measurement operator can be bounded by $\deg(f) - 1$. Now, it suffices to use Lemma 6 to bound the level of the Clifford hierarchy of the selected measurement operators. This finishes the proof given that the Krawtchouk transform always generates measurement assignments that evaluate the intended Boolean function deterministically. □

This result demonstrates that minimizing the Clifford hierarchy of the measurement operators is always possible and that the degree of the Boolean function alone defines the minimum and maximum level required. For the upper bound, we used the granularity of the Fourier spectrum, which abstractly was the same concept used in [1] for the lower bound. Furthermore, in Boolean analysis, the granularity of the Fourier spectrum is also used to prove lower bounds [194]. However, in those cases, only Boolean functions are considered. In contrast, for the $\mathrm{NMQC}_{\oplus}$ model, due to the symmetry of the cosine function, there is a large class of equivalent semi-Boolean functions that compute the intended Boolean functions correctly. Therefore, these results can not be translated to the $\mathrm{NMQC}_{\oplus}$ model. Nevertheless, it was with the concept of the minimum possible granularity of all the possible semi-Boolean functions that the author of [164] proved a lower bound for the sparsity of arbitrary Boolean functions. In the end, we highlight the importance of these measures to bound the necessary resources for the deterministic evaluation of Boolean functions in $\mathrm{NMQC}_{\oplus}$.

3.5.2 Lower bound conjecture for symmetric Boolean functions

The previous section introduced several methods to determine measurement assignments that deterministically compute a given Boolean function in the $\mathrm{NMQC}_{\oplus}$ model. Moreover, all these solutions consider the general GHZ state as the resource, which was proven to be the optimal resource [102, 192]. Nevertheless, no construction guarantees that the sparsest polynomials are created. This indicates that determining the minimum number of qubits necessary to evaluate a Boolean function is a nontrivial problem. Additionally, it is possible to show that each Boolean function in the model is equivalent to a Bell inequality [102]. Therefore, one could suspect that the difficulty of the former problem is associated to the known hardness of the latter one [209][7]. Nonetheless, finding lower bounds for specific subsets of Boolean functions was proven possible in [102], and it would be exciting to extend these results further.
Finding the smallest possible resource is generally a complex problem because there are an infinite number of valid measurement assignments for a particular Boolean function. In particular, all these solutions exist because the phase introduced after the measurement process, which defines the result of the computation, is invariant with respect to a 2π rotation (see Equation (3.3.8)). Therefore, various multi-linear polynomials of type $\mathsf{poly} : \{0,1\}^n \mapsto \pi\mathbb{Z}$ are equivalent in the model when all their outputs correspondingly map to the same equivalence class of the quotient set $\pi(\mathbb{Z}/2\mathbb{Z})$. Consequently, the minimum number of qubits necessary to evaluate a Boolean function in the $\mathrm{NMQC}_{\oplus}$ model can be defined by the sparsest element in this set.

[7] Determining the maximal violation of a Bell inequality obtained from measurements on a quantum state was proven to be NP-hard even in the tripartite case [210]

Definition 15. (Minimal sparsity). *The minimum number of qubits necessary for the deterministic evaluation of a Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ in the* NMQC$_\oplus$ *model can be defined as the sparsest multi-linear polynomial equivalent to the function, i.e.*

$$\min \left\{ \text{sparsity}\big(\mathcal{F}(g)\big) \,\middle|\, g : \{0,1\}^n \mapsto \mathbb{Z} \; s.\, t. \; \forall_{x\in\{0,1\}^n} \; f(x) \equiv g(x) \mod 2 \right\} . \tag{3.5.8}$$

Fortunately, while the number of multi-linear polynomials equivalent to the Boolean functions is infinite, the effective search space is finite [193]. Nevertheless, the size of the search space is double exponential with the size of the input string [1]. Thus, an exhaustive search for the sparsest multi-linear polynomials is not feasible.
Despite this tremendously ample search space, there are some restrictions that can make the problem considerably simpler for the set of SBF. More precisely, we will impose the constraint that only symmetric measurement assignments are allowed for their evaluation.

Definition 16. (Symmetric measurement assignment). *A symmetric measurement assignment for an* NMQC$_\oplus$ *computation is composed of a set of linear functions L, which is limited to having either all parity bases of a specific size or none,*

$$L = \left\{ \bigoplus_{l\in S} x_l \,\middle|\, S \subseteq [n],\ |S| \in T \right\},\ with\ T \subseteq [n] \, , \tag{3.5.9}$$

and the corresponding dichotomic measurement operators, which have to be equal for all parity bases of the same size, i.e. $|S_i| = |S_j| \implies M_i(x) = M_j(x)$.

Furthermore, all symmetric measurement assignments for an NMQC$_\oplus$ computation are equivalent to a symmetric multi-linear polynomial (poly$_{\text{sym}} : \{0,1\}^n \mapsto \mathbb{R}$) of the following form:

$$\text{poly}_{\text{sym}}(x) = \pi * \left(\sum_{i\in T} \widehat{f}(|S| = i) * \sum_{S_j\subseteq[n],|S_j|=i} g^{-1}\left(\bigoplus_{l\subseteq S_j} x_l \right) \right),\ with\ T \subseteq [n] \ . \tag{3.5.10}$$

Therefore, with this restriction, the multi-linear polynomials that describe SBF have at most $n+1$ different Fourier coefficients concerning the size of the input string n. Unfortunately, the set of solutions is still of an exponential size to perform a search. Nevertheless, knowing the particular polynomials with the smallest sparsities is not necessary to determine lower bounds. It is sufficient to independently learn the sparsities of the polynomials with the smallest sparsities. Hence, one could define a decision problem that only questions if there is a multi-linear polynomial equivalent to the SBF with a particular sparsity. Fortunately, the combinations of this restriction with the $\mathcal{KR}$ transform allow us to define an algorithm able of answering this decision problem in polynomial time.

Lemma 7. *An algorithm with polynomial scaling runtime can decide whether there exists a symmetric multi-linear polynomial with a selected subset of non-zero Fourier coefficients, $F = \left\{ \widehat{f}(|S| = i) \,\middle|\, i \in T,\ T \subseteq [n] \right\}$, that computes a particular symmetric Boolean function deterministically in the* NMQC$_\oplus$ *model.*

Proof. We will prove the lemma introducing a more detailed description of the problem and the algorithm to solve it. In the end, the algorithm's running time is shown to be polynomially bounded.
The algorithm has to determine whether there exits a symmetric multi-linear polynomial with a particular set of non-zero Fourier coefficients, $F = \left\{ \widehat{f}(|S| = i) \,\middle|\, i \in T,\ T \subseteq [n] \right\}$, that computes deterministically the SBF. Therefore, if this polynomial exists it will be an element of the set of multi-linear polynomials equivalent to the SBF in the NMQC$_\oplus$ model. Thus, it will be helpful to redefine this set in in a similar

way to what we have done previously for the minimal sparsity in Definition 15. First, consider that any semi-Boolean function $g : \{0,1\}^n \mapsto \mathbb{Z}$ that verifies the following equality,

$$g(x) = f(x) + 2h(x) \,, \tag{3.5.11}$$

for any other semi-Boolean function $h(x) : \{0,1\}^n \mapsto \mathbb{Z}$, has a multi-linear polynomial that computes the Boolean function $f(x) : \{0,1\}^n \mapsto \{0,1\}$ deterministically in the NMQC$_\oplus$ model. This allows us to use the semi-Boolean function $h(x)$ as the degree of freedom that determines the elements of the equivalence set. Furthermore this degree of freedom can be restricted for our specific problem. First, only symmetric multi-linear polynomials are considered. Therefore if $f(x)$ is symmetric, then also $h(x)$ has to be symmetric to generate symmetric semi-Boolean functions equivalent to $f(x)$ through Equation (3.5.11). In addition, the same equality on the functions will be true for the corresponding multi-linear polynomials. Therefore, the elements of the set of symmetric multi-linear polynomials $\mathsf{poly}_{g(x)}$ equivalent to the symmetric Boolean function $f(x)$ in the NMQC$_\oplus$ model, must verify the following equality,

$$\sum_{i=0}^{n} \widehat{g}\big(|S| = i\big) = \sum_{i=0}^{n} \widehat{f}\big(|S| = i\big) + 2\sum_{i=0}^{n} \widehat{h}\big(|S| = i\big) \,. \tag{3.5.12}$$

A second restriction imposed by the problem is that only a specific subset of Fourier coefficients $T \subseteq [n]$ can be non-zero. Therefore all other Fourier coefficients of the symmetric multi-linear polynomial have to be equal to zero. This implies that all the valid symmetric multi-linear polynomials correspond respectively to instances of $h(x)$ that verify simultaneously the following equality $\widehat{f}(|S| = i) = -2\widehat{h}(|S| = i)$ for all $i \in E$, with $E = [n] \setminus T$. This guarantees that $\widehat{g}(|S| = i)$ is equal to zero for all $i \in E$. Finally, a solution to the problem exists if there is a semi-Boolean function $h(x)$ for which the following equation is verified,

$$\sum_{x} X_{|S|=i}(x) f(x) = -\sum_{x} X_{|S|=i}(x) 2h(x) \tag{3.5.13}$$

for all $i \in E$. This problem defines a system of linear Diophantine equations. Therefore, determining if a symmetric measurement assignment exists reduces to determining if this system of linear Diophantine equations has a solution. Thus, a solution to the decision problem will always exist. However, we should define if it can be determined in polynomial time with respect to the size of the function. Unfortunately, the equations are of exponential size in relation to the input string because they represent specific lines of the Fourier transform. Therefore, any algorithm that receive these equations as input will need to manipulate an exponential size object. This means that solving the problem would also take exponential time.
Fortunately, the problem can be simplified using the Krawtchouk transform which computes the Fourier coefficient using equations of polynomial size. All the Fourier coefficients will be determined using specific lines of the Krawtchouk transform applied to the simplified value vectors of the functions f and h, designated v_f and v_h. Consequently, the problem reduces to finding a solution to the following equation,

$$\sum_{j=0}^{n} K_{i,j}^{[n]} * v_f(j) = -\sum_{j=0}^{n} K_{i,j}^{[n]} * 2 * v_h(j) \tag{3.5.14}$$

for all $i \in E$. The combination of all these equations defines a system of Diophantine equations that can be represented as,

$$\begin{bmatrix} K_{i_1,1} & K_{i_1,2} & K_{i_1,3} & ... & K_{i_1,n+1} \\ K_{i_2,1} & K_{i_2,2} & K_{i_2,3} & ... & K_{i_2,n+1} \\ K_{i_3,1} & K_{i_3,2} & K_{i_3,3} & ... & K_{i_3,n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ K_{i_k,1} & K_{i_k,2} & K_{i_k,3} & ... & K_{i_k,n+1} \end{bmatrix} * 2 \begin{bmatrix} v_h(0) \\ v_h(1) \\ v_h(2) \\ \vdots \\ v_h(n) \end{bmatrix} = \begin{bmatrix} \widehat{f}(|S| = i_1) \\ \widehat{f}(|S| = i_2) \\ \widehat{f}(|S| = i_3) \\ \vdots \\ \widehat{f}(|S| = i_k) \end{bmatrix} . \tag{3.5.15}$$

In order to verify if a system of Diophantine equations has a solution, one can use the Smith Normal Form of the matrix defining the system. In particular, the matrix defining the system of Diophantine equations required to be solved will be the submatrix of the Krawtchouk transform composed of the lines selected by the set E, $\mathcal{K}^{[n+1]}[E; 1, 2, ..., n+1]$. Therefore, the Smith Normal Form of this matrix creates three unique matrices that verify the following equality,

$$U^{-1}DV^{-1} = \mathcal{K}^{[n]}[E; 1, 2, ..., n+1] . \tag{3.5.16}$$

Then the system of equations to be solved can be decomposed as,

$$U^{-1}DV^{-1} * v_h[E] = [\widehat{f}(E)] \tag{3.5.17}$$

$$U^{-1}DV^{-1} * 2 * v'_h[E] = [\widehat{f}(E)] \tag{3.5.18}$$

$$V^{-1} * v'_h[E] = \frac{D^{-1}U[\widehat{f}(E)]}{2} \tag{3.5.19}$$

with $[\widehat{f(E)}]$ being the set of Fourier coefficients selected by E, resulting from T. Consequently, a solution for the system exists if for all $i \in E$,

$$\frac{(U[\widehat{f}(|S| = i)])}{2D[i, i]} \in \mathbb{Z} . \tag{3.5.20}$$

Finally, to prove that this process can be performed in polynomial time, it will be necessary to show that each step is polynomially bound. First, the Smith decomposition can be performed with algorithms that scale polynomially concerning the size of the matrix, as proven in [211]. This solution demonstrates that determining the Smith Normal Form of the Krawtchouk transform can be done in polynomial time and space, given that the matrix scales polynomially with the size of the input string. Then, all the matrix and vector multiplications are polynomially bounded, provided these objects are of polynomial size [212]. The last process tests if all elements of the final vector, composed of the elements described by Equation (3.5.20), are integer numbers. The previously mentioned vector is smaller or equal to n, which also makes this subprocess polynomially bounded. Therefore, the entire process is polynomially bounded, proving the lemma. □

Furthermore, this algorithm can be employed to identify the order of magnitude of the sparsity of an arbitrary SBF in polynomial time (For more details, see Appendix 3). Subsequently, applying these processes to the set of CSF functions led to the following conjecture,

Conjecture 2. *The number of qubits in a GHZ state necessary for the deterministic evaluation of a complete symmetric function C^k, of constant degree k, with an n bit input string and a symmetric measurement assignment, scales as $\Omega(n^{\lfloor k/2 \rfloor - 1})$.*

Support: This conjecture is supported empirically by a numerical verification process which was applied to all CSF that can be defined with an input size up to 345 bits, which will be described in detail subsequently.

First, the polynomially scaling algorithm described in the proof of Lemma 7 was implemented, in a way such that one could test the order of magnitude upon whether a solution exists or not. This was necessary because the number of combinations of Fourier coefficients that can be selected to be tested grows exponentially. However, applying the algorithm to a subset of these combinations suffices to test the scaling rate. The last statement is supported by the number of parity bases corresponding to each non-zero Fourier coefficient. Moreover, a non-zero Fourier coefficient relative to parity bases of size $t = |S|$ imposes the use of $\binom{n}{|S|}$ qubits and measurements. Therefore, each Fourier coefficient is related to a certain asymptotic growth rate,

$$\underbrace{\underbrace{\binom{n}{0}}_{1}}_{\widehat{f}(|S|=0)}, \underbrace{\underbrace{\binom{n}{1}}_{n}}_{\widehat{f}(|S|=1)}, \underbrace{\underbrace{\binom{n}{2}}_{\mathcal{O}(n^2)}}_{\widehat{f}(|S|=2)}, \underbrace{\underbrace{\binom{n}{3}}_{\mathcal{O}(n^3)}}_{\widehat{f}(|S|=3)}, \quad \cdots \quad , \underbrace{\underbrace{\binom{n}{n-2}}_{\mathcal{O}(n^2)}}_{\widehat{f}(|S|=n-2)}, \underbrace{\underbrace{\binom{n}{n-1}}_{n}}_{\widehat{f}(|S|=n-1)}, \underbrace{\underbrace{\binom{n}{n}}_{1}}_{\widehat{f}(|S|=n)} . \tag{3.5.21}$$

In order to test if there is a solution that has at most a certain growth rate $\mathcal{O}(n^t)$, it is sufficient to test if all the Fourier coefficients with larger growth rates could be made equal to zero. Therefore, this method reduced the number of necessary tests necessary to determine the order of the growth rate to a maximum of $n/2$ for each symmetric function. In particular, for the conjecture it was enough to test if there are solutions with a growth rate of $\mathcal{O}(n^{\mathsf{deg}(f)/2} - 1)$ and $\mathcal{O}(n^{\mathsf{deg}(f)/2})$. Therefore, all Fourier coefficients with larger scaling rates than these values must be zero for a positive answer. This statement implies that the following equation

$$\sum_{j=0}^{n} K_{i,j}^{[n]} * v_f(j) = -\sum_{j=0}^{n} K_{i,j}^{[n]} * 2 * v_h(j) \tag{3.5.22}$$

with $[n]$ defining the size and i, j the element of the Krawtchouk transform, is verified for all $\{i \big| \ i > \mathsf{deg}(f),\ i > n - \mathsf{deg}(f)/2\}$ if the function has a scaling rate of $\mathcal{O}(n^{\mathsf{deg(f)}/2-1})$. Equivalently, for $\mathcal{O}(n^{\mathsf{deg(f)}/2})$ the same equation has to be true for all $\{i \big| \ i > \mathsf{deg}(f) + 1,\ i > n - \mathsf{deg}(f)/2 - 1\}$. The numerical tests realized verified that all complete symmetric functions that can be defined with an input size up to 345 bits have solutions with the respective sparsities scaling with $\mathcal{O}(n^{\mathsf{deg(f)}/2})$ and failed to have a solution with $\mathcal{O}(n^{\mathsf{deg(f)}/2-1})$. Moreover, the main limitation of the numerical process is that for each specific CSF, only instances of a particular size of the input string are tested each time. One can not use this test to verify the sparsities of all possible input strings of a particular CSF simultaneously.
In summary, a reasonable number of functions have verified the conjectured scaling relating to their minimum sparsity. Furthermore, the same conjecture and numerical test are consistent with the lower bound found for two particular symmetric Boolean functions[8]. More numerical tests could be conducted. We believe that actually proving the conjecture will be possible resorting to a general Smith decomposition of the Krawtchouk transform.

Remember that the set of CSFs behaves as a building block for all SBFs, Equation (3.4.1). Therefore, the correctness of this conjecture implies an equivalent lower bound for families of SBF with a constant degree.

Corollary 1. *Any family of symmetric Boolean functions defined with a constant degree and evaluated with a symmetric measurement assignment in the* $\mathsf{NMQC}_{\oplus}$ *model requires a number of qubits which is asymptotically bounded by* $\Omega(n^{\lfloor \mathsf{deg}(f)/2 \rfloor - 1})$.

Support: This corollary follows directly from Conjecture 2 and the definition of SBF's.

[8] Note that the conjecture is consistent with the lower bound for the C^2 function and the AND function. In particular, for the latter, each CSF function is exactly equal to the function for a specific input size, as for all $k \in \mathbb{N}$ and $|x| = k$, $C^k(x) = \mathsf{AND}(x)$.

From an abstract point of view, the conjecture and the respective numerical support further strengthen the idea that the degree of the Boolean function and the size of the input string determine the number of qubits necessary for deterministic computations in the NMQC$_\oplus$ model. Moreover note that the symmetry condition was not imposed on the lower bounds obtained previously. However, the optimal measurements assignments for the SBF AND and C^2 are symmetric. It remains to be proven, but it could be possible that the optimal measurement assignments for SBFs concerning the minimum number of qubits are always symmetric. Supposing that this last hypothesis and the conjecture were both proven true, then the multi-linear polynomials presented in Lemma 2 would be optimal and generate the most resource-efficient measurement assignments for the respective SBFs in the NMQC$_\oplus$ model.

3.5.3 Qubit count for symmetric Boolean functions

In the previous subsection, we conjectured a lower bound for the number of qubits in a GHZ state required for an NMQC$_\oplus$ evaluation of symmetric Boolean functions. Complementary to this result, comparing this bound with the existing upper bound is particularly interesting. For this, it will be necessary to define new upper bounds based on our proposed constructions and their corresponding multi-linear polynomials.
Formerly, the upper bound for SBF resulted from the $\mathcal{EF}$ construction presented in [164]. In particular, this construction generates multi-linear polynomials that had a sparsity equal to $\sum_{i=0}^{\deg(f)} \binom{n}{i}$, which entails an asymptotical scaling of $\mathcal{O}\big(n^{\deg(f)}\big)$. Recently, Frembs *et al.* [1] presented a sparser description for the multi-linear polynomials representing CSFs that directly reduces the sparsity of the multi-linear polynomials used for SBFs. More precisely, the proposed multi-linear polynomials have a sparsity of $\mathsf{sparsity}(C^k(x)) = \sum_{i=0}^{k-1} \binom{n}{i}$, which translates to an asymptotical scaling for SBF equivalent to $\mathcal{O}\big(n^{\deg(f)-1}\big)$. Futhermore, this last upper bound was the best known up to now, to the best of our knowledge.
In order to find new upper bounds, it is necessary to define the impact of the $\mathcal{CSF}$ construction on the sparsity of the multi-linear polynomials generated for SBFs. For this purpose, we must consider that this construction creates multi-linear polynomials using the subset of CSFs with a degree equal to a power of two exclusively. Therefore, it can be proven that the sparsity of the resulting multi-linear polynomials also depends uniquely on the same set.

Lemma 8. *The sparsity of any multi-linear polynomial equivalent to a symmetric Boolean function f is bounded as*

$$\mathsf{sparsity}\big(f(x)\big) \leq \sum_{i=0}^{n} c_i \prod_{r\in R_i} \Big(\mathsf{sparsity}\big(C^{2^r}(x)\big) + 1\Big), \tag{3.5.23}$$

with $R_i \subseteq [\lceil \log(i) \rceil]$, such that $\sum_{r\in R_i} 2^r = i$.

Proof. A valid multi-linear polynomial can be generated for any SBF based on Equation (3.4.1) and the general decomposition defined in Proposition 5.5.26. Therefore, the sparsity of SBF is always upper bounded by the sparsity of the polynomials built from this representation. In particular, the same expression directly defines the relation between the resulting polynomial's sparsity and that of the CSFs used. □

Consequently, for families of SBF with a constant degree, the asymptotic upper bound depends only on the subset of CSFs with degrees that are powers-of-two,

$$\mathsf{sparsity}\big(f(x)\big) = \mathcal{O}\Big(\prod_{r\in R_d} \mathsf{sparsity}\big(C^{2^r}(x)\big)\Big) \tag{3.5.24}$$

with $R_d \subseteq [n]$, such that $\sum_{r\in R_d} 2^r = \deg\big(f\big)$. The fact that only the CSF function with the largest degree is considered can be explained by the fact that the CSFs polynomials have monotonically increasing

sparsity.
Subsequently, this bound has to be combined with the minimal sparsities known for the CSF with a degree equal to a power of two. In particular, for instances with a degree equal to or smaller than 64, the most compact multi-linear polynomials are the ones identified in Lemma 2, and, for higher degree functions, are the ones described in [1],

$$\text{sparsity}\Big(C^{2^r}(x)\Big) = sparsity\Big(C^k(x)\Big) = \begin{cases} \mathcal{O}\big(n^{2^{r-1}}\big) = \mathcal{O}\big(n^{k/2}\big) \Leftarrow & r = 2,3,4,5,6 \\ \mathcal{O}\big(n^{2^r-1}\big) = \mathcal{O}\big(n^{k-1}\big) \Leftarrow & r > 6 \end{cases}, \tag{3.5.25}$$

with $2^r = k$. Therefore when these instances are used to determine the asymptotic sparsities of CSF with arbitrary degrees, we obtain the following bounds,

$$\text{sparsity}\Big(C^k(x)\Big) = \mathcal{O}\big(n^{\rho}\big),\ with\ \rho = \left(\sum_{r\in R_k, r\leq 6} 2^{r-1} + \sum_{r\in R_k, r>6} 2^r - 1\right), \tag{3.5.26}$$

with $R_k \subseteq [\lceil \log(k) \rceil]$, such that $\sum_{r\in R_k} 2^r = k$. Likewise, the same bounds will be valid for for SBFs with a constant degree,

$$\text{sparsity}\big(f(x)\big) = \text{sparsity}\big(C^k(x)\big),\ with\ k = \deg(f(x))\ . \tag{3.5.27}$$

Notice that any CSF that resorts uniquely to the multi-linear polynomials from Lemma 2 in their decomposition (Equation (3.4.3)) has an upper bound of $\mathcal{O}(n^{k/2})$. This bound holds for all CSF with a degree equal to or smaller than 127. For these instances the upper bound is tight to the conjectured lower bound. This indicates that if the general form for multi-linear polynomials identified in Lemma 2 is proven correct, then all CSF functions would have an upper bound tight to the conjectured lower bound. Yet, without this proof, the best upper bound is the one described by Equation (3.5.26). In Figure 9 we represent this newly obtained upper bound, together with the lower bound.
In summary, the CSF construction, in combination with the use of sparse multi-linear polynomials, considerably lowers the number of qubits required to evaluate SBF in the $\text{NMQC}_{\oplus}$, as well as the respective upper bounds. Furthermore, the sparsities for SBF functions can be described as depending on a subset of CSFs. This suggests that the effort to reduce the sparsities of SBFs should be guided towards the subset of CSF.

3.6 Circuit realization of the $\text{NMQC}_{\oplus}$ model

3.6.1 $\text{NMQC}_{\oplus}$ circuits

This subsection translates the measurement assignments for the $\text{NMQC}_{\oplus}$ model discussed in the paper into realizable computations. This analysis is pertinent because the previous works related to this precise model have discussed these computations at an abstract level only. In other words, they do not describe them as circuits composed of elementary operations from a fixed gate set. Here we provide a more complete picture of the computational process, describing the necessary operations to create the GHZ states, characterizing the circuits that compute the auxiliary linear functions, and synthesizing or approximating the required measurement operators, keeping in mind that the latter are rarely part of the fixed gate set. This will enable a more direct comparison with fully classical circuits for Boolean function evaluation, in particular with respect to depth, which is associated with the temporal extension of these circuits.

Definition 17. (Depth). *The depth of an* $\text{NMQC}_{\oplus}$ *circuit is equal to the size of the longest sequence*

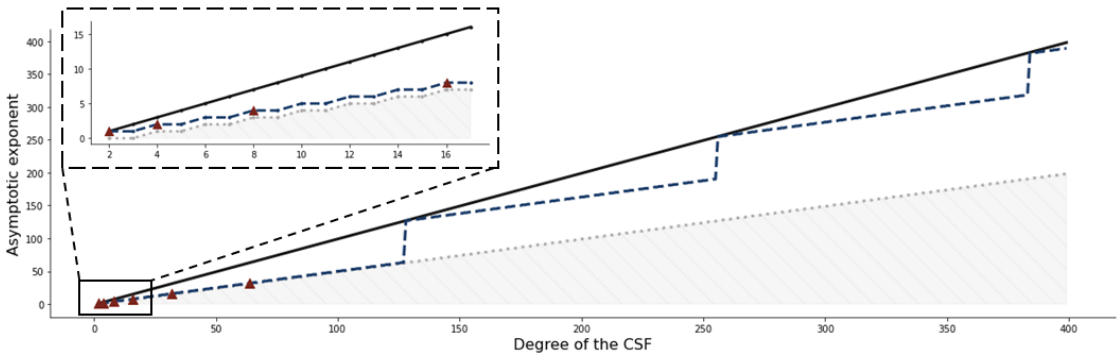


Figure 9: Representation of the relations between the individual CSF (ordered by degree) and the asymptotic growth rate of the GHZ state necessary for deterministic evaluation with the NMQC$_\oplus$ model. The continuous black line represents the most compact polynomials identified in [1]. The dashed dark blue line represents the most compact polynomials constructed using the decomposition from Proposition 1 and the multi-linear polynomials from Lemma 2. Additionally, the red triangles are the CSF computed exactly with the polynomials from Lemma 2. Finally, the dotted grey line represents our conjectured lower bound for the set of CSF (Conjecture 2).

of elementary gates, quantum or classical, in it.

The width of the circuits, which describes the maximal computational space occupied during the execution of the circuits, will be equally important. Generally, parallelizable processes can have a smaller depth but at the cost of a larger width.

Definition 18. (Width). *The width of an* NMQC$_\oplus$ *circuit is equal to the largest number of elementary gates, quantum or classical, executed simultaneously in it.*

The last property that we will account for is the number of gates that the circuits contain. This value relates to previous properties and is also an essential element for general computations as it defines the cost of an evaluation in absolute terms.

Linear pre- and post-processing circuits

The subroutine for the linear pre-processing computes all the linear functions that are part of a specific measurement assignment for a Boolean function. In particular, it contains a number of linear functions equal to the number of necessary qubits in the GHZ state. This number is defined by the sparsity of the multi-linear polynomial that generates the measurement assignment. Furthermore, each of these linear functions, $L_i = \oplus_{i \in S} x_i$, can be highly parallelized, creating a circuit composed of XOR gates with a depth of $\log_2(|S|)$, a width of $|S|$, and a number of gates equal to $|S| - 1$ (Figure 10 (a)). Consequently, the total subroutine is the parallel composition of all linear functions (Figure 10 (b)). Thus, it has a depth of $max\{|S| \mid \widehat{f}(S) \neq 0\}$, a width of $\sum_{S \subseteq [n]} |S| * \delta\big(\widehat{f}(S)\big)$, and a number of gates equal to $\sum_{S \subseteq [n]} \big(|S| - 1\big) * \delta\big(\widehat{f}(S)\big)$, based on the Fourier coefficients $\widehat{f}(S)$ of the multi-linear polynomial $\text{poly}_f(x)$ equivalent to the measurement assignment.
Likewise, the linear post-processing circuit of an NMQC$_\oplus$ computation is a very simple subroutine. It solely sums all the measurement outcomes linearly, computing the parity of the bit string. The computation of this linear function can also be parallelized such that the circuit will have a depth of $\log_2(k)$ with $k = \text{sparsity}(\text{poly}_f)$[9], a width of k, and $k - 1$ gates.

[9] We will label k as the sparsity of the multi-linear polynomial that defines the measurement assignment repeatedly throughout this section.

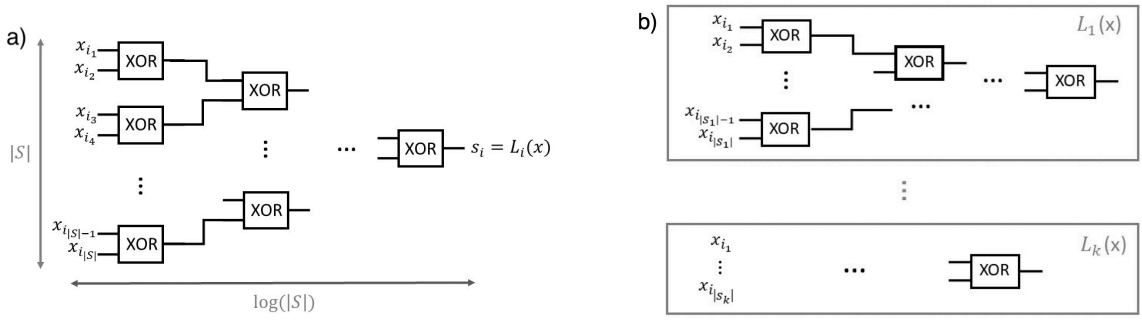


Figure 10: a) The general form and dimensions of a circuit computing a linear function with $|S|$ inputs. b) Parallel composition of the linear functions in the linear pre-processing routine.

GHZ state creation

Let us review two different circuit families used to create GHZ states. The first one is more efficient considering the number of gates while having a simpler process. However, the second compensates for the complexity and the additional operations with the capacity to prepare a GHZ state with any number of qubits in constant quantum depth.

(a) Construction in logarithmic quantum depth

The first construction, defined in [191], creates a n-qubit GHZ state from n qubits initially in the $|0\rangle$ state using only CNOT and H gates. For that, it applies a single H gate to the first qubit and a sequence of CNOT gates in a cascade pattern afterward. In particular, the first CNOT gate takes the first qubit as the control qubit and the qubit in position $n/2$ as the target. Then, in the second layer, two CNOT gates are applied, the first qubits and the qubit in position $n/2$ action as control, and the qubits in positions $n/4$ and $3n/4$ as the target qubits of these gates. This pattern repeats layer after layer, applying twice the number of CNOT gates of the previous layer while using the already entangled qubits as the control qubits and non-entangled qubits as targets. For an illustration of this process, see Figure 11 (a).
It is interesting to characterize this process for a specific NMQC$_\oplus$computation. Given a GHZ state with a number of qubits equal to the sparsity k of the polynomial is necessary, the corresponding circuit will have a depth of $\log_2(k)$, a width of k, and the circuit contains $\mathcal{O}(k)$ gates.

(b) Construction in constant quantum depth

The second technique considered to create GHZ states is particularly interesting because it uses only a constant number of quantum layers to generate a GHZ with an arbitrary number of qubits [213, 27]. Nevertheless, these fixed quantum layers have to be interleaved with classical computations of logarithmic size, which results in a total depth - adding quantum and classical processing - similar to the previous one. Also, the method relies on adaptively chosen quantum operations, which depend on previously obtained measurement outcomes. This adaptivity could be a problem as we are considering a non-adaptive measurement-based quantum computation model. However, as this adaptivity does not depend on the input string or the specific computation that will be performed, we consider this a valid alternative.
The specific protocol presented in [213] creates a n-qubit GHZ state from n qubits in the $|0\rangle$ state using six layers of quantum gates. The first quantum layer applies H gates to odd-numbered qubit wires. Then, the second quantum layer applies $n/2$ CNOT gates between neighboring qubit pairs, with odd-numbered control qubit wires and even-numbered target qubit wires. After this process, $n/2$ Bell pairs were created. The third quantum layer applies $n/2-1$ CNOT gates between the qubits of neighbouring Bell pairs, such that one qubit of the Bell pair is a target qubit for the previous Bell pair, and the other is the control qubit for the next Bell pair. Afterward, the next quantum layer (the fourth) measures all the qubits which were target

qubits of the CNOT gates applied in the previous process. Then, the results of the previous measurements define an X correction that will be applied in the fifth quantum layer. In particular, such X corrections will be applied to the qubits directly below the measured qubits, and they depend on the parity of all the measurement outcomes previous to that specific qubit. Simultaneously, in the fifth layer, the measured qubits are reset. Finally, the sixth and final quantum layer entangles the qubits recycled to the $|0\rangle$ state after the measurement, with the resulting quantum state. An illustration of the whole process is depicted in Figure 11 (b).
In the end, for an NMQC$_\oplus$ computation, this process produces a quantum state with a number of qubits defined by the sparsity of the respective multi-linear polynomial. This implies that the depth of the respective circuit will be $6 + \log_2(k/2 - 1)$, with a width of k and a total number of gates bounded above by $\frac{k^2}{8} + \frac{11}{4}k - 4$.

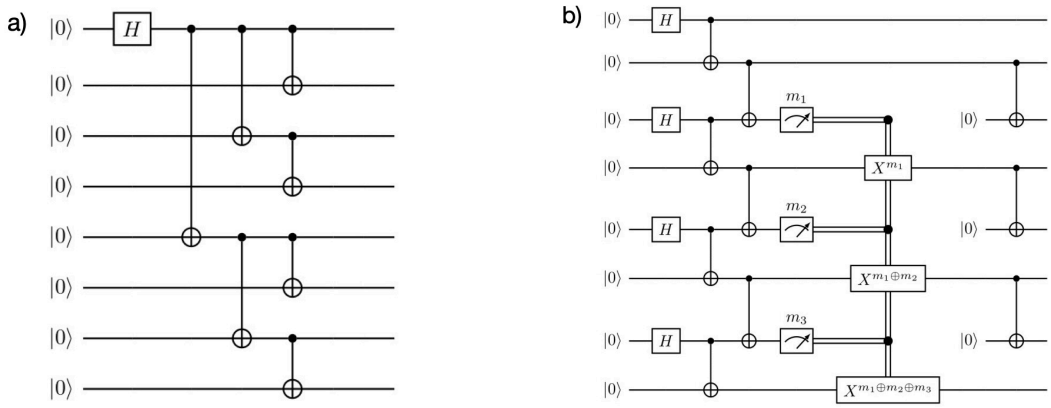


Figure 11: a) Logarithmic depth circuit for the creation of an 8-qubit GHZ state with a recursive CNOT application pattern. b) Constant quantum depth circuit with extra linear measurement post-processing the creation of an 8-qubit GHZ state.

Measurement operators

The measurement operators imposed by the measurement assignment are generally not part of the gate set. Therefore, creating these operators from the available gate set will be necessary. As is common in the literature, we will consider the Clifford+T gate-set. In particular, we need to generate single qubit Z-rotations because these are always equal to the measurement operators used by the model after applying a Pauli X operator. Therefore, the optimal circuit for the measurement operator has at most one extra elementary gate, followed by the optimal circuit for a particular Z-rotation (Figure 12). This observation allows us to use several established results on the circuit synthesis of Z-rotations [214, 215, 216, 204].

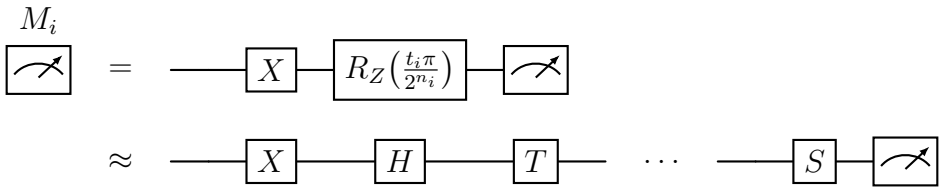


Figure 12: Exact and approximated circuit synthesis of the measurement operator. The former is based on arbitrary R_Z rotations, the latter on the Clifford+ T gate set.

The first significant result is that no $R_Z(\frac{\pi}{2^n})$ operator for $n \geq 2$ can be exactly synthesized using the

Clifford+T gate set [214, 216]. Therefore, using this gate set, Boolean functions with a degree greater than two cannot be computed deterministically in the NMQC$_\oplus$ model.

Corollary 2. *No Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ with* $\deg(f) \geq 3$ *can be computed deterministically with the* NMQC$_\oplus$ *model realized with the Clifford+T gate set.*

Proof. The proof follows from the fact that all the measurement operators necessary for evaluating Boolean functions with a degree greater than two are in third level or higher of the Clifford hierarchy, as proven in [1]. Therefore, each one of these measurement operators is equivalent to a Z-rotation of magnitude $\frac{\pi}{2^n}$ with $n \geq 2$. Thus, it is impossible to synthesize these gates exactly with the Clifford+T gate set [214, 216], and the whole NMQC$_\oplus$ computation will have a non-zero error term. □

This critical implication imposes that higher degree Boolean functions can only be computed with a certain accuracy that depends on the approximation of the measurement operators.
For our purpose, we will need to define approximations that guarantee a certain level of accuracy $\epsilon \in [0,1]$ of the NMQC$_\oplus$ computations. For that, a relationship between the approximations of the Z-rotations and the whole computation is required. From [214], it is possible to define the approximation of a Z-rotation concerning the operator norm for an arbitrary distance $\bigtriangleup$,

$$\bigtriangleup \geq \left|\left| R_Z^*\Big(\frac{\pi * t}{2^n}\Big) |\Psi\rangle - R_Z\Big(\frac{\pi * t}{2^n}\Big) |\Psi\rangle \right|\right|. \tag{3.6.1}$$

with R_Z^* representing the approximated Z-rotation, using a circuit with approximately $K + 4\log_2(1/\bigtriangleup)$ [10] T-gates. Therefore, the approximation of each single measurement operator M_i, to be applied to one of the k qubits in GHZ state, is bounded by

$$\bigtriangleup_i \geq \left|\left| \bigotimes_{j=0}^{i-1} I_j \otimes M_i \otimes \bigotimes_{l=i+1}^{k} I_l |\Psi\rangle - \bigotimes_{j=0}^{i-1} I_j \otimes M_i^* \otimes \bigotimes_{l=i+1}^{k} I_l |\Psi\rangle \right|\right| \tag{3.6.2}$$

$$\geq \left|\left| M_i |\Psi\rangle - M_i^* |\Psi\rangle \right|\right| \tag{3.6.3}$$

$$\geq \left|\left| \sigma_x R_Z\Big(\frac{\pi * t_i}{2^{n_i}}\Big) |\Psi\rangle - \sigma_x R_Z^*\Big(\frac{\pi * t_i}{2^{n_i}}\Big) |\Psi\rangle \right|\right| \tag{3.6.4}$$

$$\geq \left|\left| R_Z^*\Big(\frac{\pi * t_i}{2^{n_i}}\Big) |\Psi\rangle - R_Z\Big(\frac{\pi * t_i}{2^{n_i}}\Big) |\Psi\rangle \right|\right|. \tag{3.6.5}$$

Then, we assume that all the measurement operators have the same error bound $\bigtriangleup$. Also, as all these measurement operators commute, one can interpret that they are applied sequentially, allowing us to bound the error after all the measurements by the sum of all the individual errors [217],

$$k * \bigtriangleup = \sum_{i=0}^{k} \bigtriangleup_i \geq \left|\left| \bigotimes_{i=0}^{k} M_i |\Psi\rangle - \bigotimes_{i=0}^{k} M_i^* |\Psi\rangle \right|\right|. \tag{3.6.6}$$

Finally, we can use the fact that the total deviation distance between the resulting probability distributions between the ideal quantum state that computes the Boolean function deterministically and the resulting quantum state is at most four times the norm error defined for the quantum state [217], such that,

$$\epsilon = (-1)^{f(x)} - \left\langle \Psi_{GHZ}^k \right| \otimes_{i=1}^{k} M_i(s_i) \left| \Psi_{GHZ}^k \right\rangle \tag{3.6.7}$$

$$\geq 4 * k * \bigtriangleup \tag{3.6.8}$$

[10] K is approximately 10, and this value was proven to be lower bounded by -9. Therefore, the algorithm from [214] produces approximations that are at most an additive constant from the optimal Z-approximations.

After determining the relation between the total error of the computations and the local errors due to the measurement operators, one can define a specific number of T gates necessary for a certain precision ϵ of the NMQC$_\oplus$ computation. In particular, the corresponding circuit will have a depth of $4 * c * \log_2(\frac{4k}{\epsilon}) = 4 * c * (2 + \log_2(k) + \log_2(\frac{1}{\epsilon}))$ with c equal to a constant that accounts for all the additional gates in the circuit, which are generally proportional to the number of T gates [208] [11]. Additionally, we end up with a width of k and a total number of gates that is bounded by $4 * c * k * \log_2(\frac{4k}{\epsilon})$.

Depth, width and number of gates

To characterize the NMQC$_\oplus$ circuits, we will combine the previously described subroutines. In particular, we will combine them to minimize the circuit depth. Therefore, the linear pre-processing and the GHZ state generation will be performed simultaneously, as they have no dependencies. Then, the results of the linear functions define the set of operations that will be realized on the GHZ-state qubits in the measurement process. Then, the linear post-processing will compute the parity of the measurement outcomes to obtain the output of the Boolean function. The whole process is represented in Figure 13.
With the architecture of the circuits defined, it is possible to characterize the relevant properties resulting from specific measurement assignments. In particular, the multi-linear polynomials associated with the measurement assignments completely determine the circuits' depth, width, and number of gates.

Lemma 9. *Any Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ with $\deg(f) > 2$ can be computed with accuracy $\epsilon \in [0,1]$ based on a valid measurement assignment using a circuit composed of the Clifford+T gate set with a depth of*

$$2 * \log_2(k) + 4 * c * \left(2 + \log_2(k) + \log_2\left(\frac{1}{\epsilon}\right)\right), \tag{3.6.9}$$

a width of

$$k + \sum_{S \subseteq [n]} \left(|S| - 1\right) * \delta\left(\widehat{f}(S)\right), \tag{3.6.10}$$

and a number of gates bounded by

$$\sum_{S \subseteq [n]} \left(|S| - 1\right) * \delta\left(\widehat{f}(S)\right) + k * \left(2 + 4 * c * \left(2 + \log_2(k) + \log_2\left(\frac{1}{\epsilon}\right)\right)\right). \tag{3.6.11}$$

W.l.o.g $\widehat{f}(S)$ are the Fourier coefficient and k the sparsity of the multi-linear polynomial corresponding to the measurement assignment that will be realized.

Proof. Follows from the arguments above. □

It is essential to mention that we used the GHZ construction of Subsection 3.6.1 , with logarithmic depth. If constant quantum depth methods were used, a slight difference in the depth and the number of gates will be introduced. Nevertheless, the order of magnitude of the depth would not change. Moreover, the depth necessary to approximate the measurement operators can, in principle, be reduced with other circuit synthesis techniques [218, 219]. However, the width of the circuits would increase with the use of auxiliary qubits. In addition, we recall that the proposed solution does not guarantee the optimal realization of a given measurement assignment concerning each one of the metrics for the circuits.
Although there is a large variability in the possible realizations of a measurement assignment, we consider that the particular translations presented unravel essential details and questions about the NMQC$_\oplus$ model. For instance, one could question if there is another gate set that guarantees the exact synthesis of the

[11] In reference [208], this constant appears to be close to 2.5. However, for our analysis, we will leave it undefined

measurement operators. That existence would allow deterministic computations for Boolean functions with an arbitrary degree. However, in [204], various gate sets were considered, and none had a particular advantage, nor could they exactly synthesize arbitrary Z-rotations. This suggests that computing Boolean functions with an arbitrary degree could be a general limitation of the NMQC$_\oplus$ model that only adaptivity can bypass.

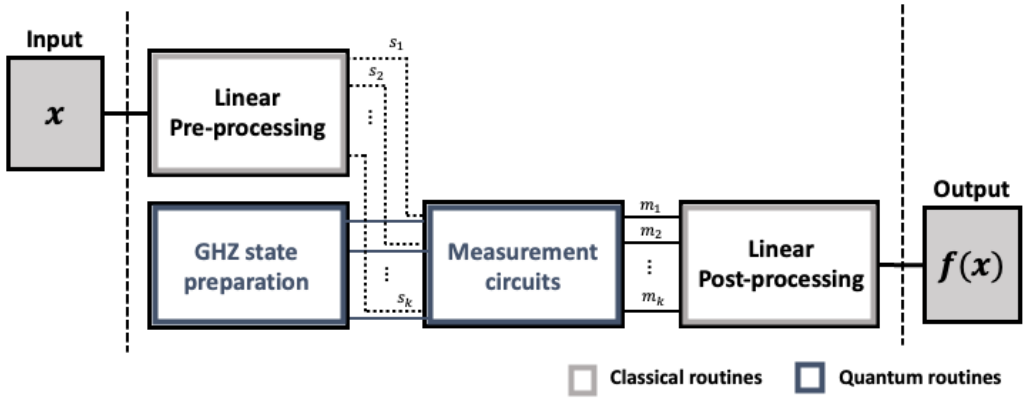


Figure 13: A schematic representation of the circuit realization proposed for the NMQC$_\oplus$ model. Vertically aligned processes are performed in parallel, while horizontally aligned processes are performed sequentially.

3.6.2 Quantum vs. classical circuits

In this last subsection, we compare the known upper and lower bounds on classical circuits for Boolean function evaluation with quantum circuits evaluating the same functions in the NMQC$_\oplus$ model. In particular, this comparison will focus on symmetric Boolean functions. This choice is justified by the contributions to the case of symmetric Boolean functions discussed in this paper and the many results on symmetric functions that can be found in the classical circuit complexity literature [220, 45, 19, 221, 196].

Degree two symmetric Boolean functions

The translation of NMQC$_\oplus$ computations from an abstract description to explicit circuits has shown that degree two Boolean functions are particularly interesting as they evaluate deterministically. Thus, their circuits' depth and number of gates do not depend on a prescribed level of approximation accuracy. Furthermore, we were fortunate to observe that some of these functions stand on the edge of the best lower bounds known for their classical circuit size [222]. In particular, this allows us to notice that there is at least one family of symmetric Boolean functions of degree two with a logarithmic advantage over any conceivable classical circuit. The advantage exists in terms of width and number of gates, considering that the classical circuits can be formed by Boolean operations from the full basis $\{\mathsf{AND}, \mathsf{OR}, \neg, \mathsf{XOR}, \mathsf{NAND}, etc.\}$.

Theorem 3.6.1. *Any classical circuit with unary and binary Boolean operators with single fan out computes the complete symmetric function of degree two C^2 with $\Theta(n * \log_2(n))$ gates and circuit width. In contrast, realizations of the* NMQC$_\oplus$ *model compute this function with $\Theta(n)$ gates and circuit width.*

Proof. It will be necessary to show that the upper and lower bounds have the same scaling rate for both functions in both circuit classes. Considering first the classical circuits, we can relate the C^2 function with the counting functions described in [222]. This is possible because these functions are defined as $\mathsf{Count}(x, k) = 1$ if $\sum_{i=0}^{n} x_i \; mod \; k = 0$. Therefore

$$\mathsf{Count}(x, 4) = \neg\big(C^2(x)\big) \; \mathsf{AND} \; \neg\big(C^1(x)\big) \tag{3.6.12}$$

provided that $C^2(x)$ and $C^1(x)$ represent the two least significant bits of the Hamming weight $|x|$. Then it is possible to relate the length of the Boolean formulas of $\mathsf{Count}(x,4)$ and $C^1(x)$ with the length of the formulas for $C^2(x)$,

$$L(\mathsf{Count}(x,4)) \leq L(C^2(x)) + L(C^1(x)) \tag{3.6.13}$$

$$L(\mathsf{Count}(x,4)) - L(C^1(x)) \leq L(C^2(x))\ . \tag{3.6.14}$$

From these relations between the length of the formula and the lower bound of $\varsigma n \log_2(\frac{n}{4})$ with $\varsigma \in [0,1]$ for $L(\mathsf{Count}(x,4))$ and the exact scaling of n for $L(C^1(x))$, we obtain that $L(C^2) \geq \varsigma n\big(\log_2(\frac{n}{4}) - 1\big)$ [222]. Thus, $\Omega\big(L(C^2(x)\big) = n * log(n)$.
The upper bound of $L(C^2(x))$ is simply obtained by considering the following expression

$$C^2(x) = \bigoplus_{i=1}^{log_2(n)} \Big(\bigwedge_{j=0}^{2^i-1} \Big(\bigoplus_{k=j*\frac{n}{2^i}}^{(j+1)\frac{n}{2^i}} x_k \Big) \Big)\ , \tag{3.6.15}$$

from which we obtain that $L(C^2(x)) \leq \mathcal{O}\big(n * \log_2(n)\big)$. Consequently, we conclude that $L(C^2(x) = \Theta\big(n * \log_2(n)\big)$. Finally, if circuits are composed of unary and binary operators with a single fan-out, the number of gates equals the Boolean formula's length [220]. In addition, the circuit width also scales with the size of the formula and the depth with $\log_2(L(C^2(x))$.
For the NMQC$_\oplus$ circuit realization of the $C^2(x)$ function, one can consider the multi-linear polynomials from Lemma 2. These have a sparsity of $n+1$ and have been proven optimal in [102]. Therefore, the process described in section 3.6.1 translates these to quantum circuits with a number of gates and width of $\Theta(n)$. Also, the computations are deterministic, given that the measurement operators can be synthesized exactly for Boolean functions with a degree equal to two. □

This separation is based on a single family of symmetric Boolean functions. However, this can be enlarged to all the functions in the $[aC^k]$ class for which the symmetrized function is C^2. The explanation is that all these functions will have an NMQC$_\oplus$ realization with a number of gates scaling with n based on Lemma 4. On the other hand, classical circuits for those functions scale with $n * \log_2(n)$ in width and number of gates.
Moreover, suppose one argues that this circuit separation is still not strong enough because it misses to account for the necessary repetitions of the input strings. For that matter, we can show that the process that generates these repetitions for the NMQC$_\oplus$ circuits will always be more efficient. In particular, it is sufficient to verify that the classical circuits require more copies of the input string than the NMQC$_\oplus$ circuits. The classical circuits require a number of copies scaling with $\log_2(n)$, while the quantum solution requires only a single copy of the input string. This single copy is enough because it can be used to compute the parity while the given input string provides all the single bits that control the rotations necessary to evaluate the function (Lemma 3.4.27). Therefore, if the cost related to this process was included, the gap between the quantum and the classical circuits would further increase.
Another possible counter-argument relates to the depth of the circuits. We obtained a quantum circuit family for the specific C^2 function whose depth scales with $2\log_2(n)$, while the classical circuit scales only with $\log_2(n)+\log_2(\log_2(n))$. Consequently, they scale with the same asymptotical rate of $\mathcal{O}(\log(n))$. Although one could argue that there will be an asymptotical constant advantage equal to two in the classical case, this constant depth advantage is insignificant compared to the logarithmic width and gate advantage, which grows infinitely. Also, one could exclude the GHZ state preparation from the evaluation process and address it as polynomial size advice. This assumption is justified by state preparation is independent of the input. In that case, the C^2 function could be evaluated with a depth of $\log_2(n)+7$, producing a small additive advantage of $\log_2(\log_2(n))$ in comparison to the optimal classical circuits for which we do not

know any equivalent advice.

Lemma 10. *Suppose an $n+1$-qubit GHZ state is provided for an* NMQC$_\oplus$ *computation. Then, the complete symmetric function of degree two C^2 can be evaluated with a depth equal to* $\log_2(n)+7$.

Proof. Considering the measurement assignment that results from the multi-linear polynomial of Lemma 2, there are n measurements that depend directly on the input bits. These can be measured and their output linearly summed while the linear side computer determines $\oplus_{i=0}^{n-1} x_i$. When the linear side computer finishes, the last qubit is measured in the determined basis, and the outcome will be included in the linear post-processing. Therefore, the depth of the computations will be equal to or smaller than the depth of operations associated with the measurements, plus the linear post-processing. The measurement process for degree two Boolean function can be executed in 6 layers using only gates of the Clifford+T gate set, given that the control Pauli X and Y gates can be synthesized within that length. Thus, this value is summed to the linear post-processing, which has a depth bounded by $\log_2(n)+1$, obtaining a total depth bounded by $\log_2(n)+7$. □

We believe this is an interesting quantum-classical separation result because it does not depend on any conjecture. Also, it is proven at the level of circuits [12], which is typically very difficult to do based on the difficulty of proving classical or quantum circuit lower bounds [113]. A related and stronger finding was the breakthrough result of [181] that identified a class of search problems with a more significant quantum-classical separation on the circuit level. Other separations at the circuit level were found in cases where the computational space was reduced [188, 187, 224]. These demonstrated a width advantage for quantum circuits at a small and fixed size. In comparison, our result does seem to indicate that this width advantage could scale for computations without limited space. Other separations of larger magnitude have been found [225, 226, 151, 227]. However, most of these are proven under the assumption of black boxes or query complexity [113]. Thus, they are not proven under the more physically realistic assumptions we consider here.
Furthermore, we observe that Theorem 3.6.1 implies that there exists a dimension of the input string above which the NMQC$_\oplus$ evaluation of the C^2 functions starts to be more efficient than any classical circuit evaluation based on Boolean formulas. This point is particularly interesting because one can relate this NMQC$_\oplus$ evaluation to the generalized Svetlichny inequalities [13]. Therefore, by doing this, we conclude that starting from some specific number of qubits, a maximal violation of the generalized Svetlichny inequalities, with a quantum state and measurement operators that can be prepared with a linear number of operations, simultaneously proves a circuit separation. This observation connects separations related to communication problems to circuit problems. Also, the magnitude of the separations is intriguing as one is exponential, and the other is only logarithmic for the function at hand.

Higher degree symmetric Boolean functions

To analyze higher degree symmetric Boolean functions, we isolate first the functions with degree three. This differentiation occurs because we identified a family of symmetric Boolean functions that presents a logarithmic advantage between the size of their measurement assignments and their optimal Boolean formulas. In particular, this family of Boolean functions refers to the $\mathsf{Count}(x,4)$ function described in the proof of Theorem 3.6.1. It is possible to find measurement assignments for these functions with sparsity equal to $2n+2$, based on Equation (3.6.12), Lemma 2 and Lemma 8. In addition, their Boolean formulas are lower bounded by $\varsigma n \log_2(n)$, with $\varsigma \in [0,1]$.

[12] Boolean formulas are a subclass of circuits because they lack memory. Therefore, they are believed to be weaker than circuits that could reuse sub-computations. However, it is still open if the class of poly-size circuits is larger than the class of poly-size Boolean formulas [223].

[13] The connection of NMQC$_\oplus$ evaluations and Bell inequalities was established in a one-to-one relation by the authors of [102].

Although we identified a separation at an abstract level, it almost disappears when these measurement assignments are translated to circuits. The depth of the circuit has an additional term for functions with a degree equal to or greater than three (Lemma 9). This implies that the depth of the resulting circuits is at least twice the classical depth. Also, the number of gates will scale as $n * \log_2(n)$, similarly to what happens in classical circuits. Therefore, the remaining advantage presents itself with the width of the quantum circuits being logarithmically smaller. However, the computation has an error term that does not exist in the classical case. A fair comparison would require the consideration of probabilistic classical circuits. We leave the analysis between the probabilistic case and circuits with different depth and width relations for future work.
For higher degree functions, we use the bounds of the Boolean formulas identified for the Counting functions, with the modular division fixed to a power of two, to define upper bounds for the complete symmetric functions.

Lemma 11. *The Boolean formula of any CSF $L\big(C^k\big)$ is bounded by $\mathcal{O}\big(n \log_2(n)^{\lfloor \log_2(k) \rfloor}\big)$.*

Proof. First, the relation between the Counting functions and the CSFs for powers of two can be established by generalizing Equation (3.6.12) to $\mathsf{Count}(x, 2^p) = \bigwedge_{i=1}^{p-1} \neg C^{2^i}$. Then, this allows us to relate Boolean formulas such that

$$L(C^{2^{p-1}}) \leq L\big(Count(x, 2^p)\big) + L\Big(\bigwedge_{i=1}^{p-2} \neg C^{2^i}\Big) \tag{3.6.16}$$

$$\leq L\big(Count(x, 2^p)\big) + \sum_{i=1}^{p-2} L\big(C^{2^i}\big) \tag{3.6.17}$$

Then, one can establish the upper bounds recursively for all CSF of the form C^{2^p} by using these relations, starting from the fact that $L(C^1) = n$, and sequentially use the upper bounds of $\mathsf{Count}(n, 2^p)$. In particular, the upper bounds of such Counting functions are $L(\mathsf{Count}(n, 2^p)) = \mathcal{O}\big(n * \log_2(n)^{p-1}\big)$. Therefore, we obtain $L\big(C^{2^p}\big) = \mathcal{O}\big(n * \log_2(n)^p\big)$ as the upper bound for the corresponding Boolean formulas. Finally, using Lemma 1, one can generalize the bound to arbitrary degrees of the CSFs given that

$$L\big(C^k\big) \leq \sum_{r \in R_k} L\big(C^{2^r}\big), \tag{3.6.18}$$

with $R_k \subseteq [\lceil \log(k) \rceil]$, such that $\sum_{r \in R_k} 2^r = k$. This relation demonstrates that the largest CSFs present in its decomposition into powers of two will define the upper bound, combining these results with the bounds established for those values, we conclude the proof. □

In comparison, our lower bound conjecture indicates that no symmetric measurement assignment in the NMQC$_\oplus$ model can compute the CSF C^k with less than $n^{\lfloor k/2 \rfloor - 1}$ qubits (Conjecture 3). Therefore, unless our conjectured lower bound deviates immensely when non-symmetric measurement assignments are considered, the NMQC$_\oplus$ computations have a polynomial disadvantage compared to the classical circuits. In particular, these would materialize in constant depth, polynomial width, and polynomial gate advantages for the classical circuits.
Moreover, given the lower bounds known from [102, 164], and the previous analysis, the NMQC$_\oplus$ model is suspected to be computationally inefficient for higher degree Boolean functions. Actually, it cannot even evaluate those functions deterministically with the Clifford+T gate set (Corollary 2). These results point to the importance of adaptivity for deterministic computations for an MBQC model. Remember that if an adaptive process is allowed, then a circuit composed of individual NMQC$_\oplus$ computations can replicate any

classical and deterministic circuit [154]. Similarly, the results of [155] and [156] constructively describe the process of reaching determinism via adaptive measurement patterns for universal quantum computations.

3.7 Discussion and conclusions

We investigated the problem of evaluating deterministically Boolean functions in the $\text{NMQC}_{\oplus}$ model, providing new constructions and a more precise understanding of the necessary instructions for each stage of the model. We characterized the complexity of two resources in this model: the number of required qubits in a GHZ state, and the required level of the Clifford hierarchy for single-qubit measurements. Regarding the number of qubits in a GHZ state, we lowered the upper bound for the number of qubits required to evaluate the entire set of symmetric Boolean functions and conjectured a lower bound for this same set, assuming symmetries between the instructions of the process. Regarding the complexity of the measurements required, we proved an upper bound on the level of the Clifford hierarchy. In particular, this bound is proven to be tight to the boundary identified in [1], which limits the degree of the Boolean functions that a set of operators can evaluate. In the end, we translated the examined $\text{NMQC}_{\oplus}$ evaluations to possible circuit realizations. This translation motivated our main result demonstrating a circuit separation for a family of degree two SBFs. Moreover, we analyzed possible separations for higher-degree functions. Although these functions have not shown any prospect for advantage under the $\text{NMQC}_{\oplus}$ model, they guide the way to other computational models with potential advantages. Also, they can still solve non-local games for which classical circuit analogues fail. Therefore, their computation using the $\text{NMQC}_{\oplus}$ model could still be an exciting element for the study of non-locality and contextuality.

Furthermore, we highlight that proving the conjectured lower bound for the number of qubits necessary to evaluate an SBF on the $\text{NMQC}_{\oplus}$ model would be an exciting result. It directly affects the understanding of the model, and it would be an interesting result for understanding non-local correlations. We state this because each Boolean function relates to a specific Bell inequality [102]. Thus, results on one front relate to the other, and optimal violations of Bell inequalities are very difficult to define for arbitrary system sizes [167]. In addition, such a result would be valuable to test the conjecture presented by [173, 174, 175], which attempts to axiomatize quantum mechanics from information theoretical assumptions. For instance, in [228, 229] the authors define axioms based on communication complexity problems. In particular, their axiom states that correlations that would allow our physical universe to solve these problems trivially are forbidden. Then, if this particular hypothesis is correct, one should retrieve quantum correlations from the maximum allowed by the non-triviality constraint. Interestingly, in [172], the C^2 was used as a non-local black box for which the authors recovered the maximal allowed efficiency at a constant distance from the corresponding maximal quantum efficiency. Therefore, introducing lower bounds for other SBFs would provide more tests for this concept, given that one can analyze any non-local box and compare the respective maximal quantum correlations to their maximal efficiency.

We showed that a circuit advantage in the $\text{NMQC}_{\oplus}$ model is possible for evaluating certain Boolean functions with a degree equal to two. However, these computations were proven inefficient for particular higher degree functions compared to the best-known classical circuits. This limitation motivates the extension of the algebraic tools used for the non-adaptive case to temporally-structured computations. In particular, one could potentially connect this computational framework to other circuit synthesis techniques [230], classical algebraic techniques [231], or some existing quantum signal processing techniques [232, 233]. The possible results from using these new techniques could be the discovery of compact quantum circuits with potential advantages in both deterministic and probabilistic settings. Also, they could describe possible trade-offs between space and computational time.

For both extreme cases of minimal depth or width, it would be interesting to understand if circuits exhibit some quantum advantage. For instance, for the depth necessary to evaluate the Boolean OR function defined on any possible size of the input string, one could expect an advantage [234, 235, 236]. In par-

ticular, in [237], the use of a single adaptive step demonstrated an exponential reduction of the number of qubits for the same function. The same separation was then generalized equally with a single adaptive step for arbitrary symmetric Boolean functions in [164] and with a constant number of adaptive steps in [103] resorting to the previously mentioned quantum signal processing techniques. Identically, suppose one does try to minimize the computational space. In that case, the quantum circuits could have some width advantage. Interestingly, for a small computational space, a separation was already identified in [188, 187, 224], and experimentally explored on IBM's quantum devices [187]. Nevertheless, we question whether the width separations for quantum and classical computations hold for computations with unbounded width and arbitrary input size. Our results seem to indicate that this could indeed be the case. Another type of analysis would be focused on an intermediate class of quantum circuits. For that, we consider the circuits class composed of quantum layers that can only pass classical information between them. Remember that for MBQC computations, the various quantum layers are entangled. Therefore, there is a quantum and a classical channel between each layer [14]. In addition, the computational capacity of this intermediate class is not entirely understood [164, 239]. However, if they present any advantage, they are easier to implement in Noisy, intermediate-scale quantum devices (NISQ), being therefore attractive candidates for experimental realizations [240]. That reduction in the hardware requirements relates to the fact that these circuits demand less connectivity between the qubits and coherence of the gates. In the end, describing such intermediate complexity classes could be a very appealing task to understand further the importance of a temporal structure in general quantum computations and other particular circuits with advantages [181, 183].

Finally, we hope this work will inspire more experiments like the ones described in [241, 103, 242], to verify the computational capabilities of quantum correlations with realizations equivalent to the $\mathrm{NMQC}_{\oplus}$ model. Moreover, as other works have shown, these realizations could work as a device benchmarking tool [243, 244, 245, 246, 247]. Independently, the ideas explored in this paper are also connected to some secure multi-party computation protocols [248], which consider clients with limited computational power. Therefore, we believe that the techniques developed here could allow the design of new and better protocols. These could then potentially motivate exciting experiments, such as the ones reported in [249, 250].

[14] Some initial ideas on how to describe a quantum wire that transmits information between layers have been proposed in [238].

Chapter 4

> "If you would be a real seeker after truth, it is necessary that at least once in your life you doubt, as far as possible, all things. "
>
> — *Rene Descartes*

Quantum circuits surpass biased thresholds circuits in constant-depth

Overview

The relevance of shallow-depth quantum circuits has recently increased, mainly due to their applicability to near-term devices. In this context, one of the main goals of quantum circuit complexity is to find problems that can be solved by quantum shallow circuits but require more computational resources classically.
In this section, we study classes of constant-depth circuits with gates that compute restricted polynomial threshold functions, recently introduced by [115] as a family that strictly generalizes AC^0. Denoting these circuit families $\mathsf{bPTF}^0[k]$ for *biased polynomial threshold circuits* parameterized by an integer-valued bias k, we prove three hardness results separating these classes from constant-depth quantum circuits (QNC^0).

- We prove that the parity halving problem [27], or alternatively the $\mathcal{R}_{\langle f,p\rangle}$ problem with $p = 2$ and f is the Least Significant Bit function, which QNC^0 over qubits can solve with certainty, remains average-case hard for polynomial-size $\mathsf{bPTF}^0[k]$ circuits for all $k = \mathcal{O}(n^{1/(5d)})$.
- We construct a new family of $\mathcal{R}_{\langle f,p\rangle}$ type relation problems based on computing Mod p for each prime $p > 2$, and prove a separation of QNC^0 circuits over higher dimensional quantum systems ('qupits') against $\mathsf{bPTF}^0[k]$ circuits for the same range of bias as above.
- We prove that both foregoing results are noise-robust under the local stochastic noise model, by introducing fault-tolerant implementations of non-Clifford $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuits, that use logical magic states as advice.

We also prove tighter lower bounds on the size of $\mathsf{bPTF}^0[k]$ circuits that are required to solve the relational problem with certainty, which we leverage to significantly reduce the estimated resource requirements for potential demonstrations of quantum advantage against shallow-depth classical circuits. The paper providing the content for this section can be found in [143].

4.2 Introduction

Realizing the theoretical gains promised by landmark quantum algorithms such as integer factorisation or search is challenged by the constraints of existing quantum hardware, requiring extensive resources including qubit counts, coherence times and fault-tolerant operations for their implementation [9, 10, 11]. A series of pioneering studies showed that under plausible complexity theoretic assumptions—e.g. that the polynomial hierarchy does not collapse to the third level—certain classes of quantum circuits are exponentially hard to simulate classically while being more resource-efficient and hardware-friendly. Conditional hardness results of this kind include IQP circuits [13], Boson sampling experiments [14], and random circuit sampling [15]. On the other hand, the need for complexity theoretic assumptions and the difficulty in accounting for the noise levels present in current quantum experiments highlights a gap between theoretical models and practical implementation.

Instead of trying to compare the capabilities of shallow-depth quantum circuits with general models of classical computation, recent work has turned towards the comparison to their classical shallow-depth counterparts. Shallow-depth circuits have been studied extensively in classical computer science [19, 20, 21, 22], and continue to be a vibrant area of research at the frontier of complexity theory [251, 25, 252, 253, 254, 255, 256, 257, 258].

The seminal work of Bravyi et al. [18] exhibited a relational problem that can be solved by constant-depth quantum circuits consisting only of 2-qubit Clifford gates, which no constant-depth classical circuit with bounded fan-in gates can solve. This first *unconditional* separation, i.e. without any complexity theoretic assumptions, between QNC^0 and NC^0 spurred renewed interest in the field, and was followed by an extension of this result to a separation of QNC^0 against classical circuits of unbounded fan-in (AC^0) [27], average-case hardness [33, 27, 34], interactive protocols separating QNC^0 from classical logarithmic-depth circuits of bounded fan-in (NC^1) [74], the case of noisy quantum circuits requiring fault-tolerance and quantum error correction [35, 36, 37], and sampling problems [38].

Constant-depth classes beyond AC^0. Going beyond the AND and OR gates allowed in AC^0, one can consider unbounded fan-in Mod p gates, and the corresponding family of $\mathsf{AC}^0[p]$ classes which contain such gates. [27] showed a separation between $\mathsf{QNC}^0/\mathsf{cat}$ and $\mathsf{AC}^0[2]$ where the former class has access to cat or GHZ advice states. [74] showed a separation between QNC^0 and $\mathsf{AC}^0[p]$ for every prime p in the setting of interactive protocols. Yet, no unconditional separations are known of this kind for relational problems.

An even more general class of gates is *thresholds* Th_k, which output one if the input string has Hamming weight at least k, and zero otherwise. Constant-depth circuits of threshold gates form the class TC^0 [259], and encompass NC^0 and AC^0. They are a canonical theoretical model for neural networks [4, 260, 261], extending even to the transformer architecture which underlies large language models such as GPT [3, 6], whose general interest has spread across all scientific fields [262]. A natural question that arises is whether quantum circuits can also surpass threshold circuits in constant depth. The results of [74] include a separation between QNC^0 and TC^0, but only in the interactive setting and furthermore conditional on the (weaker) assumption that $\mathsf{TC}^0 \subsetneq \mathsf{NC}^1$. A separation of this nature that is unconditional and in the setting of relational problems without interactivity would be desirable, but appears out of reach of current techniques; TC^0, or the subclass[1] $\mathsf{AC}^0[6]$, lies at the very frontier of classical circuit complexity.

4.2.1 Main results

In this paper, we thus explore an intermediate family of circuit classes newly introduced in [115], which we call $\mathsf{bPTF}^0[k]$ for constant-depth *biased polynomial threshold circuits*. They are built from gates

[1] For a representation of the known inclusions among some of the main circuit classes discussed, see Figure 14.

that compute *biased polynomial threshold functions* (PTF), which may be non-linear but include a bias parameter k that restricts their behaviour to be either AND or OR functions when the input string has Hamming weight at least k. Thus for non-zero and constant k (i.e. $k = \mathcal{O}(1)$) this class equals AC^0, and for bias $k = \Theta(n)$ it contains TC^0, giving us a way to study circuits beyond AC^0 in this spectrum of bias parameterized by k. This family of circuit classes is well-motivated from the theoretical standpoint [115], for its connections to polynomial threshold functions and the interplay between the polynomial hierarchy and the counting hierarchy (Section 4.2.4), and also for its relation to theoretical models of neural networks (Section 4.8.2). It thus opens up new avenues for studying the interrelationships between quantum and classical circuits and computational models.

Following previous unconditional separation results, we focus on relational or search problems, where the inputs x are n-bit strings and the outputs y are m-bit strings such that $(x, y) \in R \subset \{0,1\}^n \times \{0,1\}^m$ for some relation R. We define QNC^0 to be the class of computational problems $R : \{0,1\}^n \to \{0,1\}^m$ solvable by quantum circuits with a constant depth using a polynomial number of gates of bounded fan-in (i.e., every gate has a fixed constant number of input and output wires) drawn from a finite, universal quantum gateset. Similarly, we denote by $\mathsf{bPTF}^0[k]$ the class of constant-depth classical circuits composed of unbounded fan-in gates that compute k-biased Polynomial Threshold Functions (PTF) [115], which we define below.

Definition 19 (see Definition 24)**.** *A polynomial threshold function*[2] *with bias k is defined as follows,*

$$f(x) = \begin{cases} P(x), & \sum_{i=1}^{n} x_i \leq k \\ 1, & \sum_{i=1}^{n} x_i > k \end{cases}; \quad \textit{with } P : \mathbb{F}_2^n \to \mathbb{F}_2 \textit{ a polynomial over } \mathbb{F}_2 = \{0,1\}, \tag{4.2.1}$$

where k restricts the maximum degree of P.

Note that the value of k also limits the degree of the polynomial P, since the restriction on the Hamming weight means that any monomial of degree at least $k+1$ must be identically zero when $|x| \leq k$. For example, we have the k-OR ($= 1$ iff $> k$ bits are 1's) and k-AND ($= 0$ iff $< n - k$ bits are 1's). When $k = \mathcal{O}(1)$, we recover unbounded fan-in AND, OR gates and Not gates, and $\mathsf{bPTF}^0[\mathcal{O}(1)]$ corresponds to AC^0. When $k = n$, the polynomial P can have degree n and so $\mathsf{bPTF}^0[n]$ includes *all* Boolean functions, therefore trivially including TC^0. More interestingly, for $k = o(n)$, one can consider $\mathsf{bPTF}^0[k]$ as AC^0 circuits with additional computational power, including access to majority gates of small fan-in.

With these definitions, we are ready to present our first main result: for the parity halving problem [27], we prove average-case correlation bounds against $\mathsf{bPTF}^0[k]$, establishing a separation of QNC^0 from $\mathsf{bPTF}^0[k]$.

Theorem 4.2.1 (Informal, see Theorem 4.6.1 and Theorem 4.6.2)**.** *For every large enough input size $n \in \mathbb{N}$, there exists a relation $R \subset \{0,1\}^n \times \{0,1\}^m$ with $m = \mathcal{O}(n \log n)$ such that for any $x \in \{0,1\}^n$, a QNC^0 circuit consisting of $o(n^2)$ gates can output a $y \in \{0,1\}^m$ such that $(x, y) \in R$ with certainty. Conversely, the size s of any $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit of depth d that computes any valid y is lower bounded as shown in the table below. Here $/\mathsf{rpoly}$ means the circuit is randomised, i.e., has access to polynomially many random input bits.*

As previously mentioned, threshold circuits serve as a useful theoretical model for neural networks, and in Section 4.8.2, we highlight how biased polynomial threshold circuits of constant depth similarly exhibit a strong relationship with neural networks. Building on the later class, our results show that $\mathcal{O}(n^{1/d})$-biased threshold circuits of constant-depth, which have access to polynomially many random bits, require

[2] We note that the common definition in the literature takes $P : \mathbb{R}_2^n \to \mathbb{R}$, and sets $f(x) = \frac{1}{2}(1 + \mathrm{sgn}(P(x)))$. On the other hand here we are interested in polynomials of degree at most k over $\mathbb{F}_2^n$, with threshold behaviour determined by the bias k.

$\mathsf{bPTF}^0[k]/\mathsf{rpoly}$	$k = \mathcal{O}(1)$ $(\equiv \mathsf{AC}^0/\mathsf{rpoly})$	$k = n^{1/(5d)}$
Exact hardness	$\Omega\left(\exp\left(\left(\frac{\sqrt{n}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)\right)$	$\Omega\left(\exp\left(\left(\frac{n^{3/10}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)\right)$
Avg. hardness	-	$\Pr[\mathsf{Suc}] \leq \frac{1}{2} + \exp\left(-\Omega\left(\frac{n^{3/5-\mathcal{O}(1)}}{(\log s)^{2d-1}}\right)\right)$

Table 3: Size lower bounds for the circuit classes $\mathsf{bPTF}^0[k]$. By "exact hardness" we mean that the classical circuit is required to produce a valid output string with certainty on all inputs, and in this setting the expressions bounds the size s of the respective $\mathsf{bPTF}^0[k]$ circuits.

superpolynomial size to solve a relational problem that is solved efficiently by constant depth bounded fan-in quantum circuits. Observing that even a *single gate* of a $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit with $k = \mathrm{polylog}(n)$, would also require superpolynomial (i.e. $\Omega(n^{\mathrm{polylog}(n)})$) size AC^0 circuits to emulate it, and since $\mathsf{NC}^0 \subsetneq \mathsf{AC}^0$, we obtain $\mathsf{NC}^0 \subsetneq \mathsf{AC}^0 \subsetneq \mathsf{bPTF}^0[k]$ for $k = \omega(\log n)$ [115]. This is thus the separation of QNC^0 against the largest classical circuit class to date, obtained unconditionally. Importantly, we achieve a super-polylogarithmic value for k, demonstrating that Theorem 4.2.1 extends separations against AC^0 from prior work to this genuinely new class. Furthermore, we remark that our bounds exhibit the largest possible bias value for the type of relation problems considered. This is evidenced by the fact that for any larger k, $\mathsf{bPTF}^0[k]$ can compute parities, and thus *can* solve the relational problem with high probability and small circuit size. This suggests that a novel approach will be needed for future work to improve our results.

Additionally, we consider the case when classical circuits are required to solve the relation with certainty on all inputs, equivalent to what is achievable by quantum circuits. We establish a new separation against $\mathsf{AC}^0/\mathsf{rpoly}$ for this scenario, described as 'exact hardness' in Theorem 4.2.1, demonstrating the tightest bounds known to date. Using this result we present a resource estimation analysis in Section 4.8.1, showing that a quantum advantage could be realized using quantum devices that require orders of magnitude fewer qubits than prior work for both values of k in Table 3.

Our second contribution involves constructing a new family of relation problems based on computing Mod p for each prime $p > 2$, designated as ISMR problems (see Definition 62). With these, we extend all previous results from qubits to higher dimensions. Throughout this paper, we refer to qudit systems of prime dimension p as 'qupits'.

Theorem 4.2.2 (Informal, see Theorem 4.6.4)**.** *For every prime $p \in \mathbb{N}$ and large enough input lengths $n > n_p^* \in \mathbb{N}$, there exists a relation $\mathcal{R}_p \subset \{0,1\}^n \times \{0,1\}^m$ with $m = \mathcal{O}(n \log(n)^p)$, such that there is a QNC^0 circuit over $(\mathbb{C}^p)^{\otimes n}$ (i.e., n 'qupits') consisting of $o(n^2)$ gates that has constant correlation with $\mathcal{R}_p^m$. In contrast,* any *polynomial size $n^{1/(5d)}))/\mathsf{rpoly}$ circuit has exponentially small correlation* $\exp\left(-\Omega\left(n^{3/5-\mathcal{O}(1)}\right)\right)$ *with $\mathcal{R}_p^m$.*

Previous studies have suggested that qudit non-local games with significant quantum-classical winning probability gaps could give rise to computational separations between classical and quantum circuit classes [117]. However, such separations were not explicitly demonstrated. By proving Theorem 4.2.2, we establish the existence of such a non-local game for each prime dimension, which might be of independent interest (see Section 4.6.2). We use these non-local games to explicitly construct relational problems and quantum circuits that solve them with high correlation. Our correlation measure (see Definition 32) is a generalization of the usual correlation between Boolean functions to functions mapping $\{0,1\}^n$ to

the cyclic groups $(\mathbb{F}_p, +)$, taking into account 'how wrong' an output is: if $f(x) = 2$, then $\tilde{f}(x) = 4$ should intuitively be a worse guess than $\tilde{f}(x) = 3$ for any $f : \{0,1\}^n \mapsto \mathbb{F}_p$. Thus, the result above is stated in terms of the correlation between the relation problem $\mathcal{R}_p^m$ (defined formally in Definition 62) and bPTF$^0(k)$/rpoly circuits. For the $p = 3$ case, the correlation function directly provides a lower bound on the success probability of the classical circuits. We explicitly calculate these in Theorem 4.6.5 and Corollary 8.

Remark. The new relation problems $\mathcal{R}_p$ in Theorem 4.2.2 are still binary problems that are natural to the bPTF$^0[k]$ class. They do not simply enforce hardness via arithmetic operations over other prime fields, and Razborov-Smolensky type lower bounds are not directly applicable to this setting for the classical binary gate sets. We also emphasize that this separation is achieved in the average-case hardness setting with uniform input distributions over binary inputs. Technically, we prove novel bounds on the classical winning strategies for qupit non-local games under *non-uniform input distributions*, handle the qupit-to-binary encoding, and ensure the use of uniform (rather than biased) random restrictions over the binary input.

We note that if we were to consider an infinite gate set such as $\{\text{all single-qubit gates},\ \text{CNOT}\}$ as assumed in [263, 78, 77], we could realize all our qupit circuits with qubit circuits in the noise-free setting [108, 107]. However, the qupit QNC^0 circuits we consider are defined over a standard, minimal *finite* qupit gate set. This allows for constant-depth realizations with the gate sets available on standard quantum computing devices and fault-tolerant implementations, which are impossible if infinite gate sets are required. These explicit separations not only clarify the theoretical landscape but also hold practical relevance, as many quantum computing platforms naturally operate in higher dimensions [110, 116]. Therefore, this design choice could provide a fair and effectively demonstrable quantum advantage[3].
Additionally, in this setting, any finite minimal gate set over a certain qupit dimension, if attempting to solve one of the ISMR problems of another prime dimension, would require decomposition into its own native gates. In this context, relying on Solovay-Kitaev-type decompositions, the gates would very likely necessitate log-depth decompositions for approximation. Thus, under this hardware-realistic definition of qupit QNC^0 circuits, we raise the following conjecture.

Conjecture 3. *For each pair of distinct primes p and q, there exists a relation problem $\mathcal{R}_p$ such that $\mathcal{R}_p \notin$ QNC^0 when* QNC^0 *is defined over a minimal universal gate set for qupits of dimension q.*

We thus conjecture that the ISMR problems could be for the qupit QNC^0 circuit classes what the Razborov-Smolensky type modular problems are for the AC$^0[p]$ circuit classes.

Finally, quantum systems are unavoidably affected by noise, and error correction is a much more complex process in the quantum realm, bringing into question the *robustness* of a computational advantage under noise. This question is of significance to both theory and practice, especially as we navigate the NISQ era. Even for very powerful quantum computational models, the introduction of noise often dramatically diminishes computational advantages that they may offer over their classical counterparts: for instance, recent work shows that even small constant error rates result in a collapse of the power of multi-prover interactive proofs where the provers share entanglement (MIP*) from RE to multi-prover interactive proofs without shared entanglement (MIP) [264]. Our third main result is to prove that all our separations are robust to noise: even if all steps of the quantum computation are affected by local stochastic noise, there is a family of modified relation problems that these noisy quantum circuits, when provided with logical magic states, can solve, but is hard for noiseless bPTF$^0[k]$ circuits.

[3] Note that parallel repetition techniques [27, 34] could be applied to Theorems 4.2.1 and 4.2.2, to amplify the advantage, facilitating experimental demonstrations.

Theorem 4.2.3 (Informal, see Theorem 4.7.1)**.** *For every prime $p \in \mathbb{N}$ and large enough input lengths $n > n_p^* \in \mathbb{N}$, there exists a relation $\mathfrak{R}_p \subset \{0,1\}^n \times \{0,1\}^m$ with $m = \mathcal{O}(n \cdot \text{poly}(\log n))$ based on $\mathcal{R}_p$, such that for local stochastic noise with physical error rate below a constant threshold, there is a noisy qupit* $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ *circuit of all-to-all connectivity that has constant correlation with $\mathfrak{R}_p$. In contrast,* any *(even noiseless) polynomial size* $\mathsf{bPTF}^0[n^{1/(5d)}]$ *circuit has exponentially small correlation* $\exp\left(-\Omega\left(n^{3/5-\mathcal{O}_d(1)}\right)\right)$ *with $\mathfrak{R}_p$.*

This result improves upon prior work in three main ways. First, it accounts for noise while addressing non-Clifford gates, going beyond the framework for noise-robust separations introduced in [35], by including logical T-type magic states. Second, it generalizes shallow-depth computational separations and the error-correction mechanisms to show that they are robust to arbitrary prime qudit dimensions. Third, these constitute the state-of-the art in unconditional separations between noisy QNC^0 and the largest classical circuit classes to date. Complementarily, as a corollary, it establishes separations with NC^0 for each prime qudit dimension, potentially enabling near-term quantum advantage experiments due to the favorable error correction properties that have been demonstrated for qudits.
We remark that the difference in the correlation bounds between Theorem 4.2.3 and Theorem 4.6.4 lies in the exponent of the $\log s$ term, resulting from the additional overhead for error correction, which nevertheless does not undermine the exponential difference between the classical and quantum correlation achieved by the respective circuit classes with $\mathfrak{R}_p$. Finally, we also note that although Theorem 4.2.3 is stated with logical T-type magic states as advice, we also prove a new separation for qubits without advice (Corollary 9).

4.2.2 Technical contributions

In this section we highlight the crucial new ingredients we have used and devised for the development of our main results, some of which may also be of independent interest.

Switching lemmas. We prove a novel multi-switching lemma (Lemma 25) for multi-output $\mathsf{bPTF}^0[k]$ circuits, using inductive methods to significantly refine the parameters beyond the prior work of [115]—see Section 4.3.4 for a further discussion. Utilizing this lemma, we establish depth reduction (Lemma 30) for multi-output $\mathsf{bPTF}^0[k]$ circuits, aimed at addressing relation-type problems. These methods have not only enabled us to substantiate our lower bounds in Theorem 4.2.1 and Theorem 4.2.2 but have also facilitated a precise determination of the maximum value of the bias parameter k for $\mathsf{bPTF}^0[k]$ while elucidating the relationship between geometrical hardware constraints and the possibility for quantum advantage. Moreover, these lemmas in principle apply to circuit models of some neural network architectures with practical string-to-string mapping applications (see Section 4.8.2), and we surmise that they may have applications to search problems beyond what we consider.
Note that the essence of switching lemmas has remained unchanged since their introduction in the 80s, primarily addressing separations within the AC^0 class. The significance of our work is underscored by the fact although Kumar's introduction of a class that extends AC^0 offered new possibilities, the switching lemmas in that paper are insufficient for relation-type problems, and in fact are not tight enough to yield separations between QNC^0 and AC^0.

Decision tree depth bounds. Our second technical contribution builds on the first. More concretely, our new multi-switching lemma enables the reduction of the initial $\mathsf{bPTF}^0[k]$ circuit to an object we refer to as a forest line. This is a set of decision trees, where each one is responsible for computing the outcome bit of the relation problem for the set of variables left "alive" by random restrictions applied to the input. We leverage this mathematical structure for our new exact hardness results in Theorem 4.2.1. The novelty lies in counting the super-exponential number of valid outcome strings for the chosen relation problem, through the Algebraic Normal Forms [265] that represent the valid 'deterministic' solutions to the problem. Using

this concept, we first establish a lower bound on the number of degree two terms in the representation of any valid solution (Lemma 33). Then, we upper bound the number of terms each local decision tree can contribute to our remaining forest line, thus limiting the expressiveness of these decision trees and their efficacy in solving the problem correctly (Lemma 34). This in turn reflects the initial circuit's ability to solve the relation problem (Theorem 4.2.1). This approach captures the classical-quantum circuit separation using ideas standard in theoretical computer science, sidestepping the need for non-local or non-contextual games, which have traditionally been indispensable in prior research.

Non-local games in higher dimensions. Our next technical contribution is to constructively prove that, for all prime qudit dimensions, there exists a family of generalized XOR non-local games, which we have coined "Modular XOR games" (see Definition 39), where classical strategies, at best, achieve exponentially lower success probabilities compared to their quantum counterparts as the number of parties increases (Lemma 38). We build on this construction by revisiting the switching lemma, leveraging the locality of the resultant forest of decision trees therein to demonstrate a separation of quantum circuits over *qudits* from $\mathsf{bPTF}^0[k]$, since we show that the former can efficiently solve a family of relational problems defined via our new family of non-local games. Our method elegantly incorporates the ability to manage biased input distributions for non-local games, handling the necessary dit-to-bit mappings to ensure that we use unbiased random restrictions for $\mathsf{bPTF}^0[k]$. This approach allows us to obtain average-case $\mathsf{bPTF}^0[k]$ hardness with respect to the uniform distribution on binary inputs for the relation problems that demonstrate the separation. Additionally, it sidesteps the task of computing explicit success probability bounds for each game, while still achieving comparable separations. This obstacle has been emphasised in prior research, which predominantly relied on established quantum-classical distinctions in non-local games [18, 27]. Also, few developments beyond the conventional qubit setting were considered or achieved previously [118].
We remark that we have also integrated into the correctness proof of the candidate quantum circuit that solves the qupit relational problem a technique to efficiently describe the support of standard measurement outcomes of local unitary (LU)-equivalent generalized GHZ states with a dependence on the phases (Lemma 44). This allows us to incorporate a correction function for the output string, and compute the success probability of the quantum circuits in a clean way.

Fault tolerance. Our final technical contribution extends the fault-tolerant regime from qubit Clifford to qupit non-Clifford constant-depth quantum circuits with logical T-type magic states (see Figure 24), which can be affected by local stochastic noise (Lemma 45). We show that the qupit surface code, equipped with the hard decision renormalization decoder, enables fault-tolerant implementation of these quantum circuits. This includes the extension of single-shot logical state preparation for qupits. Specifically, we extend a 3D block construction of [124] to higher dimensions, demonstrating that a specific measurement pattern results in a reduced state that is a logical GHZ_2 state, up to local Clifford corrections (Lemma 51). Furthermore, previous work used the minimum weight perfect matching decoder, which performs poorly in higher dimensions. By using a different decoder, we show that we can still perform corrections and recover the desired states with exponentially high confidence (Corollary 10).

4.2.3 Related work

A few unconditional separations between larger classes of quantum and classical circuits have also appeared in prior work [76, 77, 78, 96]. However, the majority of these involve more complex ingredients, such as the use of quantum fan-in and fanout gates in combination with gate sets that offer infinite precision, rendering them impractical for consideration as true constant-depth circuit candidates. Although [96] later dropped the requirements for quantum fanout and infinite gate sets, the proposed quantum circuits still refer to multi-qupit controlled gates, which are in early development and require further research to

assess their potential for fault-tolerant implementation.
This challenge shifted the focus towards the study of shallow-depth quantum circuits that can, in principle, be implemented on near-term hardware. Various extensions of the first unconditional separation between constant-depth classical and quantum circuits, established by Bravyi et al., have been achieved in the last five years. These results are intimately related, sharing similar quantum solutions all based on non-local games, but employing different techniques to prove classical hardness. Here, we provide a brief overview of the most relevant results related to our work, summarised concisely in Table 4.

Bounded fan-in and worst-case. We begin by noting that the results of [18] were achieved by using a non-local game previously described by [69]. However, unlike the usual definition of non-local games, this extended GHZ game allowed for some communication between the parties. This additional communication was utilized to construct a quantum circuit whose outputs are based on winning strategies for this game, while also linking the circuit's outcome to the solution of the 2D Hidden Linear Function (2D HLF) problem. The latter is a non-oracular version of the Bernstein-Vazirani problem [70], with the difference that the linear function $f(x) = \mathbf{a} \cdot \mathbf{x}$ to be learnt is not accessed directly via a quantum oracle, but is 'hidden' by a quadratic form. The authors use lightcone arguments that exploit bounded fan-in to show that the efficacy of classical NC-type circuits in solving the 2D HLF problem is limited by the efficacy of classical strategies in winning the incorporated non-local game. Thus, it is guaranteed that the worst-case performance of such classical circuits is unconditionally inferior to the quantum solution.

Average-case and unbounded fan-in. Subsequent work extended this unconditional separation in two major directions: to average-case hardness or correlation bounds [33, 27], and to the class AC^0 of unbounded fan-in AND and OR gates [27]. The latter work achieved this through a task dubbed the Parity Halving Problem (PHP), which requires producing an outcome string with a parity equal to the result of applying the Mod 4 operation to the input string. The authors applied a different lower bound technique to the PHP, namely (multi-)switching lemmas, to relate AC^0 circuits back to NC^0 circuits, which allowed locality arguments to be ported to the former. Switching lemmas describe how the function or relation computed by an AC^0 circuit simplifies when some of its input literals are randomly set to 0 or 1. Their main result was that classical solutions to the PHP require superpolynomial size AC^0 circuits. For the quantum upper bound, the authors utilized poor-man's cat states, which are local unitary (LU)-equivalent to n-party GHZ states, to efficiently solve the PHP with certainty on all inputs. Furthermore they showed an NC^0 reduction from the PHP to 2D HLF, proving that the new average-case separation applies to the HLF problem as well.

Noise resilience. Traditional quantum fault-tolerant schemes have a significant depth overhead, and might not be compatible with earlier quantum advantage proposals. Driven by this understanding, [35] showed that for the popular local stochastic noise model, errors occurring in a given circuit layer could be corrected with constant-depth circuits. They showed that transversal Clifford gates could be realized by a constant-depth circuit equipped with single-shot state preparation and readout, thereby proving that noise resilience could be achieved within a constant depth. For the Magic Square Problem (MSP) (inspired by the Mermin-Peres magic square [71, 72]) they exhibited a QNC^0 circuit in a 1D configuration that has a quantum advantage against NC^0, and consequently a 3D configuration error-correcting code using the previous elements that offers robustness against noise while retaining the advantage against NC^0.
[37] recently proved unconditional quantum advantage of noisy QNC^0 circuits over AC^0. Focusing on the task of possibilistic simulation of single-qubit Clifford gate teleportation (TELEP)—a task previously shown to be challenging for NC^0 [73]—the authors extended the earlier separation to AC^0, while introducing a 1D-implementable quantum circuit, enabling 3D fault-tolerant construction as shown in [35].

Finally, alternative approaches have considered different types of tasks and separations, including interactive protocols [74, 36] and sampling problems [38]. Some studies have also applied geometric restrictions to classical circuits [75] to obtain separations on the depth.
As a closing remark, we recall that while all these separations between QNC^0 and classical shallow-depth circuit classes are unconditional, the quantum circuits in question can still be simulated classically in polynomial time, reiterating that the notion of separation here is more granular.

	Problem	Advantage against	Geometry	Higher dimensions	Avg.-case hardness	Noise resilience
[18]	2D HLF	$\mathsf{NC}^0/\mathrm{rpoly}$	2D	✗	✗	✗
[27]	PHP	$\mathsf{AC}^0/\mathrm{rpoly}$	2D	✗	✓	✗
[35]	MSP	$\mathsf{NC}^0/\mathrm{rpoly}$	1D	✗	✓	✓
[37]	TELEP	$\mathsf{AC}^0/\mathrm{rpoly}$	1D	✗	✓	✓
Our work	ISMR	$\mathsf{bPTF}^0[k]/\mathrm{rpoly}$	2D, 3D, ...	✓	✓	✓

Table 4: Summary of related work on unconditional separations of quantum and classical shallow-depth circuits. Please refer to the main text of Section 4.2.3 for problem abbreviations.

4.2.4 Discussion and Outlook

[266] showed the first lower bounds for depth two circuits of linear threshold functions with small bias, for the *inner product modulo* 2 function. Gates computing polynomial threshold functions as in Definition 19 may be activated not only by a linear function $\mathbf{a} \cdot \mathbf{x} - k$ of the input $\mathbf{x} \in \{0,1\}^n$, but also by higher-degree polynomials [267]. This increases the expressivity and the computational capabilities of such gates but renders almost all known lower-bound techniques ineffective. TC^0 can compute unbounded but fixed degree PTFs we consider, and the best results known for this class are barely superlinear $\Omega(n^{1+\epsilon_d})$ lower bounds on the number of wires [268]. Even for fixed depth equal to 2, only $\omega(n^{3/2})$ gate or $\omega(n^{5/2})$ wire lower bounds have been proven [269]. While our results are related to TC^0 and more general PTF circuits, and use similar higher-degree polynomials, they are distinct due to the limitations on the gates in each class $\mathsf{bPTF}^0[k]$ introduced via the bias parameter k, rendering the classes and results incomparable.
In the context of noise tolerance, our work expands the separation to larger classical classes and provides explicit quantum circuits with corresponding bounds on the correlation for all qudit prime dimensions. Our Theorems 4.2.2 and 4.2.3 necessitate all-to-all connectivity, which may require additional SWAP gates in certain device architectures, thus increasing the circuit depth. However, qudit devices with native all-to-all quantum gates are among the leading quantum hardware devices [127]. Importantly, all classical circuit classes considered for the separations—NC^0, AC^0, and $\mathsf{bPTF}^0[k]$—also inherently require all-to-all connectivity, ensuring a fair comparison. Furthermore, our lower bound techniques could potentially achieve separations with noisy geometrically local circuits when problems of a distinct nature are under consideration [37].
Our work raises further intriguing questions in this active research area. Having initiated the consideration of error-corrected qudit circuits for unconditional quantum separations, we question whether previous results, shown for high-dimensional error-correcting codes with improved parameters [270], could facilitate experimental demonstrations. Additionally, we have reiterated the interest of magic states in the constant-depth fault-tolerant regimes [65, 271]. Although we conjecture that magic state factories should not be parallelizable to the level of being in the QNC^0 circuit class, it is of great interest to understand what the simplest circuit class capable of realizing these processes is, as well as introduce the capacity for adaptive corrections during error-corrected circuit execution. This might extend the set of unconditional

separations beyond the currently known limits and lift larger conditional separations to the simplest fault-tolerant quantum circuit classes [272]. In parallel to these questions, it would be important to consider more complex and alternative noise models [273] to better align with specific practical quantum computing architectures. Following the same line of research, it would also be interesting to understand if there is a complexity phase transition for some values of the bias parameter k and the noise strength.
On a different note, our family of relational problems could generate Razborov-Smolensky type separations between qupit QNC^0 circuit classes, thereby providing a foundational understanding of qudit computations in constant depth circuit classes. Additionally, we find that $\mathsf{bPTF}^0[k]$ circuits for $k = n^{1/d}$ can solve problems that $\mathsf{AC}^0[p]$ circuits cannot, such as $\mathsf{Mod}\ p$ operations over $n^{1/d}$ bits [19]. This strengthens the potential to establish, through the family of ISMR problems related to $\mathsf{Mod}\ p$, unconditional separations between qupit QNC^0 circuits and $\mathsf{AC}^0[p]$ for any prime p, as previously conjectured [27, 74]. Thus, our results might simplify existing conditional results by eliminating the need for interactivity or advice in proving such separations. Additionally, our work achieves near-optimal separations concerning the circuit class under study. This suggests that to attain conceptually different and relevant separations, such as with TC^0, we must identify harder problems that are solvable in this setting since the ISMR problems are solvable by TC^0 [96]. This likely will necessitate the consideration of more advanced QNC^0 circuits, such as constant-depth adaptive MBQC circuits and more elaborate quantum circuit designs.

Connections between biased PTF circuits and other complexity classes. As a final note, we would like to highlight some additional interest in the $\mathsf{bPTF}^0[k]$ class and the new multi-switching lemmas. Well-known prior work relates uniform AC^0 to the polynomial hierarchy PH by associating the logical formulas that an AC^0 circuit represents with the quantified polynomial-time predicates describing the polynomial hierarchy [274]. In particular, computational separations against AC^0 have been used to establish relativized separations against the polynomial hierarchy [20], and depth hierarchy results within AC^0 have been used to establish separations between the levels of PH [275]. Similar studies have been conducted between threshold circuits and the counting hierarchy [276], albeit with less success due to the lack of strong lower-bound techniques for classes such as TC^0. Interestingly, $\mathsf{bPTF}^0[k]$ for $k = \omega(\log n)$ describes a computational hierarchy that likely extends beyond the polynomial hierarchy and has non-trivial overlaps with the counting hierarchy. Thus, new lower-bound techniques might enable oracle separations between these hierarchies. Additionally, this opens avenues for the study of unconditional separations between $\mathsf{BQLOGTIME}$ and $\mathsf{bPTF}^0[k]$, which, based on our previously raised conjectures, could enable oracle separations between BQP and computational classes beyond the PH [277, 278].

4.3 Overview of Proofs and Techniques

In this section, we provide a concise overview of the proofs for our main results. We aim to strike a balance between rigor and accessibility by offering an intuitive discussion of the classical and quantum techniques employed. We will also highlight the key improvements we have made over existing work in this area. We begin by introducing the family of computational problems that we will be addressing.

Inverted Strict Modular Relation Problems (ISMRP). In this paper, we introduce a new class of relation problems that are related to the 2D hidden linear function problem from [18] and generalize the parity halving problem studied in [27]. We leverage this class of problems to establish new unconditional separations between QNC^0 and the $\mathsf{bPTF}^0[k]$ classes of biased threshold circuits.

Definition 20. *For any positive integer p, the Inverted Strict Modular Relation Problem* (ISMRP) *with input and output lengths n and m respectively, denoted as $\mathcal{R}_p^m : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$, is defined as follows. For*

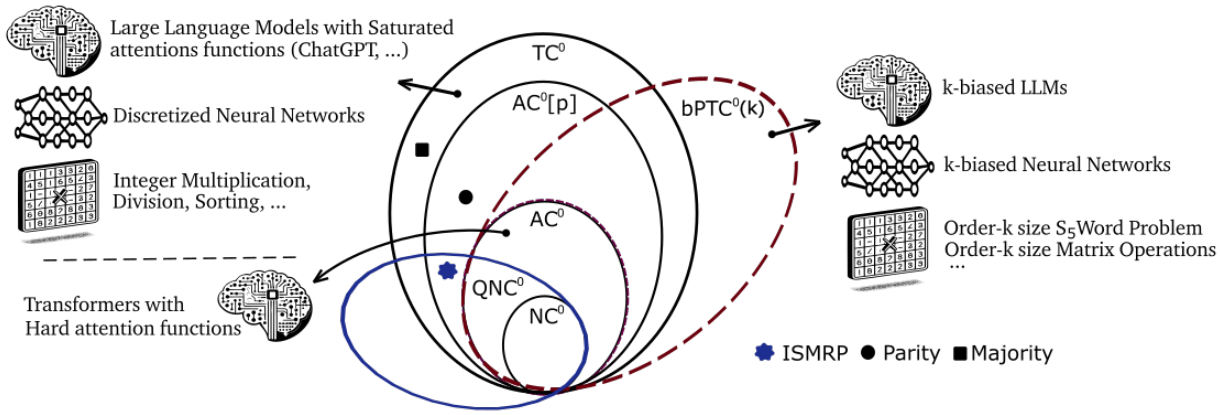


Figure 14: This illustration depicts the primary classical and quantum circuit classes analyzed in this paper, highlighting the inclusion of relevant problems that characterize the relationships and distinctions between these classes. Notably, the Parity function is computable in $\mathsf{AC}^0[2]$ but not in AC^0, while the Majority function is not computable in $\mathsf{AC}^0[p]$ but is in TC^0. We also indicate the ISMR family of problems introduced in this work (see Definition 62), which differentiate QNC^0 from $\mathsf{bPTF}^0[k]$ with $k = \omega(\log n)$. The latter class, which strictly includes AC^0, can solve instances of NC^1-complete problems for super-logarithmic size input strings. It can also solve computational problems conjectured to be strictly beyond NC^1. This suggests a potential non-trivial overlap with TC^0. We also list more general and practically relevant computational processes known to be computable within each class, particularly neural network-type circuits.

Transformers, known for their substantial use as crucial components of modern large language models (LLMs), can be simulated by AC^0 circuits when the attention mechanism in the transformer is restricted to hard attention functions [2, 3]. Furthermore, TC^0, recognized as the standard class for theoretically representing discretized neural networks [4, 5], can simulate LLMs under realistic operational conditions [6], such as restricting the precision of internal variables and auto-regression or the number of time-steps the model is run for at inference. This highlights a key distinction between AC^0 and TC^0, in that the latter class has enhanced expressivity for a wide range of algebraic operations, such as integer multiplications, division, and sorting [7, 8]. The $\mathsf{bPTF}^0[k]$ circuits that we study in this work are capable of simulating a variety of variants of standard neural networks. This class of circuits also supports attention functions that are likely to be beyond TC^0 and may accommodate substantially different LLM architectures. We refer the reader to Section 4.8.2 for a more detailed discussion on these points.

any $x \in \mathbb{F}_2^n$ *such that* $|x| \bmod p = 0$*,*

$$\mathcal{R}_p^m(x) := \left\{ y \;\middle|\; y \in \mathbb{F}_2^m :\ |y| \bmod p = -\left(|x|/p\right) \bmod p \right\}, \tag{4.3.1}$$

where $|x| = \sum_{i=1}^{n} x_i$ *is the Hamming weight of the* n*-bit string* x*.*

The relational problem $\mathcal{R}_p^m$ is defined only for inputs x whose Hamming weight is a multiple of p, say $|x| = kp$ for some $k \in \mathbb{N}$. For any such valid input, the set of valid outputs is a set of bitstrings with length specified by the parameter m and Hamming weight equal to the additive inverse of the factor k modulo p, i.e. $|y| = -k$ Mod p. For $p = 2$, $\mathcal{R}_2^m$ is exactly the parity halving problem defined in [27]. By thus generalizing PHP to $\mathcal{R}_p^m$ we will see that we can define and analyze new non-local games over qupits, allowing us to extend separation results to quantum circuits over higher dimensional systems.
For this family of problems, we provide separate analyses for quantum shallow circuits over qubits ($p = 2$) and qupits ($p > 2$ prime). In the $p = 2$ case, we demonstrate that qubit QNC^0 circuits can solve the $\mathcal{R}_2^m$ problem exactly and with certainty, and prove both exact-case and average-case hardness results for $\mathsf{bPTF}^0[k]$, in the form of lower bounds on the circuit size for a fixed depth, thereby proving Theorem 4.2.1. We give a sketch of our proof and an overview of the technical tools we developed for it in Section 4.3.1.
We will then demonstrate the average-case correlation bounds against $\mathsf{bPTF}^0[k]$ for all other $\mathcal{R}_p^m$ problems for $p > 2$, as claimed in Theorem 4.2.2. We complement these classical lower bounds with the respective quantum upper bounds, using the corresponding qupit QNC^0 circuits for each value of p. The correlation with the ISMR problems, in this case, has a more complex interpretation than in the qubit case (see Section 4.6.2 for details), and we show that our qupit QNC^0 circuits achieve $\Omega(1)$ correlation while polynomial size $\mathsf{bPTF}^0[k]$, with $k = n^{1/(5d)}$, circuits can achieve only $O(2^{-n})$ correlation.

4.3.1 Proof of Theorem 4.2.1: Separation for the qubit case

This subsection is divided into two parts: average-case and exact hardness. We start with average-case hardness, to facilitate easier understanding of the proof techniques, and lead naturally into the exact case hardness. Our results for both average and exact cases rely on the quantum solution (QNC^0 circuit) for the ISMRP $\mathcal{R}_2^m$, and diverge in how they handle the classical circuits—via upper bounds on correlation for the former, and a counting argument for the latter.

Average-case hardness

We start by constructively proving the existence of QNC^0 circuits that solve the $\mathcal{R}_2^m$ problem with certainty, while producing output strings that are of length $m = o(n^2)$[4]. Specifically, we construct QNC^0 circuits with 3D connectivity in Lemma 36 and all-to-all connectivity in Corollary 5. We do this using the Measurement-Based Quantum Computation (MBQC) paradigm, which allows for shallower quantum circuits and helps us improve the resource estimates as presented in Section 4.8.1.
Given this candidate problem that QNC^0 circuits can solve, the second and most challenging aspect pertains to establishing upper bounds for the probability with which $\mathsf{bPTF}^0[k]$ circuits of size s and depth d can solve the same problem on typical inputs of length n. This is the most technical part of the proof, as no previous lower-bound techniques exist for multi-output $\mathsf{bPTF}^0[k]$ circuits.
Moreover, the development of depth reduction lemmas, which sequentially reduce a $\mathsf{bPTF}^0[k]$ circuit to a decision forest, necessitates multi-switching lemmas. These lemmas enable the reduction of a single layer of $\mathsf{bPTF}^0[k]$ circuits to a decision forest. However, they require a sufficiently high probability guarantee that a layer of a $\mathsf{bPTF}^0[k]$ circuit will reduce to decision forests under random restrictions. The existing lemmas of this type, as demonstrated in [115], do not meet the necessary conditions. Consequently,

[4] Note that solutions to the problem with outputs of length $m = \mathcal{O}(n^2)$ would be trivially achievable even by NC^0 circuits

we develop new tools including a new multi-switching lemma for $\mathsf{bPTF}^0[k]$ circuits with higher success probabilities, as follows.

Lemma 12 (Informal, see Lemma 25)**.** *Let $\mathcal{F} = \{F_1, F_2, F_3, \ldots, F_m\}$ be a single $\mathsf{bPTF}^0[k]$ circuit layer with gates that have maximal fan-in of w over inputs in $\mathbb{F}_2^n$. Then, for sufficiently large l and small probability p it reduces under p-random restrictions to a decision forest $\mathsf{DT}(l)^m \Rightarrow \mathsf{DT}(t)$, with a global decision tree of depth t of which each leaf is a set of m decision trees of depth l. Formally,*

$$\Pr_{\rho \in R_p}[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(l)^m \Rightarrow \mathsf{DT}(t)] \leq m \cdot 2^k (80wp)^t. \tag{4.3.2}$$

This new multi-switching lemma facilitates the creation of a depth reduction lemma, which refers to our Lemma 29. More importantly, this particular lemma enables sequential depth reductions of the circuits from depth d to a circuit of depth $d-1$ with high probability. It effectively allows the reduction of constant-depth $\mathsf{bPTF}^0[k]$ circuits to decision forests, especially in cases where there are multiple output bits as follows,

Lemma 13 (Informal, see Lemma 31)**.** *Let f be a $\mathsf{bPTF}^0[k]$ circuit with m outputs and inputs in $\mathbb{F}_2^n$, of size s and depth d. Then, for a probability p parameterized by s, d, k, t and q and we have that*

$$\Pr_{\rho \sim R_p}[f\lceil_\rho \notin \mathsf{DT}(q-1)^m \Rightarrow \mathsf{DT}(2t-2)] \leq s \cdot 2^{-t+k}\,. \tag{4.3.3}$$

The previous lemmas enable us to reduce any initial $\mathsf{bPTF}^0[k]$ circuit to a decision forest. Then, a second use of random restrictions reduces this decision forest to m independent decision trees, each computing one of the outcome bits. At this point, these techniques based on random restrictions have us allowed to reduce the circuits under consideration while the computational problem kept the same structure and hardness over the variables that were kept alive. From here, we explore this asymmetry using standard locality arguments, such as the lightcone techniques, we can relate the efficiency with which these decision trees solve the initial problem to the efficiency with which classical strategies solve a Mermin's multi-player game [279] that is implicitly integrated into the $\mathcal{R}_2^m$ problem over the variables that are preserved "alive" through the random restriction. By combining the probabilities of successful reductions with the known upper bounds for classical strategies solving this specific non-local game, we obtain our average-case hardness results for $\mathsf{bPTF}^0[k]$ circuits,

Corollary 3 (Informal, see Corollary 4)**.** *For large n, any $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit of depth $d \geq 4$, size $s \leq \exp\left(n^{1/(2d-2)}\right)$ and $k \leq n^{1/(5d)}$ solves the $\mathcal{R}_2^m$ with an efficiency bounded by $\frac{1}{2} + \exp\Big(-\Omega\Big(\frac{n^{2-o(1)}}{m^{1+o(1)}(k\cdot\log(s))^{2d}}\Big)\Big)$.*

This combination of the previous results still requires careful consideration of all parameters to prove a quantum advantage. There is an interplay between the amount of bias k in the $\mathsf{bPTF}^0[k]$ circuits and the dimensionality of the QNC^0 circuits, while also optimizing all remaining degrees of freedom resulting from the proof techniques. In particular, we establish the average-case hardness, as stated in Theorem 4.6.2, by combining Lemma 36 and Corollary 5, which pertain to our quantum solutions, with Corollary 3.

Exact-case hardness

We start by clarifying that exact-case hardness refers to the capacity of the circuit to solve the proposed problem with a probability equal to 1 for all inputs. This is different from worst-case hardness, where we choose a subset of inputs and demonstrate that computing these with a certain probability provides a certain level of hardness. In particular, from the previous average-case bound in [27], one can determine a

lower bound on the size of an AC^0 circuit that solves this problem in the exact-case hardness setting. Similarly, the same follows for $\mathsf{bPTF}^0[k]$ from our average-case hardness result referenced in Theorem 4.6.2. However, these bounds are not tight, as we have demonstrated with our exact hardness bound, which implies that even larger AC^0 and $\mathsf{bPTF}^0[k]$ circuits are required to solve the problem exactly. Specifically, this allows us to ascertain that quantum-classical advantages exist with input sizes that are orders of magnitude smaller. Additionally, to the best of the authors' knowledge, this is the first exploration of this particular hardness setting.

To achieve this result we explore some specific proprieties of the $\mathcal{R}_2^m$ problem. The particular property is that the XOR of all the outcome bits has to equal a specific Boolean function, known as the function $\mathsf{LSB} : \{0,1\}^n \mapsto \{0,1\}$ which outputs the second least significant bit of the binary representation of the Hamming weight of the input $|x|$. More precisely, each output bit produced by a $\mathsf{bPTF}^0[k]$ circuit can be viewed as a distinct Boolean function $f_i : \{0,1\}^n \mapsto \{0,1\}$, and these have to fulfill the same property. Thus, by additionally leveraging the algebraic normal form (ANF) representation we have that the XOR of respective representation of all these Boolean functions must equal the one of the $\mathsf{LSB}(x)$ function,

$$\mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_1}\big) \oplus \mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_2}\big) \oplus \ldots \oplus \mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_m}\big) = \mathsf{ANF}\big(\mathsf{LSB}(x)\big). \tag{4.3.4}$$

The previous setting follows from the consideration of input distribution over the entire set of inputs in $\mathbb{F}_2^n$. However, the input considered for the quantum solution includes only even strings; thus, the same statement must be considered for partial Boolean functions. We address the entire problem with this additional difficulty in a twofold manner. First, we demonstrate a foundational structure of the ANF for all the exponential number of Boolean functions that equal the partial evaluation of the LSB function. We do this by showing that all these functions require, in their ANF representation, a minimum number of terms of a certain degree, as outlined in Lemma 33. Secondly, we examine the capacity of decision trees, to which $\mathsf{bPTF}^0[k]$ circuits can be reduced with the random restriction, in generating terms of the previously identified degree, as detailed in our Lemma 34 [5]. Finally, these two components, in conjunction with our switching lemma Lemma 13 that reduces $\mathsf{bPTF}^0[k]$ circuits to decision trees, allow us to determine a minimal depth of the decision tree's that directly translate to the minimal size for this class of circuits,

Lemma 14 (Informal, Lemma 35)**.** *For sufficiently large n, any $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit depth $d \geq 4$ and $k \leq n^{1/(5d)}$ that solves the $\mathcal{R}_2^m$ has size no smaller then $s \geq \exp\left(\widetilde{\mathcal{O}}\left(\left(nk^{-d}m^{-1/2}\right)^{1/(d-1)}\right)\right)$.*

The combination of this bound with the specific parameters of quantum upper bounds Lemma 36 and Corollary 5 complete our Theorem 4.2.1. Subsequently, we will enlarge the set of average-case hardness separation between QNC^0 and $\mathsf{bPTF}^0[k]$ with an infinite set of concretely instantiated relation problems.

4.3.2 Proof of Theorem 4.2.2: Separations for higher dimensions (qupits)

For our second result, we consider an infinite class of instances of the ISMR problems as candidates to separate qudit QNC^0 from $\mathsf{bPTF}^0[k]$, beyond the $\mathcal{R}_2^m$ problem and related problems that have been explored in previous works. These problems, however, possess some special properties, as the outcomes depend uniquely on the modular residue of their Hamming weight in $\mathbb{F}_p$. Therefore, we employ a correlation measure based on the nature of this problem as follows,

$$\mathsf{Corr}_{\mathcal{D}}(f,g) = \mathop{\mathbb{E}}_{x\sim\mathcal{D}}\left[\mathsf{Re}\left(e^{i\frac{2\pi|f(x)|-|g(x)|}{p}}\right)\right]. \tag{4.3.5}$$

[5] An illustrative example demonstrating the tree depth dependence on the ANF of Boolean functions is the parity function, which necessitates any decision tree to have a depth at least equal to the input size. This requirement ensures the tree can produce all the degree 1 terms of the parity function's ANF.

Intuitively, these correlation functions assess the distance between two relations using the natural measure of the inner product for domains that form an Abelian group [265].
Now, the proof will be divided into two parts. In the first part, we will demonstrate that polynomial-size $\mathsf{bPTF}^0[k]$ circuits exhibit a correlation that asymptotically approaches zero as the input size increases. In the second part, we demonstrate that quantum circuits over qupits solve all these instances of ISMR problems with a constant positive and far from zero correlation for inputs of arbitrary size.
For the first part, we will employ proof techniques used for the average-case hardness results of Theorem 4.2.1. However, in this higher-dimensional setting, many essential technical tools have not yet been developed or defined. We begin by noting that we are defining a family of XOR non-local games (Definition 39) related to the respective ISMR problems (Definition 38) for the first time, to the best of the authors' knowledge, and we need to determine the upper bounds of the efficiencies for any classical strategy. This contrasts with previous works, which relied on known non-local games and their predetermined optimal winning strategies.

Lemma 15 (Informal, see Lemma 38)**.** *Any local probabilistic classical strategy $w_{\mathcal{G}_p}$ that solves the Modular* XOR *non-local games, denoted as G_p, with n parties and messages over $\mathbb{F}_p^n$, taken from $\mathcal{D}_p$ a uniform distribution over $\mathbb{F}_2^{n(p-1)}$ with Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)$ Mod $p = 0$ and a bit to base-p encoding, the maximal correlation with r restricted bits is bounded by* $\mathsf{Corr}_{\mathcal{D}_p}\left(w_{\mathcal{G}_p}, \mathcal{G}_p\right) \leq (c_p)^{\frac{n-r}{p-1}}$*, with $c_p \in (0,1)$.*

To utilize the bounds determined for the non-local games, which are naturally defined as mappings of the type $\mathbb{F}_p^n \mapsto \mathbb{F}_p^m$, we need to consider encodings and decodings between binary strings in $\mathbb{F}_2^n$ and $\mathbb{F}_p^n$. Furthermore, we addressed an issue that arises in many of these encodings: they translate distributions that are uniform over one field to non-uniform distributions over another. This is particularly crucial as we aim to use uniform random restrictions in the proof process; hence, we require non-local games that correspond to uniform distributions of strings in $\mathbb{F}_2^n$. We resolved this problem by considering biased inputs for the non-local games that correspond to uniform distributions of strings in $\mathbb{F}_2^n$ when mapped back using our encoding. This involved choosing a specific encoding that allows for a linear bias in the non-local games, which we could accommodate in the particular proof technique for the winning strategies of non-local games while still obtaining sufficiently tight upper bounds.
From this upper bound on the optimal classical strategies and also the particular case for the probability of the qutrit case Lemma 40, we then have that we can now extend the lightcone arguments for NC^0 circuits such that we take advantage of the non-local for which we were able to bound its efficiency. This follows by considering blocks of bits representing the dits of the non-local game and exploring that these dits do not have any particular structure as in previous works [37] to apply lightcone arguments as in Lemma 39. This combination of results already determines a new NC^0 separation with qupits as described in Theorem 4.6.5, which we will further explore to obtain the separations against $\mathsf{bPTF}^0[k]$.
For that last step, we resort to our multi-switching lemmas Lemma 30, to reduce an $\mathsf{bPTF}^0[k]$ circuit to a decision tree that can be analyzed as of bounded causal effect, reducing it therefore to NC^0 type circuit. In particular, we obtain that through our capability of using uniform distributions of inputs we do have random restriction that are taking without any bias and obtain therefore average-case hardness results. From the previous considerations and building blocks, we obtain exactly that.

Lemma 16 (Informal, see Lemma 41)**.** *For sufficiently large n and $q \in \mathbb{N}_{>0}$, any* $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ *circuit C of depth $d \geq 4$, size $s \leq \exp\left(n^{1/(2d-2)}\right)$ and parameter $k \leq n^{1/(5d)}$ solves the $\mathcal{R}_p^m$ with correlation bounded by,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(C, \mathcal{R}_p^m\right) = \exp\left(-\Omega\left(\frac{n^{2-o(1)}}{m^{1+2/q}\log(s)^{2d-2}k^{2d}}\right)\right), \tag{4.3.6}$$

with $\mathcal{D}_p$ being the uniform distribution over string in $\mathbb{F}_2^n$ that satisfy $|x|$ Mod $p = 0$.

For the quantum lower bound, the initial step involves the creation of a generalization of the "poor man's cat state" introduced in [27], as detailed in our Definition 66. These states serve as the resource states and we demonstrate that they can be generated using QNC^0 circuits, as shown in Lemma 59. Subsequently, employing these resource states, we illustrate that another QNC^0 circuit can produce a classical outcome string related to the ISMRP. In particular, this string can classically be reduced with a NC^0 circuit to a solution to each one of the $\mathcal{R}_p^m$ problems sufficiently small outcome string m and constant and non-zero correlation.

Lemma 17 (Informal, see Lemma 44)**.** *There exists a QNC^0 circuit, denoted as C_Q, that given a generalized poor-man's cat state $\left|\mathsf{GPM}_p^n\right\rangle$, solves the* ISMRP*, represented as $\mathcal{R}_p^m$, across a uniform distribution $\mathcal{D}_p$ in $\mathbb{F}_2^n$ that satisfy the condition $\sum_{i=1}^n x_i$ Mod $p = 0$, with a correlation,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(C_Q, \mathcal{R}_p^m\right) = \frac{p-1}{p^2}. \tag{4.3.7}$$

Moreover, we obtain a separation for the qutrit case with respect to the probability of success in solving the respective ISMR problem. This is possible because the correlation function still has a direct relation with the success probability and thus compares to binary case Lemma 43.
We note that the previous lemma presented the most significant technical challenge for establishing the lower bounds for quantum correlation. This arises because poor-man's cat states in higher dimensions produce phase components that result in complex supports for standard measurement outcomes. Initially, we find a concise and clear representation for the projections of the outcome strings based on the input and the random strings z defining the specific qupit poor man's cat state at hand (as detailed in Equation (4.6.77) based on Table 5). The second major challenge is that the resulting support depends on a larger number of inner products between strings characterising the state and the input. Moreover, we require an NC^0 reduction based on these values that enhances the efficiency of solving the initial problems. Lastly, we demonstrate that at least one inner product can be evaluated within NC^0 and attached to the outcome strings, such that the entire string exhibits a larger correlation. Consequently, the quantum circuit solves the problem with a constant and input-independent correlation value, which together with the preceding steps completes the proof of Theorem 4.2.2.
Next, we extend the results of this section to the noisy setting i.e. we obtain separations between noisy quantum circuits and *noiseless* bPTF^0, taking inspiration from [35].

4.3.3 Proof of Theorem 4.2.3: Noise-resilient separations

For our computational separation between qupit QNC^0 circuits afflicted by noise and noiseless $\mathsf{bPTF}^0[k]$ circuits, we consider the qupit generalized local stochastic noise model (see Section 4.4.4), which is standard in quantum error correction. This model is favored due to its ability to account for gate-level noise, noisy input state preparation, and noisy measurement, while allowing for (weakly) non-local errors by requiring the probability of an error to decay exponentially with the number of qupits it affects non-trivially.
At a high level, we define a family of relation problems related to the noise-tolerant realization of the QNC^0 circuits solving the ISMR problems (see Definition 42). These problems are tailored to proving the noise-robustness of our QNC^0 Vs. $\mathsf{bPTF}^0[k]$ separations. We show that the ISMR problem is AC^0-reducible to a corresponding problem in the new family of problems exactly when an AC^0 circuit can successfully decode the outcome of these noise-tolerant QNC^0 circuits. This reduction allows us to lift the hardness results shown for the noiseless case to the case of noisy QNC^0 circuits, establishing the desired noise-robust separations with correlation bounds asymptotically equal to those in Theorem 4.6.4.

We address two main technical challenges that go beyond the previous noise-robust separations explored in [35, 36]. The first is that with qupits of dimension $p \geq 3$ we need to consider *non-Clifford* circuits. The need for these more complex quantum circuits arises from the inability to violate Bell inequalities with stabilizer states for any qupit dimension beyond qubits, as shown in [121, 122, 123], which also suggests that the same limitation extends to quantum-classical separations in circuit complexity. To overcome this, we must show that such circuits can also be implemented fault-tolerantly, which represents a novel departure from prior work on unconditional separations between quantum and classical circuit classes. The second challenge is the need to generalize the techniques used in [35] for quantum-error correction—such as single-shot state preparation, decoders, and transversal constant-depth gate execution—for arbitrary qupit dimensions, in addition to techniques related to fault-tolerantly implementing non-Clifford circuits.
We overcome the first challenge by introducing logical advice states over qupits, which enable the fault-tolerant realization of non-Clifford operations through a qupit magic state injection protocol. In particular, we prove in Section 4.7.2 that for a specific CSS-type error correction code that fulfills a set of conditions listed in Definition 43, we obtain fault-tolerant implementations of non-adaptive qupit Clifford circuits with advice states, for arbitrary prime dimensions.

Lemma 18 (Informal, see Lemma 45)**.** *Let $\mathcal{EC}_C$ be the error-corrected version of C, a constant-depth Clifford circuit of depth d with advice state $|A\rangle$, using a code that meets the conditions in Definition 43. The code has distance $l = \mathcal{O}(\text{poly}\log n)$ and the circuit is affected by local stochastic noise $\mathcal{E} \sim \mathcal{N}(\varrho)$ and the advice by $\mathcal{E}_A \sim \mathcal{N}(\rho)$. Then, whenever* $\max\{\varrho, \rho\} < p_{th} = 2^{-2^{\mathcal{O}(d)}}$ *for any input $x \in \mathbb{F}^n$ we can show that*

$$\Pr[\text{dec}^*(\mathcal{EC}_C(x)) = C(x)] > 0.99, \tag{4.3.8}$$

with dec* *being the combined correction and decoding functions needed to retrieve the logical outcome from the encoded state.*

We divide the proof into two parts. The first part demonstrates that the qupit surface code meets all the necessary conditions, generally implying that it supports single-shot state preparation, transversal gate implementation in constant depth, and single-shot information retrieval using the selected decoder. Making these steps precise in higher-dimensional QNC^0 circuits requires new insights that extend beyond prior work, which we outline below. In parallel with verifying that the qupit surface code ensures the fault-tolerant implementation of the circuit architecture as shown in Lemma 45 and Figure 24, we must prove a second element. This refers to the existence of QNC^0 circuits that are non-adaptive Clifford circuits with input-independent advice states capable of solving the ISMR problems; this requires redesigning new quantum circuits and forms the second piece of the proof.

Noise-resilient qupit Clifford circuits with quantum advice

The first condition, concerning the transversal implementation of Clifford operations, follows from the work of [280]. However, achieving single-shot state preparation and information retrieval in the qupit surface code is more complex than implementing transversal gates. In higher dimensions, errors do not simply manifest as defects at the endpoints of the qupit lattice, as they do in the qubit case. Instead, for every error, defects are likely to be scattered throughout the lattice. This distribution of defects motivated the use of the HDRG decoder, as it has been shown to have good error-correction properties beyond the qubit case, overcoming issues with the previously considered minimum weight perfect matching decoder [281, 282]. The information retrieval property necessitates that the code and decoder accurately perform logical $\overline{Z}$ measurements even under noisy conditions. In Lemma 49, we extend a result from [283] to the definition of local stochastic noise discussed in this text. We demonstrate that the HDRG decoder's probability of failure decreases exponentially as the qupit lattice size increases. Consequently, the HDRG decoder reliably yields the correct outcome for logical measurements, provided the noise rate does not exceed a

certain threshold. Specifically, we establish that $\Pr[\text{Success}] \geq 1 - \exp(-\Omega(m^{\eta}))$, where m is the surface code distance and η is a constant.
We also demonstrate that single-shot state preparation can be achieved using the qudit surface code and the HDRG decoder for qupits of dimension $p \geq 2$. More precisely, we show that the 3D block construction from [124] preparing logical Bell pairs can be adapted for single-shot state preparation of GHZ_2 qupit states, a capability not previously demonstrated. This adaptation involves defining functions necessary for the single-shot state preparation process, typically categorized as recovery and repair. The recovery function entails applying operations to retrieve the correct state from the randomness inherent in the noise-free process, while the repair function corrects the states based on the effects of noise that may occur during the described noise-free state preparation circuit and the recovery function.
The recovery function is derived directly from our generalization of the state preparation process. Regarding the repair function, we establish its feasibility up to the single-shot properties, using the alternative lifting properties proposed in [35] and considering a repair function based on the HDRG decoder (Lemma 50). This shows that these logical states can be prepared for errors below a certain threshold with exponentially high confidence for increasing the lattice size.

Non-adaptive Clifford circuits with magic state injection

To redesign the quantum circuits proposed in Lemma 44 for the required noise-resilient architecture from Lemma 45, we retain all Clifford operations as they present no difficulties. However, handling the non-Clifford rotations requires a new approach. Initially, we could consider using a gate teleportation gadget that integrates the non-Clifford gate into the advice state, as illustrated in Equation (4.7.65). Unfortunately, while these gadgets are adaptive, we only have access to non-adaptive gadgets for a noise-resilient execution. Consequently, we implement the non-adaptive gadgets as depicted in Section 4.7.2, and in Lemma 53, we discuss the additional phases introduced by the lack of adaptive corrections. Due to their diagonal nature, we demonstrate that it is still possible to address the ISMR problems using a more complex reduction from the outcome of these qupit QNC^0 circuits, involving a considerably more intricate NC^0 reduction.

Lemma 19 (Informal, see Lemma 54)**.** *There exist constant-depth non-adaptive Clifford circuits with the advice state $\left|T^{1/p}\right\rangle^{\otimes n}$, denoted as Cliff_{+T}, that solve the ISMR problems $\mathcal{R}_p^m$ and $m = \mathcal{O}(n \cdot g^{p^3})$, for a uniform distribution $\mathcal{D}_p$ of input strings that satisfy the condition $\sum_{i=1}^{n} x_i \bmod p = 0$ within $\mathbb{F}_2^n$ with a correlation, $\mathsf{Corr}_{\mathcal{D}_p}\left(|\mathsf{Cliff}_{+T}|, |\mathcal{R}_p^m|\right) = \frac{p-1}{p^2}$.*

In our final step, we show that there exists a noisy quantum circuit that solves the proposed noise-resilient relation problems with large constant correlation. We consider the encoded version of the circuits proposed in Lemma 54 using the qupit surface code and the HDRG decoder. This does provides us with the quantum lower bound on the correlation described in Lemma 55. Additionally, we obtain the classical lower bound through the AC^0-reduction from the noise relation problems to the ISMR problems.
Furthermore, we demonstrate that a minimal universal gate set, which requires only the set of Clifford gates and a single magic gate, suffices. We prove that this magic gate is related to T-type qubit gates, as shown in Lemma 52. Thus, we recover the standard noise-resilient circuit architecture, incorporating magic state injection gadgets for all qupit dimensions, thereby extending the scope of these results.
As a corollary, we also obtain the same separation for qubits. However, since these are based on Clifford gates, we obtain a direct separation between noisy QNC^0 circuits and $\mathsf{bPTF}^0[k]$ circuits without any advice states.

4.3.4 New switching lemma for biased PTF circuits

As mentioned in Section 4.3.1, we prove a new multi-switching lemma for $\mathsf{bPTF}^0[k]$ circuits, which provides tighter concentration bounds compared to previous work. We then use it to establish depth

reduction for $\mathsf{bPTF}^0[k]$ circuits with multiple output bits. Our approach parallels the strategies employed for multi-output AC^0 circuits [284, 27], tailored specifically to $\mathsf{bPTF}^0[k]$ circuits. Using this tool, we analyze how $\mathsf{bPTF}^0[k]$ circuits reduce under random restrictions, showing that we obtain *decision forests* of low complexity with high probability. These more tractable decision forests can then be compared with quantum circuits via lightcone arguments.

Tighter multi-switching lemma for $\mathsf{bPTF}^0[k]$

On the technical side, our main contribution is a multi-switching lemma, which shows that a finite set of depth-2 $\mathsf{bPTF}^0[k] \circ \mathsf{AND}_w$ circuits reduces to a decision forest of the type $\mathsf{DT}(w)^m \Rightarrow \mathsf{DT}(t)$ with high probability converging to unity as $1 - O(2^{-t})$. We provide the formal statement in Lemma 12 and the proof in Lemma 25. Note that while Theorem 4.4.2, as proven in [115], also functions as a multi-switching lemma, a tighter bound on the probability is necessary to reduce the initial circuit to a tree-like structure. In particular, in the success probability of the switching lemma, $4(64(2^k m)^{1/q} pw)^t$, the dependence on the size t of the global and q of local decision trees in the decision forest must be removed from the term m, that accounts for the number of $\mathsf{bPTF}^0[k]$ circuits simultaneously reduced by the lemma. Expanding on this technical limitation, the primary bottleneck arises when employing the lemma for the depth reduction step. This dependency necessitates the utilization of large fixing probabilities for the random restrictions, ultimately leading to the fixation of nearly all the free variables. Consequently, this is insufficient to prove a separation from QNC^0. In contrast, our tighter bound for the various lemmas allows for subsequent separation from QNC^0 circuits via locality arguments.
Our proof of Lemma 25 for $\mathsf{bPTF}^0[k]$ employs an inductive approach similar to the one used by [253] for AC^0 circuits. We conduct induction over the number of variables fixed by random restrictions and the circuits f_i corresponding to each output bit, aiming to lower bound the probability with which the latter reduce to decision trees of the type $\mathsf{DT}(l)$. Simultaneously, when we encounter a circuit that does *not* reduce to a decision tree of the type $\mathsf{DT}(l)$ with a set of variables fixed by random restrictions, we query the variables that remain "alive" to forcefully simplify this circuit. These variables then become part of the global decision tree $\mathsf{DT}(t)$, sequentially contributing to the construction of a decision forest. To make this approach effective, we address the issue of fixing variables and clauses that describe these circuits by combining our induction with canonical decision trees from the witness method technique (also used by [115] to prove Theorem 4.4.2).
The induction technique described above necessitates the use of *downward-closed* random restrictions. This property ensures that the size of our canonical decision tree does not increase when additional variables, left alive by the initial random restriction, are further restricted. In other words, it guarantees the monotonicity of decision tree size under random restrictions, which is crucial for the inductive reasoning process. This is accomplished through Algorithm 1, which constructs our canonical decision trees $\mathsf{DT}(l)$ corresponding to the leaves of the decision forest. We prove two ingredients that ensure downward closedness. First, Lemma 27 shows that all random restrictions reducing the circuit to a fixed decision tree $\mathsf{DT}(l)$ overlap in the variables that they fix. Then, Lemma 28 shows that for any arbitrary random restriction, fixing more variables does not increase the depth of the decision tree to which the initial circuit reduces.

Depth reduction for $\mathsf{bPTF}^0[k]$

Adopting the proof technique of [284] for AC^0 circuits, we use Lemma 25 to prove more efficient depth reduction for $\mathsf{bPTF}^0[k]$ circuits.
The multi-switching lemma first reduces the depth by 1, from d to $d-1$, for a computational object of the type $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)^{s_1} \Rightarrow \mathsf{DT}(t-1)$. This describes a decision forest that provides the input for a $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d]$ circuit of depth d, with layers having $s_1, \ldots, s_d$ gates

each, which will subsequently have its depth reduced, as demonstrated in the following lemma.

Lemma 20 (Informal, see Lemma 29)**.** *For $f \in \mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)^{s_1} \not\Rightarrow \mathsf{DT}(t-1)$ parameterized by circuit depth d, decision tree depth t, and restriction probability p, we have that*

$$\Pr_{\rho \sim R_p}\left[f\restriction_\rho \notin \mathsf{bPTF}[k; d; s_2, \ldots, s_d] \circ \mathsf{DT}(l)^{s_2} \not\Rightarrow \mathsf{DT}(t-1)\right] \leq s_1 \cdot 2^k (400wp)^{t/2}, \quad (4.3.9)$$

where $l \geq \log s_1 + k + 2$.

Using this result, we can reduce the depth of a $\mathsf{bPTF}^0[k]$ circuit while simultaneously constructing the decision forest to which it will be reduced. More precisely, iterative application of the depth reduction lemma to the more elaborate computational object considered above leads to the following lemma.

Lemma 21 (Informal, see Lemma 30)**.** *For parameters $d, t, k, q, m, s_1, l_1, \ldots, s_{d-1}, l_{d-1} \in \mathbb{N}$ and $p_1, \ldots, p_d \in (0,1)$, we have for $f \in \mathsf{bPTF}^0[k; d; s_1, \ldots, s_{d-1}, m](k)$ with n inputs and m outputs that*

$$\Pr_{\rho \sim R_p}\left[f\restriction_\rho \notin \mathsf{DT}(q-1)^m \not\Rightarrow \mathsf{DT}(2t-2)\right] \leq \sum_{i=2}^{d-1} s_i \cdot 2^k \mathcal{O}(p_i l_i)^{t/2} + 2^k m^{1/q} \mathcal{O}(p_d \cdot l_d)^t, \quad (4.3.10)$$

where $p = p_1 \cdot p_2 \cdot \ldots \cdot p_d$.

Finally, we assign values to the probabilities $p_1, \ldots, p_d$ of different random restrictions and the depths $l_1, \ldots, l_d$ of the local decision trees such that for global decision tree depth t, initial circuit size s, and circuit type parameterized by k, we determine the probability of successfully reducing the circuit to decision forests. We define these decision forests informally in Lemma 13 and formally in Lemma 31.

4.4 Preliminaries and Notation

In this work, we expand on the set of classical circuit classes against which we can establish a separation of computational power of QNC^0, for a relational problem. Towards this goal, in Section 4.4.1 we first set up the notation and definitions for the constant-depth classical circuit classes we study, and then in Section 4.4.2 introduce the most important techniques we use: random restrictions and switching lemmas. We then present some background on the quantum information theory of non-local games in Section 4.4.3, and noisy quantum circuits in Section 4.4.4.
Throughout the text we use the notation $[n] = \{1, 2, \ldots, n\}$, log denotes the logarithm to base two, and exp denotes the exponential function, with the base e or 2 that will be clear from context.

4.4.1 Low-depth complexity classes

We now explicitly introduce the circuit classes and computational models that we analyze in this work. For a detailed introduction to these topics, we refer to standard textbooks such as [43]. The *depth* of a circuit is the maximum number of gates that are composed sequentially. The *width* of a circuit refers to the largest number of gates applied simultaneously in any of the its layers. In terms of standard computational resources, depth is related to time, and width is related to the space of a computation. We say that a circuit is composed of bounded fan-in gates when the number of wires feeding logical values into each gate for processing is limited by a fixed constant. Conversely, unbounded fan-in refers to gates that can accommodate a polynomial number of input wires, relative to the input size, feeding into the logical gate for processing.

We assume the standard notation and definitions for classical notions such as NC^0 and AC^0. NC^0 is the only circuit class that uses only bounded fan-in gates, but unbounded fan-out and arbitrary wiring. A family of circuits $\{\mathcal{C}_n\}_{n\geq 1}$ is said to be X-uniform for a complexity class X if there is a Turing machine T in X which on input 1^n outputs a description of $\mathcal{C}_n$ for each n. All of our classical circuits will be non-uniform, and admit randomised advice strings of polynomial length in the input size (denoted by $/\mathsf{rpoly}$), while our quantum circuits will always be uniform.

Definition 21 (NC^0 circuits)**.** *The class* NC^0 *consists of classical constant-depth circuits composed of bounded fan-in* $\mathsf{AND}, \mathsf{OR}$*, and* NOT *gates, with a polynomial number of gates in the input size* n*.*

The subsequent two circuit classes permit unbounded fan-in, which enhances their computational capabilities. They are distinguished by their computational power and the selection of gate sets.

Definition 22 (AC^0 circuits)**.** *The class* AC^0 *consists of classical constant-depth circuits composed of unbounded fan-in* $\mathsf{AND}, \mathsf{OR}$*, and* NOT *gates, with a polynomial number of gates in the input size* n*.*

Note that in fact circuits with sub-logarithmic depth, $o(\log n)$, are considered to be of constant depth. Our primary focus will be on QNC^0, which as defined in Section 4.2.1 is the class of constant-depth quantum circuits constructed from a universal and finite set of gates.

Definition 23. $(\mathsf{QNC}^i$ *class*$)$*. The* QNC^i *class consists of quantum circuits composed of bounded fan-in quantum gates from a universal finite set, with depth* $\mathcal{O}(\log^i(n))$*, and a polynomial number of gates concerning the size of the input string* n*.*

In addition to the standard classical circuit classes, our analysis will focus on a novel class featuring parameterized gates, first defined by [115]. However, we will present these using a different notation, in the form commonly used in relation to polynomial threshold functions [285, 286, 287]. We first introduce the class of polynomial threshold gates restricted by a bias parameter.

Definition 24 (Biased polynomial threshold gates)**.** *We denote by* $\mathsf{bPTF}[k]$ *the set of unbounded fan-in gates that implement two types of Polynomial Threshold Functions with bias* k*, defined as follows. The first type of gates implement* OR*-type PTFs, of the form*

$$f_{\mathsf{OR}}(x) = \begin{cases} P(x), & \sum_{i=0}^{n} x_i \leq k \\ 1, & \sum_{i=0}^{n} x_i > k \end{cases}; \quad \text{with } P : \mathbb{F}_2^n \to \mathbb{F}_2 \text{ a polynomial over } \mathbb{F}_2 = \{0,1\}. \tag{4.4.1}$$

These gates permit arbitrary mappings of strings with a Hamming weight bounded by $\leq k$ *and identically equal to* 1 *for all inputs with a Hamming weight* $> k$*. The second type of gates includes all unbounded fan-in gates that implement the analogous* AND*-type PTFs:*

$$f_{\mathsf{AND}}(x) = \begin{cases} P(x), & \sum_{i=0}^{n} x_i \geq n-k \\ 0, & \sum_{i=0}^{n} x_i < n-k \end{cases}; \quad \text{with } P : \mathbb{F}_2^n \to \mathbb{F}_2 \text{ a polynomial over } \mathbb{F}_2 = \{0,1\}. \tag{4.4.2}$$

AND*-type PTFs allow for arbitrary mappings of strings with Hamming weight* $\geq n-k$ *and are identically equal to* 0 *for all inputs with a Hamming weight* $< n-k$*.*

Note that when $k = 0$, we recover the usual unbounded fan-in AND and OR gates. Given this definition of biased threshold gates, we now introduce circuit classes composed of these gates, which will be the main objects of study in this work.

Definition 25 (Biased polynomial threshold circuits)**.** *The class* $\mathsf{bPTF}^0[k]$ *consists of classical circuits composed of unbounded fan-in* $\mathsf{bPTF}[k]$ *gates with constant depth, and a polynomial number of gates in the input size* n*.*

Note that for $k = \mathcal{O}(1)$, $\mathsf{bPTF}^0[k]$ is equal to AC^0. In addition to these circuit classes, we also work with decision trees, which are fundamental in the study of these classes due to their role in random restriction techniques. We denote by $\mathsf{DT}(t)$ the class of Boolean functions computed by depth-t decision trees with a single output bit.
Furthermore, we will examine more complex decision trees, whose leaves do not just bear single binary values, but additional sets of decision trees. This structure allows for a hierarchical querying process: a global decision tree first addresses a set of variables, which in turn delineates a collection of local decision trees responsible for computing the bits of the final output string. To denote non-binary values at the leaves of a decision tree, we introduce the symbol $\Rightarrow\!\!|$ to signify that the entity to the left represents its leaves.

Definition 26 (Decision forests)**.** *We use* $\mathsf{DT}(w)^m \Rightarrow\!\!| \mathsf{DT}(t)$ *to denote the class of* (t, m, w)*-decision forests, defined as depth-*t *decision trees whose leaves are labeled by* m*-tuples of depth-*w *decision trees.*

Building upon these objects, we introduce a new type of decision trees whose outputs become the inputs to a $\mathsf{bPTF}^0[k]$ circuit. In the following, we write $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d]$ to denote a $\mathsf{bPTF}^0[k]$ circuit with depth d, and s_i gates for each layer i for $i \in \{1, \ldots, d\}$[6].

Definition 27 (Biased threshold circuits with decision tree inputs)**.** *The object* $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)$ *defines the class comprised of* $\mathsf{bPTF}^0[k]$ *circuits with depth* d*, and* s_i *gates per layer whose inputs are labeled by decision trees in* $\mathsf{DT}(w)$*.*

We also introduce a second object of this type, comprising a global decision tree whose leaves specify a collection of local decision trees; these, in turn, determine the input to a $\mathsf{bPTF}^0[k]$ circuit.

Definition 28. $(\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w) \Rightarrow\!\!| \mathsf{DT}(t))$*. The object* $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w) \Rightarrow\!\!| \mathsf{DT}(t)$ *defines the class comprised of depth-*t *decision trees, whose leaves are objects in* $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)$*.*

Finally, we also introduce the Algebraic Normal Form (ANF) which refers to polynomial representation of Boolean functions over $\mathbb{F}_2$.

Definition 29 (Algebraic Normal Form.)**.** *Every Boolean function* $f : \{0,1\}^n \mapsto \{0,1\}$ *can be represented uniquely as,*

$$f(x) = \bigoplus_{S \subseteq [n]} c_S \prod_{i \in S} x_i. \tag{4.4.3}$$

with $c_S \in \{0, 1\}$*.*

4.4.2 Random restrictions and Switching lemmas

Our analysis relies heavily on random restrictions, a technique that has been widely utilized in computational circuit complexity.

Definition 30 (p-random restriction)**.** *A* p*-random restriction* $\rho \in R_p$ *is a function mapping from a set of literals* $\{x_i\}_{i \in I}$ *to values* $\{0, 1, *\}$*. It is parameterized by the probability* $p \in (0, 1)$*: each variable is independently kept 'alive' using the symbol* $*$ *with probability* p*, or assigned the value* 0 *or* 1 *with equal probability* $\frac{1-p}{2}$*.*

[6] These computational objects will be our $\mathsf{bPTF}^0[k]$ equivalent of those introduced in [284] for AC^0 and used to establish the respective depth reduction techniques.

Incorporating the concept of random restriction, the application of such constraints to a function f, denoted by $f\lceil\rho$, results in the definition of a new function evaluated according to $f(x \circ \rho)$. Put simply, the restrictions imposed by ρ are applied to the inputs of f, where the surviving variables adopt the values from x that are allowed to fluctuate. Thus, a restricted function can be viewed as a new function characterized by a reduced domain of input.
A random restriction can also be described by first choosing a ground truth $z \in \{0,1\}^n$ and a set of indices $\Lambda \subseteq [n]$. The random restriction, denoted as $\rho(z, \Lambda)$, is then defined such that Λ determines the indices x_i for which the corresponding bits are set to the symbol $*$. The remaining bits adopt the corresponding values from z. For instance, a p-random restriction in this context corresponds to a stochastic process where each index in Λ is selected independently with probability p, and z is chosen uniformly at random from the set $\{0,1\}^n$.
With the use of random restriction, it is possible to reduce the circuits from the classes defined previously to decision trees based on switching lemmas. In particular, we will discuss two important switching lemmas derived by [115] for $\mathsf{bPTF}^0[k]$ circuits, along with some additional useful lemmas from the same text.

Theorem 4.4.1. *[115] Let f be computable by a depth-2 circuit[7] F comprised of a* $\mathsf{bPTF}[k]$ *gate which has as inputs* AND *or* OR *gates with maximal fan-in w, designated as* $\mathsf{bPTF}[k] \circ \{\mathsf{AND}, \mathsf{OR}\}_w$*. Then,*

$$\Pr_{\rho \in R_p}[F\lceil_\rho \notin \mathsf{DT}(t)] \leq (20pw)^t 2^k. \tag{4.4.4}$$

The previous theorem indeed provides a switching lemma for $\mathsf{bPTF}^0[k]$, exactly with the same structure as Hastad's switching lemma for AC^0. Now, we will also consider a multi-switching lemma developed in the same work.

Theorem 4.4.2. *[115] Let $\mathcal{F} = \{F_1, \ldots, F_m\}$ be a list of* $\mathsf{bPTF}[k] \circ \mathsf{AND}_w$ *circuits on $\{0,1\}^n$. Then,*

$$\Pr_{\rho \in R_p}[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(q)^m \twoheadrightarrow \mathsf{DT}(t)] \leq 4(64(2^k m)^{1/q} pw)^t. \tag{4.4.5}$$

Finally, we can leverage a useful lemma that facilitates the mapping between AND_w and OR_w clauses and decision trees.

Lemma 22. *[115] Any depth-2 circuit of the form* $\mathsf{bPTF}[k] \circ \mathsf{DT}(w)^m$ *comprised of a top (i.e., last) gate, which is a* $\mathsf{bPTF}[k]$ *gate with fan-in m, with* $\mathsf{DT}(w)$ *as inputs can be expressed as a circuit in* $\mathsf{bPTF}[k] \circ \mathsf{AND}, \mathsf{OR}_w$ *with a size of $m2^w$.*

4.4.3 Non-local games and Correlation

Non-local games are a fundamental element of study for quantum-classical separation in foundational terms. These games involve specific rules played by participants. It is possible to demonstrate that, for certain games where entanglement is a key resource, classical players cannot achieve the same winning probabilities as quantum players. Particularly of interest are XOR and generalized XOR games in the multi-party setting, as we relate these to computational relation problems. They are defined as follows.

Definition 31 (XOR and generalized XOR games)**.** *These classes of multiparty non-local games involve n parties, denoted $p_1, \ldots, p_n$. In these games, each party p_i receives information x_i, where $(x_1, \ldots, x_n)$ is drawn from a distribution $\mathcal{D}$. The parties, unable to communicate post-distribution, must each provide an individual response y_i, contributing to a collective output $y_1 \otimes y_2 \otimes \ldots \otimes y_n$. The probability of winning*

[7] We will refer to a circuit that computes a certain function using upper case letters, and reserve lower case letters for the corresponding abstract function.

the game is determined by the expression,

$$\Pr_{(x_1,\ldots,x_n)\sim\mathcal{D}}\left[\sum_{i=1}^{n} y_i \text{ Mod } p = f(x_1,\ldots,x_n)\right], \tag{4.4.6}$$

where $f : \mathbb{F}_p^n \mapsto \mathbb{F}_p$ is a predetermined function. This function maps the input tuple from the finite field $\mathbb{F}_p^n$ to a single element in $\mathbb{F}_p$, thereby dictating the winning condition based on the collective responses y_i.

These are typically associated with Bell inequalities [166, 44], which define the statistics that can distinguish between local hidden variable theories (classical models) and quantum observations with quantum states. Specifically, we will utilize established definitions to determine the winning probabilities for games in this context. It is also worth mentioning that our primary interest lies in achieving the maximal classical winning probability, while aiming for quantum winning probabilities that surpass the classical limit, achievable through constant-depth quantum circuits.
Additionally, we need to establish a measure of efficiency for higher-dimensional non-local games and computational problems. Specifically, we consider that the correctness of the outcome of ISMR problems and the generalized XOR non-local games depends solely on the modular residue of their Hamming weight. We can define proximity between relations based solely on that value. Furthermore, considering $\omega^{|C(x)|}$ again, with ω being the roots of unity, we encounter group arithmetic over the finite fields $\mathbb{F}_p$ with respect to the addition operation. Thus, we can assess the distance between two relations using the natural measure of the inner product for domains that form an Abelian group, defined as $\langle f, g\rangle = \sum_{|G|} f\overline{g}$ [265]. More precisely, we can use a general correlation measure for any circuit attempting to approximate an ISMR problem or a generalized XOR non-local game as follows,

Definition 32 ($\text{Corr}_{\mathcal{D}}(f,g)$)**.** *We define the correlation between functions $f, g : \mathbb{F}_p^n \mapsto \mathbb{F}_p$ on a distribution $\mathcal{D}$ over $\mathbb{F}p^n$ as,*

$$\text{Corr}_{\mathcal{D}}(f,g) = \mathop{\mathbb{E}}_{x\sim\mathcal{D}}\left[\text{Re}\left(e^{-i\frac{2\pi|f(x)|-|g(x)|}{p}}\right)\right]. \tag{4.4.7}$$

More precisely, it will discount for deviation will be defined by its size and the respective real part of the roots of unity involved in each dimension (see Figure 15).

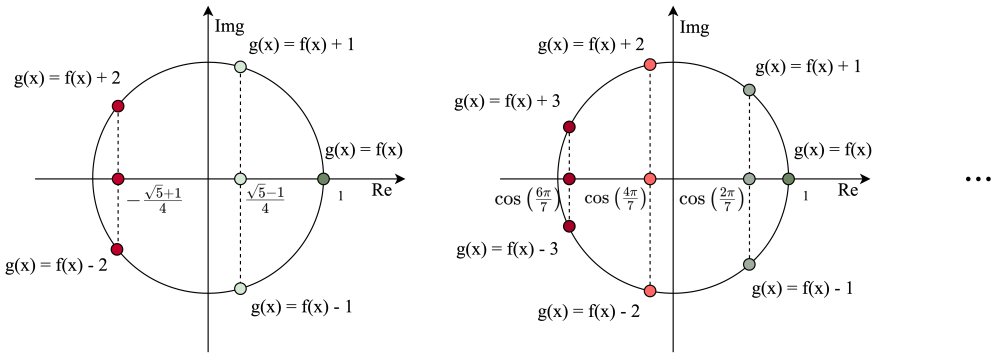


Figure 15: Illustration of the respective contributions to the correlations function based on the deviation between two arbitrary function f and g over $\mathbb{F}_3$ and $\mathbb{F}_5$.

4.4.4 Noise and quantum error correction in higher dimensions

We will analyze noisy quantum circuits over qubits and qudits in Section 4.7. To do this, we first introduce the noise model that we adopt and identify key properties for later use. We will then introduce the topological code that we use to attain noise-resilient quantum advantage.

Qupit operations and local stochastic noise

We first establish preliminaries on the standard operations and gates used to characterize the impact of noise. In particular, for qupits, we need to introduce the generalized Pauli operators [288], which are higher-dimensional versions of the single-qubit Pauli operators. Specifically, the Pauli X and Z operators generalize to the shift and clock operators, respectively

$$\mathsf{X} = \sum_{j=0}^{p-1} |j\rangle \langle j \oplus 1| , \;\; \mathsf{Z} = \sum_{j=0}^{p-1} \omega^j |j\rangle \langle j| . \tag{4.4.8}$$

Both are non-Hermitian unitary operators, where $j \in \mathbb{F}_p$, $\omega = e^{\frac{2\pi i}{p}}$ and we will use $\oplus$ to denote addition mod p respective to each qupit dimension under consideration. These operators have order p, and commutation relation

$$\mathsf{Z}^p = \mathsf{X}^p = \mathbb{1}, \;\; \mathsf{ZX} = \omega \mathsf{XZ}. \tag{4.4.9}$$

For a system of n-qupits, we use the notation $\omega(a), \mathsf{X}(a), \mathsf{Z}(a)$ with $\mathsf{Z}(a) = \mathsf{Z}^{a_1} \otimes \cdots \otimes \mathsf{Z}^{a_n}$ for $a \in \mathbb{F}_p^n$.

Definition 33 (Generalized Pauli group)**.** *The generalized Pauli group $\mathcal{P}_p$ is generated by the single qupit operators* X *and* Z*. The elements of $\mathcal{P}_p$ are given by $\omega^k \mathsf{Z}^a \mathsf{X}^b$ for $a, b, k \in \mathbb{F}_p$. The generalized Pauli group $\mathcal{P}_p^{\otimes n}$ over n qupits is the n-fold tensor product of the single-qupit Pauli groups, and each operator can be expressed as $\omega(k)\mathsf{Z}(a)\mathsf{X}(b)$ for $a, b \in \mathbb{F}_p^n$ and $k \in \mathbb{F}_p$.*

In the rest of the paper we drop the subscript p in $\mathcal{P}_p$ when it is clear from the context. Using the generalized Pauli operators, we can define the generalized version of the Clifford group.

Definition 34 (Qupit Clifford group)**.** *The Clifford group $\mathcal{C}^n$ over n qupits is the normalizer of the n-qupit Pauli group, defined as follows.*

$$\mathcal{C}^n = \{C \in U(p^n) \mid \forall P \in \mathcal{P}^n, CPC^\dagger \in \mathcal{P}^n\}. \tag{4.4.10}$$

Interesting elements of the generalized Clifford group include the qupit versions of the Hadamard and CNOT gates, known as the Fourier and Sum gates respectively, as well as the involution gate INV. These can be defined as follows.

$$\mathsf{F} = \frac{1}{\sqrt{p}} \sum_{x=0}^{p-1} \sum_{y=0}^{p-1} \omega^{xy} |x\rangle\langle y| , \; \mathsf{SUM} = \sum_{x=0}^{p-1} |x\rangle\langle x| \otimes |x \oplus y\rangle\langle y| , \text{ and } \mathsf{INV}\, |j\rangle = |-j\rangle . \tag{4.4.11}$$

We are now ready to introduce the error model we consider, along with some of its properties. In the following, for an n-qupit Pauli operator $\mathcal{E}$ we write $\mathsf{Supp}(\mathcal{E})$ to denote the subset $\mathcal{I} \subset [n]$ of indices of the qupits on which $\mathcal{E}$ acts non-trivially. That is, $\mathsf{Supp}(\mathcal{E})$ is the subset of qupits that are corrupted by $\mathcal{E}$. The local stochastic noise model is defined as follows [289, 290].

Definition 35 (Qupit local stochastic noise)**.** *Let $0 \leq \tau \leq 1$. A random n-qupit Pauli error $\mathcal{E}$ with support $\mathsf{Supp}(\mathcal{E}) \subseteq [n]$ is said to be τ-local stochastic if for all $F \subseteq [n]$, we have*

$$\Pr[F \subseteq \mathsf{Supp}(\mathcal{E})] \leq \tau^{|F|}. \tag{4.4.12}$$

We write $\mathcal{E} \sim \mathcal{N}(\tau)$ to denote random Pauli errors that follow a τ-local stochastic noise model.

This is a very versatile noise model, being able to account for independent, correlated, and even adversarial noise, given that the probability of an error occurring decays as its support size increases. This means that local errors are more likely to occur than non-local ones.

Lemma 23. *The qupit local stochastic noise defined in Definition 35, over qupit dimension p, satisfies the following properties.*

1. *Let $\mathcal{E} \sim \mathcal{N}(\tau)$. Then for any random error $\mathcal{E}'$ such that $\mathsf{Supp}(\mathcal{E}') \subseteq \mathsf{Supp}(\mathcal{E})$ with certainty, we have that $\mathcal{E}' \sim \mathcal{N}(\tau)$.*

2. *Let $\mathcal{E} \sim \mathcal{N}(\tau)$ and $\mathcal{E}' \sim \mathcal{N}(\varrho)$ be (possibly dependent) random Paulis. Then, $\mathcal{E} \circ \mathcal{E}' \sim \mathcal{N}(\tau')$ where $\tau' = p \cdot \max(\sqrt{\tau}, \sqrt{\varrho})$.*

These properties are up to some constants equal across all qudit dimensions, and the proof for the qubit case presented in [35] extends to the qupit case with only minor differences.
When a circuit is affected by the local stochastic noise model, each layer of the circuit, defining the support of the error, will be impacted by a random Pauli operator, adhering to the local stochastic property parameterized by the respective probability.
Importantly, this noise model allows for commuting the individual error operators through Clifford circuits. This enables the overall effect of noise, represented by random Pauli errors occurring after each layer of gates, to be described as amplified local stochastic noise occurring only at the end of the entire Clifford circuit.

Lemma 24. *For any n-qupit Clifford circuit U_d of depth d composed of one- and two-qupit gates, the noisy implementation of U_d described as $C_1\mathcal{E}_1C_2\mathcal{E}_2 \dots C_d\mathcal{E}_d$ with $\mathcal{E}_i \sim \mathcal{N}(\tau)$ when applied to $|0\rangle$ and measured produces an outcome $z \in \{0,1\}^n$ with the following conditional probability.*

$$\mathsf{Pr}(z|\mathcal{E}) = |\langle z| U_d\mathcal{E}_t |0\rangle|^2, \tag{4.4.13}$$

with $\mathcal{E}_t \sim \mathcal{N}(p^2\tau^{4-d-1})$.

The proof in [35] for qubit systems with local stochastic noise extends straightforwardly to the qupit case.

Qupit surface code

To address noise, we must select a quantum error correcting code with an appropriate decoder to correct the errors based on the chosen noise model. We identify the surface code over qupits as a candidate for our general purpose.
Analogously to the qubit case, we have a square lattice of qupits of size $L \times L$, with smooth (square tiles) and rough (tiles without all the edges) boundaries. In particular, this topological code is characterized by the stabilizer subgroup $\mathcal{S} = \langle A_v, B_p \rangle$ defined as follows,

$$A_v = X_e \otimes X_e^\dagger \otimes X_e \otimes X_e^\dagger \; \forall e \in V \tag{4.4.14}$$

$$B_p = Z_e \otimes Z_e \otimes Z_e^\dagger \otimes Z_e^\dagger \; \forall e \in P, \tag{4.4.15}$$

where V defines all the edges incident to a single vertex of the lattice v, and P defines the plaquette constituted by all the edges that close a square/tile of the lattice. The entire $L \times L$ lattice, containing $2L^2 - 2L - 1$ qudits, defines a single logical qupit. Given this logical qupit, the logical operators $\overline{\mathsf{X}}$ and $\overline{\mathsf{Z}}$ are defined as follows

$$\prod_{e \in \mathsf{Column}} X_e, \quad \prod_{e \in \mathsf{Line}} Z_e, \tag{4.4.16}$$

which also fixes the code distance to be L. Note that throughout the entire text, we will use the notation $|\overline{a}\rangle$ for the logical version of a state $|a\rangle$ and $\overline{A}$ for an operator A.

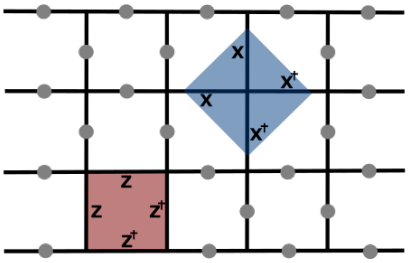


Figure 16: Illustration of A_v and B_p type surface code stabilizers on a 3×3 lattice.

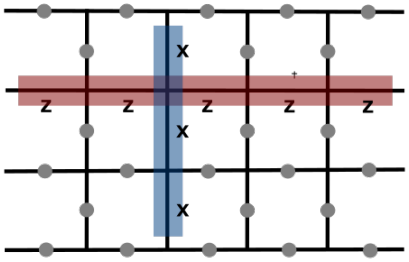


Figure 17: Representation of the logical operators $\overline{\mathsf{X}}$ and $\overline{\mathsf{Z}}$ for the surface code.

Complementary to the code, we use a hard decision renormalization decoder [283], which can be implemented for any stabilizer code with topological order, as is the case for the surface code that we will rely upon.

4.5 New multi-switching lemma for biased PTF circuits

In this section we state and prove our multi-switching lemma and depth-reduction lemma for $\mathsf{bPTF}^0[k]$ circuits. We begin by proving Lemma 12, and then proceed to proving Lemma 13.

4.5.1 Tight multi-switching lemma

The first lemma of this type will be a multi-switching lemma that reduces a finite set of depth-2 $\mathsf{bPTF}^0[k] \circ \mathsf{AND}_w$ circuits, which are defined by $\mathsf{bPTF}[k]$ gate with AND gates of bounded fan-in w as input, to a $\mathsf{DT}(w)^m \not\Rightarrow \mathsf{DT}(t)$ decision forest. It is worth noting that Theorem 4.4.2 serves as a multi-switching lemma of this type. However, a higher probability of success is required to reduce the initial circuit to a tree-like object, enabling the subsequent proof of quantum advantage with respect to the QNC^0 circuit class.
In particular, better parameters can be obtained with additional assumptions on the decision trees to which the initial circuit will be reduced. This is demonstrated in the subsequent lemma, where we combine the inductive technique of [251] with the canonical decision trees proposed by [115] for $\mathsf{bPTF}^0[k]$ circuits restricted by a random restriction ρ.

Lemma 25. *Let $\mathcal{F} = \{F_1, F_2, F_3, \ldots, F_m\}$ be a list of $\mathsf{bPTF}[k] \circ \mathsf{AND}_w$ circuits on $\{0,1\}^n$, $t > 0$, $l \geq \log(m) + k + 2$ and $p < \frac{1}{40w}$, then,*

$$\Pr_{\rho \in R_p}[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(l)^m \not\Rightarrow \mathsf{DT}(t)] \leq m \cdot 2^k (80wp)^t. \tag{4.5.1}$$

Proof. Our proof employs an inductive approach similar to the one used by [251] for his multi-switching lemma. However, this necessitates the use of downward closed random restrictions. This attribute ensures that the size of the canonical decision tree does not expand when additional variables from the initial random restriction are restricted. It guarantees the monotonicity of the canonical decision tree's size under random restriction, which is crucial for the inductive reasoning process. Therefore, we will first establish some properties of the random restrictions and then utilize them in the proof of the new multi-switching lemma using inductive techniques.

Downward closedness. We start with defining the downward closeness of random restrictions.

Definition 36. *Let $\mathcal{P}$ be a set of downward closed restrictions, and consider any $\rho \in \mathcal{P}$. Then for any other restriction ρ' which satisfies $\rho'(x_i) = \rho(x_i)$ for all $i \in I = \{i \mid \rho(x_i) \in \{0,1\} \wedge \rho'(x_i) \in \{0,1\}\}$, we have that $\rho' \in \mathcal{P}$.*

Building on this definition, our interest lies in combining restrictions, necessitating the set of restrictions used in this process to be downwards closed. This requirement can be ensured with the following lemma.

Lemma 26. *Let $\mathcal{P}$ and $\mathcal{P}'$ be two downward-closed sets, then $\mathcal{P} \cap \mathcal{P}'$ is also a downward-closed set.*

Proof. This follows directly by the definition of the downward-closed set. □

Now, we need to demonstrate that the random restrictions chosen to bind the entire probability for the multi-switching lemmas are also downwards closed. To achieve this, we must analyze the canonical decision trees generated from $F\lceil_\rho$, where F is an initial circuit of the type $\mathsf{bPTF}[k] \circ \{\mathsf{AND}, \mathsf{OR}\}_w$, and ρ represents the random restrictions of interest. We will consider two objects defined in [115]. The first is the algorithmic process that constructs the canonical decision tree from F and ρ, while the other is the resulting description of the canonical decision tree in a non-standard format as considered by the author.

Algorithm 1 Canonical Decision Tree

procedure $CDT_F(F = G(C_1, \ldots, C_n)$+ black-box access to a string $x \in \{0,1\}^n$.)
  $j* \leftarrow 0; x \leftarrow (*)^n; ctr \leftarrow 0$
  while $j* \leq m$ **do**
    Find $j \geq j*$ such that $C_j \not\equiv 0$. If no such j exists, exit the loop.
    $B_j \leftarrow$ the set of variables not used by $C_j(x)$
    Query $\alpha_{x_{B_j}}$
    Set $x_{B_j} \leftarrow \alpha_{B_j}$.
    if $C_j(x) = 1$ **then**
      $ctr \leftarrow ctr + 1$
      if $ctr = k$ **then**
        return $G(1^m)$
      end if
    end if
    $j* \leftarrow j$
  end while
  return $F(x \circ 0^n)$
end procedure

Given Algorithm 1 that produces the canonical decision tree $T_{F|_\rho}$ for $F\lceil_\rho$, we can define the following object known as an s-witness. At an information-theoretical level, an s-witness describes the depth s subtree of the canonical decision tree of $F\lceil_\rho$.

Definition 37 (t-witness [115])**.** *Let F be a* $\mathsf{bPTF}[k] \circ \mathsf{AND}_w$ *circuit and ρ a restriction. Let $t \geq 1$. Consider the tuple $(r, l_i, s_i, B_i, \alpha_i)$ where*

- *$r \in [1, t+k]$ is an integer;*
- *$(l_1, \ldots, l_r) \in [m]^r$ is an increasing list of indices;*
- *$(s_1, \ldots, s_r)$ is a list of non-negative integers, at most k of which are allowed to be 0, such that $s := \sum_{i=1}^r s_i \in [t, t+w-1]$;*

- $(B_1, \ldots, B_r)$ *is a list of subsets of* $[w]$ *satisfying* $|B_i| = s_i$*;*
- $(\alpha_1, \ldots, \alpha_r)$ *is a list of potential bit strings satisfying* $|\alpha_i| = s_i$*.*
- $(r, l_i, s_i, B_i, \alpha_i)$ *is called a* t*-witness for* ρ *if there exists an* $\alpha \in \{0,1\}^n$ *such that*
  - *When we run* $T_{F|_\rho}$ *on a* α*,* C_{l_i} *is the* i*-th term queried by* $T_{F|_\rho}$*.*
  - $T_{F|_\rho}$ *queries* s_i *variables in* C_{l_i}*, and the relative location of those variables within* C_{l_i} *are specified by set* B_i*.*
  - $T_{F|_\rho}$ *receives* α_i *in response to its* i*-th batch query.*

The size of the witness $(r, l_i, s_i, B_i, \alpha_i)$ *is defined to be* $s := \sum_{i=1}^{r} s_i$*. We may denote the size of a witness* W *as* $size(W)$*.*

Given Algorithm 1 and a t-witness as defined in Definition 37, the previous statements can be made precise with the following fact.

Fact 1. *[115] For every* ρ *such that* $F\restriction_\rho \geq \mathsf{DT}(t)$*, there exists a t-witness for* ρ*.*

Now, we can begin to establish precise statements about the random restriction of interest. Our first property is as follows, which will aid in establishing downward closedness for two types of random restrictions.

Lemma 27. *For all random restrictions* ρ *and* $F = \mathsf{bPTF} \circ \{\mathsf{AND}, \mathsf{OR}\}_w$*, for which* $F\restriction_\rho$ *has an* l*-witness, let* L *be the set of variables in the canonical decision tree* $\mathsf{DT}(l)$*. If* ρ' *is a random restriction such that* $\rho'(x_i) = \rho(x_i)\ \forall i \notin T$*, where* $T = \{*\}^t$ *and* $L \cap T = \emptyset$*, then* $F\restriction_{\rho'}$ *has the same* l*-witness.*

Proof. If a variable within T is assigned a Boolean value, it implies that all clauses preceding the first non-fixed clause, denoted by l_1 in the l-witness, retain their assigned values. Crucially, any 'alive' variables that could influence the outcome of this clause are encompassed within B_1, which are, by definition, elements of L. This assertion is supported by the "Canonical Decision Tree" algorithm and the definition of an s-witness. Thus, altering any variable in T leaves the initial segment of the l-witness unchanged. By inductively applying this reasoning to each element of the l-witness, we establish that variable assignments within T do not alter the overall structure of the l-witness. Therefore, $F\restriction_{\rho'}$ has exactly the same l-witness as $F\restriction_\rho$. □

Now we will prove that a particular set of random restrictions is downward closed.

Lemma 28. *The set of restrictions* ρ *for which a circuit* $F = \mathsf{bPTF} \circ \{\mathsf{AND}, \mathsf{OR}\}_w$ *reduces to a decision tree of the type* $\mathsf{DT}(l)$*, with* $l \in \mathbb{N}$*, is downward closed.*

Proof. We will demonstrate the lemma by considering a general scenario where any restriction ρ satisfies the condition $F\restriction_\rho \in \mathsf{DT}(l)$. For such a restriction ρ, we observe that it designates $*$ to all variables within the decision tree $\mathsf{DT}(l)$, which we will refer to as the set L. Additionally, there exists a distinct set T of variables also assigned $*$ that do not intersect with L, i.e., $T \cap L = \emptyset$.
When a variable in L is set to a Boolean value, the downward closedness property is maintained. This is because the application of the "Canonical Decision Tree" algorithm would yield an identical tree, albeit with the fixed variables omitted or potentially with further reductions. Such reductions occur if the fixed variables directly satisfy a complete clause C_{l_i}, with l_i being an index of the l-witness. Consequently, the new s-witness encompasses at most a subset of the elements from the previous l-witness.
Now, by Lemma 27, we can conclude that assigning a value to any variable in T generates a new random restriction that does not change the s-witness of the initial function under this new restriction. Therefore, the resulting canonical decision tree is the same as for ρ.
We can hence deduce that the set of restrictions for which $F\restriction_\rho \in \mathsf{DT}(l)$ is indeed downward closed. □

Induction. As mentioned previously, the precise proof of the multi-switching lemma will follow the inductive idea of the initial multi-switching theorem by [251]. Specifically, we will apply induction to the circuits rec_i^{*-1} and the input size. Whenever we encounter a circuit that does not reduce to a decision tree of the type $\mathsf{DT}(l)$, we make a query on the variables that are alive to reduce this circuit and use these as variables of the global decision tree $\mathsf{DT}(t)$.
The theorem trivially holds when $m = 0$ and $n = 0$, where m represents the number of $\mathsf{bPTF}[k] \circ \{\mathsf{AND}, \mathsf{OR}\}_w$ circuits considered, and n denotes the input size. Now, we can divide the required induction step into two cases: one where the first $F_1 \lceil \rho$ does have a decision tree of depth l, and another where it does not.

$$\Pr\left[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t)\right] = \Pr\left[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil_\rho \notin \mathsf{DT}(l)\right] \cdot \Pr\left[F_1\lceil_\rho \notin \mathsf{DT}(l)\right] \tag{4.5.2}$$

$$+ \Pr\left[\mathcal{F}\lceil_\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil_\rho \in \mathsf{DT}(l)\right] \cdot \Pr\left[\, F_1\lceil_\rho \in \mathsf{DT}(l)\right]. \tag{4.5.3}$$

For $F_1\lceil_\rho \in \mathsf{DT}(l)$, we obtain that we can simply induct on the $\mathcal{F} \setminus F_1$. By doing that we obtain the following probability,

$$\Pr\left[\mathcal{F}\lceil\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil_\rho \in \mathsf{DT}(l)\right] =$$
$$\Pr\left[(\mathcal{F} \setminus F_1)\lceil\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil_\rho \in \mathsf{DT}(l)\right] \leq (m-1) \cdot 2^k (80wp)^t.$$

To completely establish Equation (4.5.3), we would need to bound $\Pr[F_1\lceil_\rho \in \mathsf{DT}(l)]$. However, this is unnecessary as we will observe when determining the conditional probability in Equation (5.5.26). In addition to the previous argument, we need to ensure that the random restrictions guaranteeing F_1 to reduce to a decision tree of depth l are downward closed, i.e., $F_1\lceil_\rho \in \mathsf{DT}(l) \wedge \rho \in \mathcal{F}$. This has already been proven in Lemma 28, thus ensuring Lemma 26 that the same property holds for all random restrictions used in the inductive step of Equation (4.5.3).
Now, we need to analyze the case where the first circuit does not have a reduction to a small decision tree. For that, we will use the following fact.

Fact 2. *For every ρ such that $\mathsf{DT}(F\lceil_\rho) \geq t$ there exists a $t' > t$, such that $\mathsf{DT}(F\lceil_\rho) = t'$.*

This allows us to rewrite the expression which facilitates the application of the intended inductive step.

$$\Pr\left[\mathcal{F}\lceil\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil\rho \notin \mathsf{DT}(l)\right] \cdot \Pr\left[F_1\lceil\rho \notin \mathsf{DT}(l)\right] \leq$$
$$\sum_{l'>l} \Pr\left[\mathcal{F}\lceil\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \,\middle|\, F_1\lceil_\rho \in \mathsf{DT}(l')\right] \cdot \Pr\left[F_1\lceil_\rho \in \mathsf{DT}(l')\right].$$

Now, we will use precisely the size of the canonical decision tree resulting from the restriction on the first circuit $F_1\lceil_\rho$ to provide us with a second restriction τ. The idea is that we can rewrite the same probability $\Pr\left[\mathcal{F}\lceil\rho \notin \mathsf{DT}(t) \circ \mathsf{DT}(l)^m \middle|\, F_1\lceil\rho \in \mathsf{DT}(l')\right]$ as $\Pr\left[\mathcal{F}\lceil_{\rho\circ\tau} \notin \mathsf{DT}(t-l') \circ \mathsf{DT}(l)^m |\, F_1\lceil\rho \in \mathsf{DT}(l')\right]$ when $l' < t$, where τ is the assignment of the l' variables of the canonical decision tree resulting from $F_1\lceil_\rho$.

$$\sum_{l'>l} \Pr\left[\mathcal{F}\lceil\rho \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t) \;\middle|\; F_1\lceil_\rho \in \mathsf{DT}(l')\right] \cdot \Pr\left[F_1\lceil_\rho \in \mathsf{DT}(l')\right] = \tag{4.5.4}$$

$$\sum_{l'>l} \sum_{\tau \in \{0,1\}^{l'}} \Pr\left[\mathcal{F}\lceil_{\rho\circ\tau} \notin \mathsf{DT}(l)^m \Rrightarrow \mathsf{DT}(t-l') \;\middle|\; F_1\lceil_\rho \in \mathsf{DT}(l')\right] \cdot \Pr\left[F_1\lceil_\rho \in \mathsf{DT}(l')\right]. \tag{4.5.5}$$

In particular, when $l' < t$ we can bind the first term of Equation (4.5.5) inductively, while the second can

be very largely bound by the probability that $F_1\lceil_\rho \notin \mathsf{DT}(l')$,

$$\sum_{l'>l}\sum_{\tau\in\{0,1\}^{l'}} \Pr\left[\mathcal{F}\lceil_{\rho\circ\tau}\notin \mathsf{DT}(t-l')\circ\mathsf{DT}(l)^m \;\middle|\; F_1\lceil_\rho\in \mathsf{DT}(l')\right]\cdot\Pr\left[F_1\lceil_\rho\in \mathsf{DT}(l')\right]$$
$$\leq \sum_{l'>l}(20pw)^{l'}2^k2^{l'}\left(m\cdot(80wp)^{t-l'}2^k\right)$$
$$\leq (80wp)^t2^{k-1}\sum_{l'>l}2^{k+1}2^{-l'}m.$$

Using our initial assumption that $l \geq \log(m)+k+2$ yields that,

$$\sum_{l'>l}2^{-l}\cdot m\cdot 2^k \leq 1. \tag{4.5.6}$$

Therefore, the second term in Equation (5.5.26) in our initial division is inductively bounded above by $(80wp)^t2^{k-1}$ for cases where $l' < t$. When we consider this first case for the second term in Equation (5.5.26), it's imperative to ensure that the random restrictions in question maintain the property of downward closedness. To this end, we invoke Lemma 26 inductively, but we must also affirm the same property for the composition of restrictions $\rho\circ\tau$ when $F_1\lceil_\rho\in \mathsf{DT}(l')$.
Since τ assigns definitive values to all variables within $\mathsf{DT}(l')$, our task simplifies to confirming that all variables not specifically assigned by the composition $\rho\circ\tau$ do not alter the conditional probability that our induction is predicated upon. This condition is indeed satisfied, as established by Equation (4.5.5), thus ensuring the downward closedness of the random restrictions we are considering. With this confirmation, our inductive argument remains robust and consistent.
Now, to show that an equal bound holds for the case that $l' \geq t$, we obtain that the first conditional probability of Equation (5.5.26) is fixed and equal to 1, while second, we bound again with $F_1\lceil_\rho\notin \mathsf{DT}(l')$.

$$\sum_{l'>l}\Pr\left[\mathcal{F}\lceil_\rho\notin \mathsf{DT}(t)\circ\mathsf{DT}(l)^m\,\middle|F_1\lceil_\rho\in \mathsf{DT}(l')\right]\cdot\Pr\left[F_1\lceil_\rho\in \mathsf{DT}(l')\right] \leq 2^k\sum_{l'>t}(20pw)^{l'}$$
$$\leq 2^k\cdot 2(20pw)^t = 2^k\cdot 2^{2t-1}(80pw)^t.$$

For any $t>0$ we obtain that this case is smaller or equal to $2^{k-1}(80pw)^t$. Thus, if we use the two values that build up Equation (5.5.26) and the value for Equation (4.5.3) obtained, we obtain that,

$$\Pr[\mathcal{F}\lceil\rho\notin \mathsf{DT}(t)\circ\mathsf{DT}(l)^m] \leq (m-1)2^k(80wp)^t + 2^k(80wp)^t$$
$$\leq m\cdot 2^k(80pw)^t,$$

completing the proof. □

4.5.2 Depth reduction lemmas

Provided with the previous multi-switching lemma we can construct a more efficient depth reduction lemma for $\mathsf{bPTF}^0[k]$ circuits. This follows the proof technique from [284].

Lemma 29. *If* $d,t\geq 1$ *and* $F\in \mathsf{bPTF}[k;d;s_1,s_2,\ldots,s_d]\circ\mathsf{DT}(w) \not\Rightarrow \mathsf{DT}(t-1)$, $l\geq \log s_1+k+2$ *and* $p<\frac{1}{40w}$, *then*

$$\Pr[F\lceil_\rho\notin \mathsf{bPTF}[k;d-1;s_2,\ldots,s_d]\circ\mathsf{DT}(l) \not\Rightarrow \mathsf{DT}(t-1)] \leq s_1\cdot 2^k(400wp)^{t/2}. \tag{4.5.7}$$

Proof. Suppose that F is computed by a depth $t-1$ decision tree T, each of whose leaves λ is labelled by a circuit $C_\lambda \in \mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)$, we will assume the following two events,

$$\mathcal{A} \iff T\restriction_p \text{ has depth } \leq t/2-1 \tag{4.5.8}$$

$$\mathcal{B} \iff C_\lambda\restriction_p \in \mathsf{bPTF}[k; d-1; s_2, \ldots, s_d] \circ \mathsf{DT}(l) \Rightarrow \mathsf{DT}(t/2-1) \text{ for every leaf } \lambda \text{ of T} \tag{4.5.9}$$

such that,

$$\mathcal{A} \wedge \mathcal{B} \implies F\restriction_\rho \in \mathsf{bPTF}[k; d-1; s_2, \ldots, s_d] \circ \mathsf{DT}(l) \Rightarrow \mathsf{DT}(t-1). \tag{4.5.10}$$

Now, the probability that neither of the events occurs will be considered. For the first event, we obtain the value this event to be

$$\mathsf{Pr}[\neg\mathcal{A}] = \mathsf{Pr}[T\restriction_p \text{ has depth } \geq t/2] \leq (4ep)^{t/2}. \tag{4.5.11}$$

For the second, we sort to Lemma 22 to translate $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{DT}(w)$ to $\mathsf{bPTF}[k; d; s_1, s_2, \ldots, s_d] \circ \mathsf{AND}_w$, and with Lemma 25 we obtain

$$\begin{aligned}
\mathsf{Pr}[\neg\mathcal{B}] &\leq \sum_\lambda \mathsf{Pr}[C_\lambda\restriction_p \notin \mathsf{bPTF}[k; d-1; s_2, \ldots, s_d] \circ \mathsf{DT}(l) \Rightarrow \mathsf{DT}(t/2-1)]. \\
&\leq \sum_\lambda s_1 \cdot 2^k (80wp)^{t/2} && (4.5.12)\\
&\leq 2^{t-1}\big(s_1 \cdot 2^k (80wp)^{t/2}\big). && (4.5.13)
\end{aligned}$$

Combining both, we obtain,

$$\begin{aligned}
\mathsf{Pr}\big[F|_{R_p} \notin \mathsf{bPTF}[k; d; s_2, \ldots, s_d] \circ \mathsf{DT}(l) \Rightarrow \mathsf{DT}(t-1)\big] &\leq \mathsf{Pr}[\neg\mathcal{A}] + \mathsf{Pr}[\neg\mathcal{B}] \\
&\leq (4ep)^{t/2} + 2^{t-1}\big(s_1 \cdot 2^k (80wp)^{t/2}\big) && (4.5.14)\\
&\leq (4ep)^{t/2} + \big(s_1 \cdot 2^{k-1} (400wp)^{t/2}\big) && (4.5.15)\\
&\leq s_1 \cdot 2^k (400wp)^{t/2}. && (4.5.16)
\end{aligned}$$

□

The previous lemma allows one to reduce the depth of a $\mathsf{bPTF}^0[k]$ circuit. Now, this lemma can be employed sequentially to reduce the entire circuit.

Lemma 30. *Let $d, t, k, q, m, s_1, l_1, \ldots s_{d-1}, l_{d-1} \in \mathbb{N}$; $d, t > 0$; $l_i \geq \log s_i + k + 2$; $p_1, \ldots, p_d \in (0,1)$; $p_i < \frac{1}{40 l_i}$. Let $F \in \mathsf{bPTF}^0[k; d; s_1, \ldots, s_{d-1}, m]$ with n inputs and m outputs. Let $s = s_1 + \ldots + s_{d-1} + m$. Let $p = p_1 \cdot p_2 \cdot \ldots \cdot p_d$. Then,*

$$\begin{aligned}
\Pr_{\rho \sim R_p}\big[F\restriction_\rho \notin \mathsf{DT}(q-1)^m \Rightarrow \mathsf{DT}(2t-2)\big] &\leq s_1 \cdot 2^k \mathcal{O}(p_1)^{t/2} \\
&\quad + \left(\sum_{i=2}^{d-1} s_i \cdot 2^k \mathcal{O}(p_i l_i)^{t/2}\right) + 2^k m^{1/q} \mathcal{O}(p_d \cdot l_d)^t.
\end{aligned}$$

Proof. The first step is to consider our function f as a $\mathsf{bPTF}^0[k; d; s_1, \ldots, s_{d-1}, m] \circ \mathsf{DT}(1) \Rightarrow \mathsf{DT}(t-1)$ circuit. Then, to reduce this object sequentially through random restrictions we use Lemma 29 and p_1, such that

$$\mathsf{Pr}\big[F\restriction_\rho \notin \mathsf{bPTF}^0_{d-1}(k; d-1; s_2, \ldots, s_n) \circ \mathsf{DT}(l_1) \Rightarrow \mathsf{DT}(t-1)\big] \leq s_1 \cdot 2^k (400 p_1)^{t/2}$$

$$\leq s_1 \cdot 2^k \mathcal{O}(p_1)^{t/2}.$$

We will denote $F\lceil_\rho \notin \mathsf{bPTF}^0_{d-1}(k; d-1; s_2, \ldots, s_n) \circ \mathsf{DT}(l_1) \not\Rightarrow \mathsf{DT}(t-1)$ as event E_1, and subsequent depth reductions equally for E_i. Building on this and based on the same Lemma 29, we obtain that,

$$\Pr[\neg E_i | E_1 \wedge \ldots \wedge E_{i-1}] \leq s_i \cdot 2^k (400 l_i p_i)^{t/2} \tag{4.5.17}$$

$$\leq s_i \cdot 2^k \mathcal{O}(p_i \cdot l_i)^{t/2}. \tag{4.5.18}$$

Now after applying $d-1$ times Lemma 29 we obtain that G defined as $G = F|_{p_1 \circ \ldots \circ p_{d-1}}$ is

$$G \in (\mathsf{bPTF}[k] \circ \mathsf{DT}(l_{d-1}))^m \not\Rightarrow \mathsf{DT}(t-1), \tag{4.5.19}$$

with a probability equal to the probability with the following sequence of event $E_1 \wedge E_2 \wedge \ldots \wedge E_{d-1}$ hold. Now, to the leaves of each of the global decision tree $\mathsf{DT}(t-1)$, we apply Theorem 4.4.2, such that for each leave, we have that,

$$\Pr\left[G_{\lambda|_{\rho_d}} \notin \mathsf{DT}(q-1)^m \not\Rightarrow \mathsf{DT}(t-1)\right] \leq 4\Big(64(2^k m)^{1/q} p_d \cdot l_d\Big)^t \tag{4.5.20}$$

$$\leq 2^k m^{1/q} \mathcal{O}(p_d \cdot l_d)^t. \tag{4.5.21}$$

Then, we apply the union bound to all the leaves, which multiplies these values by 2^{t-1}. However, this value will be upper bounded by the same value previously defined. Thus, we obtain that,

$$\Pr\left[\neg E_1 \wedge \neg E_2 \wedge \ldots \wedge E_d\right] = \sum_{i=1}^{d} Pr[\neg E_i | E_1 \wedge \ldots \wedge E_{i-1}] \tag{4.5.22}$$

$$\leq s_1 \cdot 2^k \mathcal{O}(p_1)^{t/2} + \left(\sum_{i=2}^{d-1} s_i * 2^k \mathcal{O}(p_i \cdot l_i)^{t/2} \right) + 2^k m^{1/q} \mathcal{O}(p_d \cdot l_d)^t. \tag{4.5.23}$$

□

In the final step, we provide a value for the probabilities of the various random restrictions and dimensions for the local decision trees to derive a new lemma. This describes based on the previous values, and the size of the global decision tree t, the size of the initial circuit s, and the type of circuit parametrized by k the probability of the success of reducing the circuit to decision tree of the type $\mathsf{DT}(q-1)^m \not\Rightarrow \mathsf{DT}(2t-2)$.

Lemma 31. *Let $F : \{0,1\}^n \mapsto \{0,1\}^m$ be an $\mathsf{bPTF}^0[k]$ circuit of size s, depth d. Let $p = 1/(s_d^{1/q} \cdot \mathcal{O}(\log(s)^{d-1} \cdot k^{d-1}))$. Then,*

$$\Pr_{\rho \sim R_p} [F\lceil_\rho \notin \mathsf{DT}(q-1)^m \not\Rightarrow \mathsf{DT}(2t-2)] \leq s \cdot 2^{-t+k}\,. \tag{4.5.24}$$

Proof. This follows by applying Lemma 30 with $p_1 = 1/\mathcal{O}(1)$ and $p_2 = \ldots = p_{d-1} = 1/\mathcal{O}(\log(s)+k)$ and $p_d = 1/\mathcal{O}(s_d^{1/q} \cdot (\log(s) + k))$. □

4.6 Quantum advantage in the noiseless case

We now proceed to formally state and prove our separations of noiseless shallow quantum circuits against biased threshold circuits. We divide the analysis into two sections, the first focusing on qubit QNC^0

circuits, and the second on higher dimensional qupit QNC^0 circuits. In each case we deal with the respective family of ISMR problems that demonstrates the separation from $\mathsf{bPTF}^0[k]$. We also highlight qualititative and quantitative differences in the results achieved in the two cases, as well as the nature of the corresponding challenges and proof methodologies.
Let us first recall the ISMR family of problems.

Definition 38. *For any positive integer p, the Inverted Strict Modular Relation Problem* (ISMRP) $\mathcal{R}_p^m : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ *with input and output lengths n and m respectively is defined as follows. For any $x \in \mathbb{F}_2^n$ such that $|x|$ Mod $p = 0$*

$$\mathcal{R}_p^m(x) := \left\{ y \;\middle|\; y \in \mathbb{F}_2^m :\ |y| \text{ Mod } p = -\left(|x|/p\right) \text{ Mod } p \right\}, \tag{4.6.1}$$

where $|x| = \sum_{i=1}^n x_i$ is the Hamming weight of the n-bit string x.

Here, we use the term "strict" to denote that the problems require the outcome to be congruent modulo p to the additive *inverse* of $|x|/p$, rather than taking the interpretation of the Hamming weight of the input being congruent to zero modulo p^2.
As observed in Section 4.3, for each p the ISMRP is a generalization of the Parity Halving Problem introduced and studied by [27]. While several generalizations of PHP may be possible, ours was influenced by a specific family of non-local games for which we have derived an exponential separation between the success probabilities of the best classical and quantum winning strategies, as we shall see in Section 4.6.2. Note that we consider this binary setting as it provides a fair basis for a separation between the binary $\mathsf{bPTF}^0[k]$ circuits and QNC^0 circuits over qupits, not changing the arithmetic of the problem to dits, which could favor the quantum circuits and weaken the classical circuits class.
In the case of the qubit, we will prove that QNC^0 circuits can solve the $\mathcal{R}_2^m$ problem exactly. Additionally, we establish exact-case and average-case hardness bounds for $\mathsf{bPTF}^0[k]$. We use the term "exact case" to refer to the scenario where both quantum and classical circuits are required to solve the proposed problem with probability one (i.e. with certainty) for all inputs. Although this setting was implicitly considered in previous lower bounds, such as those determined in [27], it was not explicitly explored. We address this gap and demonstrate that it is possible to establish stronger lower bounds on the size of the classical circuits for this scenario. Following this, we will present the average-case hardness for all other ISMR problems, $\mathcal{R}_p^m$ where $p > 2$ is prime, along with the respective quantum (i.e. qupit QNC^0) upper bounds.

4.6.1 Separations for qubit cases

The qubit QNC^0 case is special because it enables us to solve one of the problems in $\mathcal{R}_2^m$ exactly, meaning it produces a correct outcome with probability 1 for all input strings. We leveraged this property to obtain larger bounds for the size of the first AC^0 and then $\mathsf{bPTF}^0[k]$ circuits required to solve the same problems, thereby achieving greater separations with QNC^0 beyond the previous results for AC^0 and simultaneosly completely new bounds with $\mathsf{bPTF}^0[k]$. The following theorem formally establishes this separation for QNC^0 with $3D$ and all-to-all connectivities.

Theorem 4.6.1. *There exists a* QNC^0 *circuit that solves the $\mathcal{R}_2^m$ exactly with a subquadratic number of gates, which an abritary* $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ *circuit of depth d with access to a random string* rpoly *and parameter k that does require size s circuit lower bounded as follows,*

$\Omega(s)$	*3D*	*All-to-all*
$k = \mathcal{O}(1)$ $(\mathsf{AC}^0/\mathsf{rpoly})$	$\exp\left(\left(\frac{n^{1/3}}{(\log n)^{1+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)$	$\exp\left(\left(\frac{\sqrt{n}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)$
$k = n^{1/(5d)}$ $(\mathsf{bPTF}^0[k]/\mathsf{rpoly})$	$\exp\left(\left(\frac{n^{2/15}}{(\log n)^{1+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)$	$\exp\left(\left(\frac{n^{3/10}}{(\log n)^{3/2+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right)$

In particular, for this result, we will utilize the fact that we can reduce the AC^0 and $\mathsf{bPTF}^0[k]$ circuits to a first decision forest using Lemma 31, and then to forest lines with a second application of random restrictions. We define a forest line as a set of individual decision trees, where each computes one of the final outcome bits. We will then demonstrate that for this object to produce exact solutions, the decision trees comprising the forest line require a minimum depth with Lemma 34, which consequently imposes a minimum size on the initial AC^0 and $\mathsf{bPTF}^0[k]$ circuits. More concretely, this is achieved by showing that the decision trees in this forest line require a minimum depth to produce terms of degree 2 with respect to the Algebraic Normal Form of the Boolean function they evaluate. Additionally, we establish a lower bound on the number of such degree 2 terms necessary to evaluate the relation problem exactly Lemma 33. Therefore, by combining Lemma 31 and Lemma 34 we obtain our classical lower bound enunciated in Lemma 35.
In a subsequent subsection, we present quantum circuits that solve these problems exactly with $3D$ and all-to-all connectivities in Lemma 36 and Corollary 5 respectively. By integrating the quantum upper bounds, from which we obtain all the parameters, with the classical lower bounds, we derive the previously stated separation.
Our second result concerning the qubit QNC^0 circuits is an average-case hardness separation from the $\mathsf{bPTF}^0[k]$ circuits with a super-logarithmic size bias parameter k, addressing the same $\mathcal{R}_2^m$ problem, as stated in the following theorem.

Theorem 4.6.2. *There exists a* QNC^0 *circuit that solves the* $\mathcal{R}_2^m$ *over a uniform distribution of even strings exactly with a subquadratic number of gates, which an arbitrary* $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ *circuit of depth* d *with access to a random string* $rpoly$ *and parameter* k *solves the problem with probability at most,*

$\Omega(s)$	*3D*	*All-to-all*
$k = n^{1/(5d)}$	$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n^{4/15-\mathcal{O}(1)}}{(\log s)^{2d-2}}\right)\right)$	$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n^{3/5-\mathcal{O}(1)}}{(\log s)^{2d-1}}\right)\right)$

In this setting, we once again utilize Lemma 31 to reduce the initial $\mathsf{bPTF}^0[k]$ circuit to a decision forest and then, with a second application of random restrictions, to a forest line. However, in contrast to the previous case, we can directly apply Theorem 4.6.3, which determines the maximal average-case hardness of the forest line. This can then be correlated with the hardness of the $\mathsf{bPTF}^0[k]$ circuits in solving the $\mathcal{R}_2^m$ problem, a relationship we formalize in Corollary 4.
Afterward, we once again refer to the same quantum upper bounds Lemma 36 and Corollary 5 to characterize all the parameters and optimize the free parameters to achieve the largest possible separation.

Exact and Average-case $\mathsf{bPTF}^0[k]$ hardness

To solve exactly and deterministically an ISMR problem $\mathcal{R}_q^m$, in the sense that to each input x a single valid output string y is created, one can demonstrate that for any of these solutions, the outcome bits y_i of y are individual Boolean functions of the type $f_{y_i} : \{0,1\}^n \mapsto \{0,1\}$. Furthermore, these individual functions do not have to fulfill any specific conditions, some of these could even be constant functions. However, the parity of their outcomes must adhere to a property dictated by the relation problem. Specifically, the parity of these output bits computes another Boolean function.
Now we would like to specify this Boolean function specifically for the case of $\mathcal{R}_2^m$. In this case, for all the strings that are even we have,

$$\mathcal{D}_2(x) = \begin{cases} 1, & |x| \text{ Mod } 4 = 0 \\ 0, & |x| \text{ Mod } 4 = 2. \end{cases} \tag{4.6.2}$$

This, however, does not define a total Boolean functions, as it only defines the images over even input strings. Moreover, this implies that any mapping of the odd input strings while respecting the previous mapping for the even input strings creates a valid Boolean function for the parity of the outcome bits. In particular, we can now derive which is exactly the set of valid Boolean functions.

Lemma 32. *The parity of the outcome string y of a* bPTF$^0[k]$ *circuit that does solve the $\mathcal{R}_2^m$ for abritary input $x \in \{0,1\}^n$ is contained in the following set*

$$S_{\mathcal{D}_2} = \Bigg\{ \underbrace{\bigoplus_{i_1=1}^{n-1} x_{i_1} \Bigg(\bigoplus_{i_2=i_1+1}^{n} x_{i_2} \Bigg)}_{\textit{Even mapping}} \oplus \underbrace{\Bigg(\bigoplus_{j=1}^{n} x_j \Bigg) \wedge g}_{\textit{Odd mapping}} \;\Bigg|\; g \in \{0,1\}^n \mapsto \{0,1\} \Bigg\}. \tag{4.6.3}$$

Proof. The correctness of the previous functions can be analyzed simply by considering two cases: one where the input strings are even, which obliges it to map the same decision version of $\mathcal{D}_2$, and another where the input strings are odd, necessitating consideration of all possible mappings.

Even case. In the case where the input strings are even, the expression $\left(\bigoplus_{j=1}^{n} x_j\right) \wedge g$ contributes nothing to the outcome. This means that only the first expression has any impact on the outcome. Furthermore, the first expression is known as the "Second least significant bit" function, which computes the second bit of a binary representation of the Hamming weight of the initial bit strings [95]. More precisely, it has the following definition:

$$\mathsf{LSB}(x) = \begin{cases} 1, & |x| \text{ Mod } 4 \in \{0,1\} \\ 0, & |x| \text{ Mod } 4 \in \{2,3\}. \end{cases} \tag{4.6.4}$$

It immediately follows that for even input strings, all the elements of this set map the function to the same outcome as the parity of $\mathcal{R}_2^m$.

Odd case. For the odd case, we can be assured that the set contains all possible mappings of odd input strings. This is deduced by recognizing that the contribution of the second term is now equivalent to an arbitrary Boolean function g. Additionally, this allows for the mapping of any function such that it can be decomposed as $\bigoplus_{i_1=1}^{n-1} x_{i_1} \left(\bigoplus_{i_2=i_1+1}^{n} x_{i_2} \right) \oplus h$, with h being another arbitrary Boolean function. Thus, through this decomposition, we can cancel out the effect of the first term and obtain any arbitrary mapping for odd input strings. □

We now focus on a specific representation of these Boolean functions: their Algebraic Normal Form (ANF), which we previously used in Lemma 32. This representation is crucial for capturing some common structure shared by all elements in the set of valid Boolean functions. In particular, the subsequent lemma formalizes a key aspect of this structure - the number of degree-two terms present in the ANF.

Lemma 33. *Every Boolean function belonging to the set $S_{\mathcal{D}_2}$ contains at least $\Omega(n^2)$ degree-two terms in its Algebraic Normal Form (*ANF*).*

Proof. First, we will demonstrate that any function of the set $S_{\mathcal{D}_2}$ does have a $\Omega\left(n^2\right)$ term of degree two. We can conduct this analysis using the expression for $S_{\mathcal{D}_2}$ in Equation (4.6.3).
Consider the left-hand side, where the number of such terms is given by $\binom{n}{2}$. Now, let's examine how terms of degree-two can arise from the second part of the expression. Specifically, we consider terms generated by the operation $\bigoplus_{j=1}^{n} x_j \wedge g$, where g represents a function of degree-one terms, as higher degree terms do not generate any degree-two terms. Note that degree-three terms either maintain or increase the degree, and any degree-two terms either increase the degree further or occur twice, cancelling each

other out. It's crucial to note that if a degree-three term is generated in this process, it is subsequently cancelled out in the final ANF. Given this setup, we can generate the parity function from the first part of the expression. However, the degree-two terms that emerge from the k degree-one terms in f are bounded above by,

$$h(k) = \sum_{i=0}^{k}(n-i) - \binom{k}{2} = k(n-1) - \frac{k(k-1)}{2}. \tag{4.6.5}$$

This expression accounts for the number of terms that are generated in total minus the ones that repeat and are immediately canceled out. Taking the derivative with respect to k, we find $\frac{dh}{dk} = n - 2k$. This derivative indicates that $h(k)$ reaches its maximum when $k = \frac{n}{2}$. At this point, the maximum number of degree 2 terms generated by f can be computed. By examining the limit as n approaches infinity, we find $\lim_{n\to\infty} \frac{g(n/2)}{\binom{n}{2}} = \frac{1}{2}$. This implies that for large n, approximately half of the potential degree 2 terms are effectively generated and not canceled out by the expression. Consequently, we conclude that any valid Algebraic Normal Form (ANF) must contain $\Omega(n^2)$ degree 2 terms, as intended. □

Simultaneously, we obtain that the ANF of the various $\mathsf{bPTF}^0[k]$ circuits have to equal one element of the set $S_{\mathcal{D}_2}$ if the circuit does compute correctly the $\mathcal{R}_2^m$ problem. We can formalize the previous property as

$$\mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_1}\big) \oplus \mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_2}\big) \oplus \ldots \oplus \mathsf{ANF}\big(\underbrace{\mathsf{bPTF}^0[k](x)}_{y_m}\big) \in S_{\mathcal{D}_2}. \tag{P1}$$

Subsequently, we will build a set of arguments to prove a lower bound on the size of these circuits to fulfill this property. However, we first solve a simpler instant, where we consider decision trees $\mathsf{DT}(q)^m$ in place of the $\mathsf{bPTF}^0[k]$ circuits. This will not be a problem because later on we will be able to reduce these circuits to the previous objects. Moreover, each decision tree $\mathsf{DT}(q)$ of this forest line $\mathsf{FL}(m,d) := \bigoplus_{i=0}^{m} \mathsf{DT}(q_i)$ with $d = \max q_i$, has a unique description as an ANF over the variables that are part of its nodes. Obtaining with that a second property over this computational structures as follows,

$$\mathsf{ANF}\big(\mathsf{FL}(m,q)\big) = \mathsf{ANF}\big(\underbrace{\mathsf{DT}(q)}_{y_1}\big) \oplus \mathsf{ANF}\big(\underbrace{\mathsf{DT}(q)}_{y_2}\big) \oplus \ldots \oplus \mathsf{ANF}\big(\underbrace{\mathsf{DT}(q)}_{y_m}\big) \in S_{\mathcal{D}_2}. \tag{P2}$$

We want to prove the minimum decision tree depth concerning the largest decision tree depth in the previous object so that property $P2$ is fulfilled.
For that, we will consider that for any valid Boolean function, the number of degrees two terms in its ANF is lower bounded by a quadratic term, as shown in Lemma 33. Simultaneosly, we will consider Algorithm 2 which translates an arbitrary decision tree to its ANF. Here, again one can bind the number of terms of finite degree that a certain decision tree can have in its ANF. Combining these two ideas we obtain that there exists a minimum decision tree depth such that one can compute exactly one of the Boolean functions at hand.

Lemma 34. *For any forest line* $\mathsf{FL}(m,q)$ *that computes a Boolean function from the set* $S_{\mathcal{D}_2}$ *does have a depth bounded by* $q = \Omega\big(\frac{n}{m^{1/2}}\big)$.

Proof. This involves establishing an upper bound on the number of degree-2 terms generated by a decision tree, we examine the algorithm described in Algorithm 2 for converting a binary decision tree into its Algebraic Normal Form (ANF). Initially, the algorithm identifies paths that culminate in a Boolean true value, each represented as a clause. These clauses are subsequently merged using the logical OR operation, thereby formulating a polynomial in $\mathbb{F}_2$ that encapsulates the decision tree's Boolean function.
Focusing on the clauses formed with the logical AND operation during ANF creation. We note that these can be described by the path they follow in the tree that can be decomposed by the various left (L) and right (R) turns at each node assigned with a variable x_i.

Left Turn (L). Incorporates the variable x_i directly as

$$path_{new} = path_{old} \wedge x_i \tag{4.6.6}$$
$$= \ldots \wedge x_{i-1} \wedge x_i \text{ OR } = \ldots \wedge (x_{i-1} \oplus 1) \wedge x_i \tag{4.6.7}$$
$$= \ldots x_{i-1}x_i \text{ OR } = \ldots x_{i-1}x_i \oplus x_i. \tag{4.6.8}$$

This case increases the degree of the new clause designated as $path_{new}$ by 1.

Right Turn (R). Incorporates the complement of x_i as

$$path_{new} = path_{old} \wedge (x_i \oplus 1) \tag{4.6.9}$$
$$= (path_{old} \wedge x_i) \oplus path_{old}. \tag{4.6.10}$$

While the second case does not alter the degree of the new clause designated as $path_{new}$.
Given these transformations, degree-2 terms arise from paths with specific sequences of turns. A single left turn surrounded by right turns R . . . RLR . . . , R can contribute to a degree-2 term. Furthermore, a path with two left turns, each potentially followed by right turns R . . . RLR . . . RLR . . . R, also contributes to degree-2 terms, as each left turn increases the degree by 1. The OR operations, which combine these paths, preserve the degrees of the terms in the resulting expression or increase it. Consequently, to enumerate the maximum number of degree-2 terms, it suffices to count the paths that can yield such terms. For a decision tree of depth q, the relevant paths are those with one or two left turns. The total count of such paths is given by the sum of paths with one left turn and those with two left turns $\binom{q}{2} + \binom{q}{1}$. This formulation provides a definitive upper bound on the number of degree-2 terms in the ANF representation of a Boolean function derived from a decision tree, predicated on the tree's depth.
In the end, we consider that in forest line $\mathsf{FL}(m, d)$, there are m decision trees generating degree two terms and each can generate at most $\binom{q}{2} + \binom{q}{1}$ such terms[8]. Finally, combining the previous bound with the minimum number of degree two terms necessary to prove in Lemma 33 we obtain that the minimum depth will depend on the size of the forest line and the size of the input strings as follows $q \geq \frac{n}{m^{1/2}}$.

□

Now we will be able to prove a lower bound for the size of any $\mathsf{bPTF}^0[k]$ circuit that does compute the $\mathcal{R}_2^m$ exactly. This will be followed by applying our new switching lemmas created in Section 4.3.4 to reduce the initial circuits to a decision tree of the type $\mathsf{DT}(w)^m \Rightarrow \mathsf{DT}(t)$. Subsequently, we will apply a second set of random restrictions, that reduce the same object to a forest line $\mathsf{FL}(m, q)$ for which we will be able to prove a minimum depth, and consequently link this depth to the minimum size of the initial circuit.

Lemma 35. *For sufficiently large n, any $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit depth $d \geq 4$ and $k \leq n^{1/(5d)}$ that solves the $\mathcal{R}_2^m$ has size no smaller then $s \geq 2^{\left(\frac{n}{k^d \cdot q \cdot m^{1/2+1/q}}\right)^{1/(d-1)}}$, with $q \in \mathbb{N}_{>0}$.*

Proof. This proof will be decomposed first into the application of a random restriction ρ with a probability equal to $p = \frac{1}{m^{1/q} \cdot \mathcal{O}(\log(s)^{d-1} \cdot k^d}$. Notice that this is the largest probability one can use without disrespecting the various switching lemmas, which have the minimal conditional of generating probabilities below 1. We additionally choose $t = pn/8$ and that $s \leq \exp\left(n^{1/(2d-2)}\right)$, which will not interfere with our result given that this value will be asymptotically larger than the lower bounds of s derived with this assumption. With that, we first show that $s \leq 2^{t/2}$, this results from the following comparison and the previously defined

[8] Notice, that we assume the existence of an efficient method to distribute the n variables across m decision trees in a way that generates all the corresponding degree-two terms. Specifically, this task aligns with the Steiner system problem, where we seek solutions for the $S(2, m/n, n)$ system. However, the existence of such a solution is not guaranteed, and the determination of the minimal set size remains an unresolved problem. Consequently, we will limit our consideration to the trivial upper bound.

Algorithm 2 Decision Tree to ANF Converter

procedure ANF(DT).
  $S = \emptyset$, $\mathsf{ANF}_{\mathsf{DT}} = \mathsf{False}$.
  $S' =$RecCl$(\mathsf{DT}, root, \mathsf{True}, S)$
  for $i = 0$ until $|S'|$ **do**
    $S[i] \vee \mathsf{ANF}_{\mathsf{DT}}$.
  end for
  return Reduce$(\mathsf{ANF}_{\mathsf{DT}})$.
end procedure

procedure RecCl(DT, $node$, $path$, set).
  if DT[node]=True **then**
    return $set = set \uplus \{path\}$.
  else
    if $\mathsf{DT}[node] = \mathsf{False}$ **then**
      return set.
    else
      return RecCl$(\mathsf{DT}, \mathsf{Left}(\mathsf{DT}, node), \mathsf{Var}(node) \wedge path, set)$
      $\uplus$ RecCl$(\mathsf{DT}, \mathsf{Right}(\mathsf{DT}, node), (1 \oplus \mathsf{Var}(node)) \wedge path, set)$
    end if
  end if
end procedure

procedure $\wedge(\bigoplus_{i=0}^{k_i} x_i, \bigoplus_{j=0}^{k_j} y_j)$
  return $\bigoplus_{i\in[k_i],j\in[k_j]} x_i y_j$
end procedure

procedure $\vee(x, y)$
  return $x \oplus y \oplus xy$
end procedure

values for the parameters,

$$2^{t/2} = exp\big(\Omega(pn)\big) = exp\Big(\Omega\Big(\frac{n^{(d-1)/(2d-2)-\frac{1}{5}}}{m^{o(1)}}\Big)\Big). \tag{4.6.11}$$

Also, we have that for any $d \geq 4$ then $s = \exp\big(n^{1/(2d-2)}\big) \leq exp\Big(\Omega\Big(\frac{n^{(d-1)/(2d-2)-\frac{1}{5}}}{m^{o(1)}}\Big)\Big)$. This ensures, as established by Lemma 31, that with a probability exceeding $1 - s^2 \cdot 2^{-t}$, which is greater than $1 - 2^{-t/2}$ and consequently surpasses $\exp(-\Omega(pn))$, the initial $\mathsf{bPTF}^0[k]$ circuit simplifies to a tree of the type $\mathsf{DT}(q-1)^m \nRightarrow \mathsf{DT}(2t-2)$.
The second element is that at least $pn/2$ variables remain active with very high probability. This conclusion is reached directly through the application of the Chernoff bound, given that the probability of a variable staying active is $\Omega(pn)$. In conjunction with the aforementioned probability, this suggests that

$$\Pr\big[F\restriction_\rho \in \mathsf{DT}(q-1)^m \nRightarrow \mathsf{DT}(2t-2), (x \circ \rho)_{S\subseteq[n]} = * \text{ for } |S| \geq pn/2\big] = 1 - \exp(-\Omega(pn)). \tag{4.6.12}$$

A second random restriction τ will be applied to the $2t$ variables that are in the global decision tree of

$\mathsf{DT}(q-1)^m \not\Rightarrow \mathsf{DT}(2t-2)$ after applying the first random restriction ρ. This random restriction then reduces with high probability each one of the initial $\mathsf{bPTF}^0[k]$ circuits that compute the resulting outcome bits to a forest line $\mathsf{FL}(m, q-1)$. Therefore, for all the variables that are kept alive, the m output bits are computed with local decision tree $\mathsf{DT}(q-1)$,

$$\Pr\left[F\lceil_{\tau\circ\rho} \in \mathsf{FL}(m, q-1), (x\circ(\tau\circ\rho))_{S'\subseteq[n]} = \text{ for } |S| \geq pn/4\right] = 1 - \exp(-\Omega(pn)). \quad (4.6.13)$$

Now as we have been using p-random restrictions we obtained that sampling even strings consistent with the restrictions does provide us with a uniform distribution over the even input strings. Exactly, the same argument does work over the restriction τ given that the variables of the global decision tree are selected randomly in $\{0,1\}^{2t}$. This property of p-random restrictions will now be applied to the Lemma 32 defining the necessary outcome that these resulting forest line $\mathsf{FL}(m, q-1)$ has to produce. In particular, we can simply consider the effect of the random restrictions on the variables as,

$$S_{\mathcal{D}_2}\lceil_{\tau\circ\rho} = \left\{ \bigoplus_{i_1,i_2\in[n]\setminus[\tau\circ\rho]} x_{i_1}x_{i_2} \right) \oplus \left(\bigoplus_{j\in[n]\setminus[\tau\circ\rho]} x_j \right) \wedge g\lceil_{\tau\circ\rho} \;\middle|\; g \in \{0,1\}^n \mapsto \{0,1\} \right\} \quad (4.6.14)$$

with $[\tau\circ\rho]$ representing the set of indexes of the variables that are assigned values by the respective random restriction. We obtain that the reduced expression does keep exactly the same structure. This demonstrates that over the $pn/4$ variables alive Lemma 33 applies equally. Thus, we can apply Lemma 34 considering the variables alive and the resulting forest line. Consequently, such that this object computes correctly our relation problem we have that,

$$\begin{aligned} q &\geq \frac{(p\cdot n)}{m^{1/2}} = \frac{n}{m^{1/2+1/q}\log(s)^{d-1}k^d} \\ \log(s)^{d-1} &\geq \frac{n}{m^{1/2+1/q}k^d q} \\ s &\geq 2^{\left(\frac{n}{m^{1/2+1/q}k^d\cdot q}\right)^{1/(d-1)}}. \end{aligned}$$

The proof remains applicable in cases where the $\mathsf{bPTF}^0[k]$ circuit receives a random string as advice. This is grounded in our demonstration that no deterministic circuit can accurately compute the function. Therefore, any probabilistic strategy relying on random advice is destined to fail, as it merely selects from the available deterministic solutions. As a result, $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuits are equally susceptible to failure under these conditions. □

After obtaining a tight lower bound for the size of the $\mathsf{bPTF}^0[k]$ circuits that compute exactly the modular relation problem that one can solve with QNC^0 circuits on qubits. We intend to also create the average-case hardness scenario which will be of interest by itself as it does also serve to build a noise-resistance version of this quantum-classical separation. In particular, we will make use of the following theorem which will be applied to the forest line to obtain the correlation bound in contrast to our previous size bound for the exact case.

Theorem 4.6.3. *[27] Let C be an $\mathsf{NC}^0/\mathsf{rpoly}$ circuit with n inputs, m outputs, and locality l. Then C solves $\mathcal{R}_2^m$ on a random even-parity input with probability at most $\frac{1}{2} + 2^{-\Omega(\min(n, \frac{n^2}{l^2 m}))}$*

The previous theorem does provide us with the tool to relate the resulting forest line, which we obtain from the set of random restrictions reducing the $\mathsf{bPTF}^0[k]$ circuits, with the probability with which it computes correctly the $\mathcal{R}_2^m$ problem.

Corollary 4. *For sufficiently large n and $q \in \mathbb{N}_{>0}$, any* $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ *circuit of depth $d \geq 4$, size $s \leq \exp\left(n^{1/(2d-2)}\right)$ and parameter $k \leq n^{1/(5d)}$ has small correlation with $\mathcal{R}_2^m$, bounded by,*

$$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n^2}{2^{2q} m^{1+2/q} (\log s)^{2d-1} k^{2d}}\right)\right). \tag{4.6.15}$$

Proof. The initial part of the proof follows equally to Lemma 35 with a first random restriction with ρ with a probability equal to $p = \frac{1}{m^{1/q} \cdot \mathcal{O}(\log(s)^{d-1} \cdot k^d}$ while choosing $t = pn/8$, and second random restriction τ will be applied to the $2t$ variables of the global decision tree of $\mathsf{DT}(q-1)^m$ resulting from the random restriction with probability $1 - \exp(-\Omega(pn))$.
Thus, using p-random restrictions, we find that sampling even strings that are consistent with the restrictions provide a uniform distribution over even input strings. Similarly, this argument applies to the restriction τ, considering that the variables for the global decision tree are randomly selected from $\{0,1\}^{2t}$. This characteristic of p-random restrictions enables the direct use of Theorem 4.6.3. Additionally, considering that there is a high probability that more than $pn/4$ variables remain alive, and the circuit is required to solve the $\mathcal{R}_2^m$ problem with local decision trees of depth $q-1$, it follows that these trees depend on at most 2^{q-1} of the active variables. Consequently,

$$\begin{aligned}\Pr\left[\mathsf{bPTF}_d^0(k) \text{ solves } \mathcal{R}_2^m(x),\ |x| \bmod 2 = 0\right] &\leq \frac{1}{2} + 2^{-\Omega\left(\frac{(pn)^2}{2^{2q} m}\right)} \\ &\leq \frac{1}{2} + 2^{-\Omega\left(\frac{n^2}{2^{2q} m^{1+2/q} \mathcal{O}(\log(s)^{2d-2} k^{2d}}\right)}.\end{aligned}$$

The proof also holds for scenarios where the $\mathsf{bPTF}^0[k]$ circuit is provided with a random string as advice. This is because any $\mathsf{bPTF}^0[k]$ circuit is reduced to the same entity after the application of a random restriction. Furthermore, as established in Theorem 4.6.3, we know that any probabilistic strategy informed by random advice will fail. Consequently, $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuits are also subject to failure under these conditions. □

Both these classical lower bounds left the parameters q, m, and k undefined as we intend to study the optimal values after the consideration of the quantum circuit solving the problem. This is motivated not only by the optimal asymptotic bound but yes by the smallest values of the input for which the quantum circuits demonstrate an advantage over the classical circuit classes.

Qubit upper bound

We will consider the optimal QNC^0 circuit for the quantum upper bound that exactly solves the $\mathcal{R}_2^m$ that maximizes the quantum-classical separation. For that, we rewrite the circuit in the measurement-based quantum computation description with 3D connectivity, while also considering all-to-all connectivity which is effectively a type of connectivity realizable by certain quantum hardware platforms [127].

Lemma 36. *There exists a* QNC^0 *circuit with, 3D connectivity, of depth 4 and subquadratic size that does solve the $\mathcal{R}_2^m$ exactly with $m = \mathcal{O}(n^{4/3})$.*

Proof. The proof will follow based on the demonstration that the subsequent circuit does compute the $\mathcal{R}_2^m$. More precisely, by describing and proving the correctness of the following 4 stages marked on it.

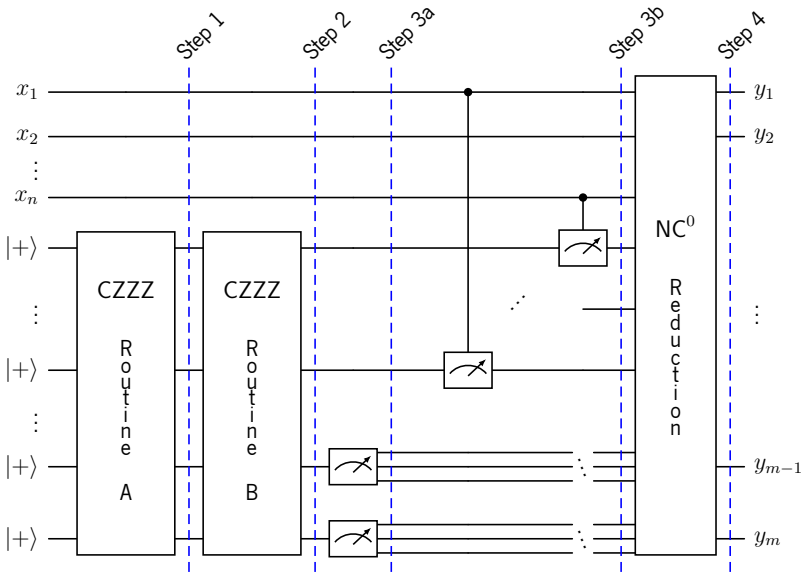


Figure 18: Illustration of the MBQC-based constant-depth quantum circuit solving an instance of the inverted strict modular relation problem. Stages 1 and 2 involve creating the graph states. In Stage $3a$, the edge qubits of the graph are measured, producing the poor-man's cat state and fanning out the measurement outcomes of the edge measurements. In Stage $3b$, phase operations are applied based on the input, and the entire state is measured. Stage 4 involves the classical post-processing.

Step 1 & **2.** The circuit begins in the Hadamard basis, and both CZZZ Routine A and CZZZ Routine B aim to generate 3D graph states without cycles. For simplicity we will first show that we can create a more densely connected 3D graph state with a simple cubic unit cell in the lattice. To achieve this structure, one needs to apply a CZ gate between each vertex (acting as the control) and its neighboring edge qubit (acting as the target). Employing single CZ gates or finite multi-control $(\mathrm{C})^{\otimes t}\mathrm{Z}$ gates would necessitate a depth at least equal to the coloring number of this 3D graph, resulting in a minimum depth of $d = 6$. However, from the perspective of the edges, each serves as the target for only two vertices. Hence, we suggest utilizing CZZZ gates, drawing inspiration from CZZ gates, which have been demonstrated to be native operations in certain platforms [291]. Specifically, we can leverage these gates to simultaneously target all vertices and edges. The primary challenge lies in ensuring a two-step selection of CZZZ gates, such that each vertex directs the target ZZZ gates to all adjacent edges. This can be effectively achieved using the following grid pattern in Figure 19 and afterward this using the following pattern in the second round the pattern in Figure 20.

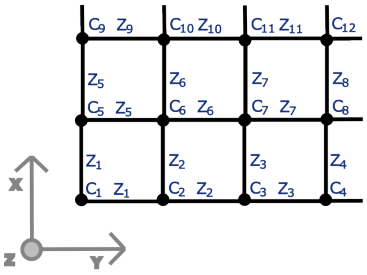


Figure 19: First horizontal entangling routine

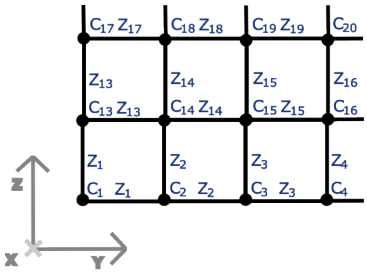


Figure 20: Second vertical entangling routine

These two sequential entangling layers effectively result in a depth-2 process that generates a larger resource state than desired. Now, one only has to remove some of the utilized gates to eliminate paths

in the previously described resource state and, with that, obtaining the desired resource state, utilizing practically feasible and finite fan-in quantum gates.

Step 3a. This step resorts to simply measuring the Hadamard basis of all the edge qubits. This is known to create on the vertex qubits an LU equivalent to the GHZ [292]. These local unitaries being all X gates do generate the state previously named poor-mans cat state, with the following description

$$|\Psi_{3a}\rangle = \frac{|v, e_1, \ldots, e_{3n}\rangle + |\overline{v}, e_1, \ldots, e_{3n}\rangle}{\sqrt{2}}, \text{ with } v \in \mathbb{F}_2^n. \tag{4.6.16}$$

while the measurement outcomes from the edges e_i do define the relations between the bits of the string v as follows,

$$v_i = v_j \oplus_{l \in path(v_i, v_j)} e_l. \tag{4.6.17}$$

Thus, by considering $v_1 = 0$ one can compute the entire string v by the previous relations, which would derive that to obtain effectively a GHZ state one requires to apply a X to all the vertexes with value 1.

Step 3b. Although the current step does occur simultaneously with the previously described, we can analyze it as independent, given the nature of the measurement operations. Also, we will decompose the represented controlled measurement as an input-dependent controlled S^{x_i} gate based on the entire input strings x, and a set of measurements in the Hadamard computational basis.
With that, we recover that before the measurements we have the state,

$$|\Psi_{3b}\rangle = \frac{i^{\langle v,x\rangle}\,|v, e_1, \ldots, e_{3n}\rangle + i^{\langle \overline{v},x\rangle}\,|\overline{v}, e_1, \ldots, e_{3n}\rangle}{\sqrt{2}} \tag{4.6.18}$$

$$= \frac{|v, e_1, \ldots, e_{3n}\rangle + i^{\langle v,x\rangle + |x|/2}\,|\overline{v}, e_1, \ldots, e_{3n}\rangle}{\sqrt{2}}. \tag{4.6.19}$$

Afterward, depending on the value of $\langle v, x\rangle + |x|/2$ having parity either 0 or 1 does translate into a measurement of a superposition of even or odd strings respectively. Therefore, the parity of the string v directly relates to the parity of $\langle v, x\rangle + |x|/2$.

Remark. Before proceeding to the final step, it's important to note that although Steps 3a and 3b were initially considered to be sequential, they take place simultaneously. The realization that analyzing these steps as sequential yet observing that their parallel execution leads to the same outcome stems from the principle that the sequence of single-qubit measurements does not alter the outcome probabilities for any quantum state.

Step 4. This last NC^0 reduction does the post-processing that guarantees that the outcome strings are the result of the $\mathcal{R}_2^m$ problem. For that we will consider that one can describe $\langle v, x\rangle$ as follows,

$$\langle v, x\rangle = \bigoplus_{i=1}^{n} \mathsf{AND}(x_i, v_i) = \bigoplus_{i=1}^{n} \mathsf{AND}\left(x_i, v_c \oplus \bigoplus_{l \in path(v_i, v_c)} e_l\right) \tag{4.6.20}$$

$$= \bigoplus_{i=1}^{n} \bigoplus_{l \in path(v_i, v_c)} \mathsf{AND}(x_i, e_l), \tag{4.6.21}$$

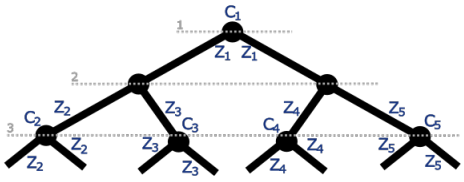
with v_c being a central vertex that we use as reference.
Now all the AND terms of this expression are computed in the routine, producing a single processor of depth 1. The final string provided does contain the result of all these computations and the outcome of the string v from the procedure described in stage $3b$. This guarantees, as the parity of this string is equal to $\langle x, v\rangle$, that the outcome string has a parity equal to $|x|/2$. This is exactly the correct outcome for $\mathcal{R}_2^n$,

as for even input strings $|x|/2 = \mathsf{LSB}(x)$.
Finally, we have only to consider the size of the outcome string. This resorts to account for the size of the paths considered in Equation (4.6.21). These do depend on the distance of any vertex in the lattice to a central vertex in the 3D lattice. One can see that even for a cubic disposition this distance is upper bounded by $n^{1/3}$, and given that the number of paths is at most n we obtain that the number of these terms is upper bounded by $\mathcal{O}(n^{4/3})$, as is the final outcome string. □

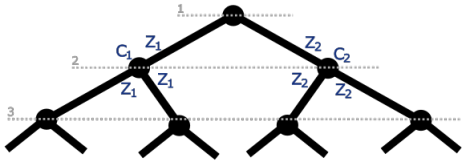
Now, we do additionally consider the use of QNC^0 circuit with arbitrary all-to-all connectivity which will guarantee tighter bounds.

Corollary 5. *There exists a* QNC^0 *circuit with, all-to-all connectivity, of depth 4 and polynomial size that does solve the* $\mathcal{R}_2^m$ *exactly with* $m = \mathcal{O}(n \log n)$.

Proof. The circuit for this new connectivity does not change in any manner from the circuit of figure Figure 18, except for the two CZZZ routines. The graph state intended to be created will follow that of a binary decision tree. Again, taking advantage of CZZZ gates and considering that each edge of the decision tree relates to only two vertices, we can achieve a depth of 2 in the circuit creating the resource state. In particular, this entanglement structure requires applying the CZZZ gate to all the vertices at even levels of the tree and their respective edges, and then in the second pass, performing the same procedure on the odd levels, as will be illustrated subsequently.



(a) CZZZ Routine A



(b) CZZZ Routine B

Afterward, every other part of the circuit remains the same. Consequently, the length of the final outcome string is solely determined by the distance from any vertex to the root node vertex of the tree, as we designate this as the central vertex. From this, we deduce that the maximum distance is at most $\log n$, which means the length of the outcome string is upper bounded by $\mathcal{O}(n \log n)$. □

This establishes two QNC^0 circuits that solve the specified relational problem. We will now examine the lower bounds for $\mathsf{bPTF}^0[k]$ derived in the preceding section alongside this quantum upper bound and present new separations between these circuit classes. To do this, we will consider the minimum value of k, which is 0, and demonstrate a new exact-case separation between QNC^0 and AC^0, with the tightest bound yet established.
Also, in this context, we consider a different scenario where the last reduction is performed by an AC^0 or a $\mathsf{bPTF}^0[k]$ circuit, and we prove a quantum advantage for the problem being solved at step $3b$ of Figure 18 with a contrapositive argument against these circuit classes using a ISMR problem. This allows us to obtain even tighter quantum-classical separations. Furthermore, we find that such a separation is also feasible with the highest value of k allowed by our classical lower bounds, indicating the existence of a problem within QNC^0 that requires superpolynomial-sized $\mathsf{bPTF}^0[k]$ circuits, where $k = n^{1/(5d)}$. It is important to note that the earlier class specifies a circuit class that is strictly larger than AC^0.

Proof of Theorem 4.6.1. The determined expressions do follow from the fact that we have defined the values of m for which there exists a QNC^0 solving the $\mathcal{R}_2^m$ problem exactly in Lemma 36 and Corollary 5. This allows, us to use the exact lower bound for the size s of $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuits of depth d derived

in Lemma 35 with these values of m and the values selected by us for k. For instance, by considering the outcomes of the 3D QNC^0 circuit with $m = n^{4/3}$, we obtain that the size of an AC^0 circuit is larger or equal to

$$s \geq 2^{\left(\frac{n^2\left(n^{4/3}\right)^{-(1+2/q)}}{q^2}\right)^{\frac{1}{d-1}}}. \tag{4.6.22}$$

The expression still has a non-circuit-dependent parameter q that relates only to the random restrictions used to obtain this bound. As we intend to maximize the quantum-classical separation, we would like to obtain the largest value possible for the previous expression. Therefore, we search for the value of q that archives exactly that. For that, we derive the expression in the function of q first,

$$\frac{ds}{dq} = -\frac{2^{1+\left(\frac{n\left(n^{\frac{4}{3}}\right)^{-\frac{2+q}{2q}}}{q}\right)^{\frac{1}{d-1}}}\left(\frac{n\left(n^{\frac{4}{3}}\right)^{-\frac{2+q}{2q}}}{q}\right)^{\frac{1}{d-1}}\left(q-\log\left(n^{\frac{4}{3}}\right)\right)}{(d-1)q^2}. \tag{4.6.23}$$

Subsequently, it requires only determining for which values of the variable of q, for which the expression is equal to 0, to determine the maximums and minimums of the size depending on q. Moreover, this results in the solution $q = \log\left(n^{4/3}\right)$ that we do requires only so substitute in our initial expression, as we know that for small integer values of q and large values close to n the expression is smaller, proving that this value a maximum as desired. Finally, we obtain that the size of an AC^0 solving the $\mathcal{R}_2^m$ problem exactly requires a size no smaller then

$$s \geq \exp\left(\left(\frac{n^{1/3}}{(\log n)^{1+\mathcal{O}(1)}}\right)^{\frac{1}{d-1}}\right). \tag{4.6.24}$$

All the other expressions follow equally with the respective parameters of m and k. □

The final set of separations refers to the average case hardness of solving the same problem. For these, we do focus uniquely on the case where we have the largest value of k, so we drop the case where k is zero as this was already studied previously. This again will refer to the case where we have either 3D or all-to-all connectivity in the corresponding QNC^0 circuit.

Proof of Theorem 4.6.2. Again the expressions do follow from the specific values of m for which there exists a QNC^0 solving the $\mathcal{R}_2^m$ problem (Lemma 36 and Corollary 5). This is now coupled with the average-case lower bound for the efficiency with a bPTF$^0[k]$/rpoly circuits of depth d and size s with maximum parameter k does solve the $\mathcal{R}_2^m$ from Corollary 4 does provide us the enunciates bounds up to a value for q. To determine the value of q, we employ the same methodology used in the proof of Theorem 4.6.1, thereby deriving the optimal values for the entire set of expressions. □

4.6.2 Separation for qupit cases

In this section, we will examine a larger set of ISMR problems, denoted as $\mathcal{R}_p^m$, and the correlation of our candidate circuit with the correct solutions. More concretely, we use these problems to establish distinctions between QNC^0 and bPTF$^0[k]$ circuits. For that, we demonstrate that a QNC^0 circuit can solve these problems with constant and non-zero correlation while generating outcome strings of sub-quadratic size, $m = o(n^2)$. In contrast, we will show that a bPTF$^0[k]$ circuit of polynomial size solves the problem with a correlation asymptotically approaching zero as the input size increases for equal-size

output strings, even for large and nearly optimal parameter k, as formalized in the subsequent theorem.

Theorem 4.6.4. *For all* ISMRP *problems defined for a prime dimension p, denoted by $\mathcal{R}_p^m$, there exists a* QNC^0 *circuit with all-to-all connectivity, considering a uniform distribution $\mathcal{D}_p$ over $\mathbb{F}_2^n$ of strings with a Hamming weight that satisfies $(\sum_{i=1}^n x_i)$ Mod $p = 0$ as the input distribution, that solves the problem for $m = n \cdot \log n^{p-1}$, achieving a correlation of $\frac{p-1}{p^2}$. In contrast, for any $C \in \mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit with fixed depth d, size polynomial s, and parameter $k \leq n^{1/(5d)}$ solves each one of these problems with a correlation that is bounded by,*

$\mathsf{bPTF}^0[k]/\mathsf{rpoly}$	$k = \mathcal{O}(1)$ $(\equiv \mathsf{AC}^0/\mathsf{rpoly})$	$k = n^{1/(5d)}$
$\mathsf{Corr}_{\mathcal{D}_p}\left(C, \mathcal{R}_p^m\right)$	$\exp\left(-\Omega\left(\frac{n^{1-\mathcal{O}(1)}}{(\log n)^{p-1}(\log s)^{2d-2}}\right)\right)$	$\exp\left(-\Omega\left(\frac{n^{3/5-\mathcal{O}(1)}}{(\log n)^{p-1}(\log s)^{2d-2}}\right)\right)$

.

The previous theorem follows in three steps, delineated in Section 4.6.2, Section 4.6.2, and Section 4.6.2. In the first subsection, we introduce a family of non-local XOR games related to the ISMR problems and the respective mapping between bits and dits, which will later facilitate the translation between these two objects. Then, in Lemma 38, we establish upper bounds on the efficacy of any classical strategy in solving this family of non-local XOR games.

In the second subsection, we establish the classical lower bound for $\mathsf{bPTF}^0[k]$ circuits that address the ISMR problems. We integrate the upper bounds determined for the non-local XOR games with light cone arguments, enabling us to constrain the maximum efficiency of NC^0 circuits operating on bits in addressing the ISMR problems. Additionally, we revisit our random restriction techniques to simplify $\mathsf{bPTF}^0[k]$ circuits into forest lines using Lemma 31, thereby connecting with the NC^0 upper bounds. This approach allows us to define upper bounds for $\mathsf{bPTF}^0[k]$ circuits concerning the correlation with which they resolve the problems, which equivalently translates to size lower bounds for $\mathsf{bPTF}^0[k]$ circuits that achieve a constant positive correlation with the ISMR problems.

Finally, in Section 4.6.2, we define the quantum resource state that will be utilized for the quantum solutions. This involves a generalization of the poor man's cat state, and we also demonstrate that these states can be prepared using qupit QNC^0 circuits as shown in Lemma 59. Subsequently, in Lemma 44, we demonstrate the existence of QNC^0 circuits that address the aforementioned problems with a constant positive correlation. The synthesis of the classical lower bounds and the quantum upper bounds culminates in our previous theorem. Additionally, it is worth noting that Theorem 4.6.5 serves as an intermediate result, delineating separations between NC^0 and qupit QNC^0 circuits for specific problem instances.

Simultaneously with the previous result, we establish a separation for the qutrit case that parallels the qubit case, with simple success probabilities of being correct or incorrect. This equivalence arises from the correlation measure for the qutrit case, which assesses deviations from the correct outcome by magnitudes of 0 or 1, mirroring the correlation measure in the qubit case that is translatable to the commonly used success probability. Furthermore, this problem is particularly intriguing because, while the quantum solution can address the bounded error probabilistic version of the $\mathcal{R}_p^m$—solving all inputs with a probability distinctly greater than $1/2$—the average success probability of a polynomial-size $\mathsf{bPTF}^0[k]$ circuit decreases asymptotically to $1/3$ as the input size increases.

Corollary 6. *For the* ISMR *problems defined for a prime dimension $p = 3$, denoted by $\mathcal{R}_3^m$, there exists a* QNC^0 *circuit with 3D connectivity, considering a uniform distribution over $\mathbb{F}_2^n$ of strings with a Hamming weight that satisfies $(\sum_{i=1}^n x_i)$ Mod $3 = 0$ as the input distribution, that solves the problem for $m = n^{4/3}$, achieving a sucess probability strictly larger then $\frac{1}{2}$. In contrast, for any $C \in \mathsf{bPTF}^0[k]/\mathsf{rpoly}$*

circuit with fixed depth d, size polynomial s, and parameter $k \leq n^{1/(5d)}$ solves each one of these problems with a success probability that is bounded by,

$\mathsf{bPTF}^0[k]/\mathsf{rpoly}$	$k = \mathcal{O}(1)$ $(\equiv \mathsf{AC}^0/\mathsf{rpoly})$	$k = n^{1/(5d)}$
$\Pr[\textit{Success}]$	$\frac{1}{3} + \exp\left(-\Omega\left(\frac{n^{2/3-\mathcal{O}(1)}}{\log(s)^{2d-1}}\right)\right)$	$\frac{1}{3} + \exp\left(-\Omega\left(\frac{n^{4/15-\mathcal{O}(1)}}{\log(s)^{2d-1}}\right)\right)$

.

This corollary follows the same proof techniques as Theorem 4.6.4, with the difference being that the correlation measure can always be directly translated into the respective success probabilities.

Quantum non-local games in higher dimensions

We will now define the precise family of XOR non-local games, enabling us to connect the correlation upper bounds of these games to the upper bounds of the ISMR problems for the previously described separations.
Recall that to define a family of non-local games of a particular type we need to describe only the function that maps the modular remainder of the responses by each player and the messages sent to the parties playing the game. In our specific game, the function will also be a modular remainder of the message sent to the parties. Additionally, we extend the XOR non-local games defined in Definition 31 for the correlation measure presented in Section 4.4.3, resulting in the following family of non-local games.

Definition 39 (Modular XOR games)**.** *The class $\mathcal{G}_p$ of multiparty non-local* XOR *games involves n parties, with each party P_i receiving information $x_i \in \mathbb{F}_p$. After each party receives its input, without further communication, each party must provide an individual response $y_i \in \mathbb{F}_p$, contributing to a collective output which is the concatenation $y = (y_1, y_2, \ldots, y_n)$ of their outputs. This defines the classical strategy $w_{\mathcal{G}}$. The correlation with the game is determined by the following expression,*

$$\mathsf{Corr}(w_{\mathcal{G}_p}, \mathcal{G}_p) = \mathop{\mathbb{E}}_{(x_1,x_2,\ldots,x_n)\sim\mathcal{D}} \left[\mathsf{Re}\left(e^{-i\frac{2\pi(\sum_{i=1}^n y_i - \sum_{i=1}^n x_i/p)}{p}}\right)\right],$$

for an arbitrary given input distribution $\mathcal{D}$.

These games are naturally related to the ISMR problems (Definition 38) because the messages to the parties and the responses follow a similar mapping between the input and output for the inverted strict modular relation problems. However, there is a technical misalignment since the non-local games have messages to the parties in $\mathbb{F}_p$, while the ISMR problems involve binary strings. Furthermore, we require uniform input distributions over $\mathbb{F}_2^n$ to achieve average-case hardness separation for the ISMR problems. This will be resolved by establishing correlation upper bounds for the Modular XOR non-local games with biased input distributions, allowing us to relate uniform binary input distributions for the ISMR problems to the corresponding biased distributions over $\mathbb{F}_p^n$ for the non-local games.

Encoding. To bridge our techniques from Section 4.3.4 with the higher-dimensional non-local games described in Definition 39, we will consider a specific encoding between binary and p-ary, for prime p. Subsequently, we will demonstrate that a Hamming weight-based encoding, defined as follows, meets all the necessary criteria.

Definition 40 (Hamming encoding)**.** *We define the Hamming encoding as a mapping of the form $H : \mathbb{F}_2^{n(p-1)} \mapsto \mathbb{F}_p^n$, defined as follows*

$$y = \bigotimes_{j=1}^{n} \left(\sum_{i=1}^{p-1} x_{j\cdot p+i}\right) \tag{4.6.25}$$

with $x \in \mathbb{F}_2^{n(p-1)}$ and $y \in \mathbb{F}_p^n$.

The previous encoding does work very well in the binary to the base-p, however, the inverse from the fact that it is not injective provides a nondetermined inverse. As we will consider later, this will not be a problem to handle the ISMR problems. However, the fact that a uniform distribution over $\mathbb{F}_2^{n(p-1)}$ generates non-uniform distributions over $\mathbb{F}_p^n$ has to be handled. A simple demonstration of this effect can be seen with the trit case, where we have

$$00 \mapsto 0; \quad 01 \mapsto 1; \quad 10 \mapsto 1 \text{ and } 11 \mapsto 2. \tag{4.6.26}$$

One does see that the frequency of 1's in the resulting strings will be larger than 0 or 2. We can account for this by defining the functions that map us the precise frequency of a term in $\mathbb{F}_p$ resulting in a binary string with the Hamming encoding, as follows,

$$\Delta_2^p(x) = \frac{\binom{x}{2}}{2^{p-1}}. \tag{4.6.27}$$

This allows us to account for and determine how the winning probabilities of the respective non-local games are altered by this input distribution. Additionally, we will consider the uniform distribution over $\mathbb{F}_p$, which is defined simply as $\Delta_p^p(x) = \frac{1}{p}$. The latter distribution will allow us to prove stronger separations between QNC^0 qupit circuits and NC^0 circuits.

Now that we have created the necessary elements to bridge our techniques, we can start establishing the sequence of correlation upper bounds for the selected family of non-local games. However, as these demonstrations become increasingly complex when dealing with all prime dimensions p, we will first, for a better exposition, demonstrate the case for qutrits. In addition, this case is of particular interest because $p = 3$ represents the first prime dimension where quantum circuits lack exact solutions to the ISMR problems, and the deviation of quantum-classical efficiencies captures the existing advantage.

Ternary correlation bound

In the case of qutrits, we demonstrate that the correlation between any classical strategy, including the optimal classical strategy denoted by $w^\star_{\mathcal{G}_3}$, and the non-local game defined by $\mathcal{G}_3$ decreases exponentially fast as the number of participants increases. Furthermore, we aim to establish a corresponding bound for the following expression,

$$\mathsf{Corr}_{\mathcal{D}}\left(w^\star_{\mathcal{G}_3}, \mathcal{G}_3\right) = \mathop{\mathbb{E}}_{x\sim\mathcal{D}}\left(\mathsf{Re}\left(e^{i\frac{2\pi\left(\left|w^\star_{\mathcal{G}_3}(x)\right| - \sum_{i=1}^n x_i/3\right)}{3}}\right)\right). \tag{4.6.28}$$

Additionally, this correlation measure can in the trinary case be used to compute the probability with which the XOR non-local game can be solved[9],

$$\Pr_{x\sim\mathcal{D}}[w^\star_{\mathcal{G}_3}(x) \in \mathcal{G}_3(x)] = \frac{1 + 2\cdot \mathsf{Corr}_{\mathcal{D}}\left(w^\star_{\mathcal{G}_3}, \mathcal{G}_3\right)}{3}. \tag{4.6.29}$$

Building on the previous objects, we will show that the winning probability of any classical strategy for the non-local game defined by the $\mathcal{G}_3$ problem is bounded below $1/2$ for sufficiently large number of participants n and asymptotically approaches $1/3$ for uniform distributions over messages in $\mathbb{F}_3^n$ and biased distribution over $\mathbb{F}_3^n$ resulting from the translation from a uniform distribution over $\mathbb{F}_2^n$ using the Hamming encoding.

[9] Note that in the case of having a maximal mismatch between the functions we obtain that $\Pr[f(x) = g(x)] = 0$ given that will be equal to $\mathsf{Corr}_{\mathcal{D}}(*,*) = -0.5$. Also, whenever $\Pr[f(x) = g(x)] = 1$ we have that $\mathsf{Corr}_{\mathcal{D}}(*,*) = 1$ which is again consistent, and all the values in between do follow equally.

Lemma 37. *Any local probabilistic classical strategy $w_{\mathcal{G}_3}$ that solves the non-local* XOR *games denoted as $\mathcal{G}_3$, with messages over $\mathbb{F}_3^n$, achieves a winning probability, for uniform input distributions over $\mathbb{F}_3^n$ with Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)$ Mod $3 = 0$ and r restricted dits, bounded by $\frac{1}{3} + \left(\frac{17}{20}\right)^{n-r}$. Similarly, for a uniform distribution over $\mathbb{F}_2^n$, where the Hamming weight satisfies $\left(\sum_{i=1}^n x_i\right)$ Mod $3 = 0$ and the inputs are translated using a bit-to-ternary encoding, the winning probability with r restricted bits is bounded by $\frac{1}{3} + \left(\frac{9}{10}\right)^{\frac{n-r}{2}}$.*

Proof. In this setting, classical strategies allow each participant, identified as i within a set of n parties, to generate an output of either x_i or $x_i + b$, with b in $\mathbb{F}_3$. Therefore, the efficiency of any classical strategy $w_{\mathcal{G}_3}$ in solving the XOR non-local game $\mathcal{G}_3$ depends on the linear function $l_a = \sum_{i=1}^n a_i \cdot x_i$ (with a in $\mathbb{F}_3^n$) resulting from the messages send to the parties and the function $f(x) = (|x|/3)^{-1}$Mod p for all messages where $|x|$Mod $3 = 0$, which defines the non-local game.
With this defined, we would like to compute the maximal value for the correlation and consequently the probability Pr$[f(x) = l_a(x)]$. In particular, we have the following expression for the correlation,

$$\text{Corr}_*\left(w^\star_{\mathcal{G}_3}, \mathcal{G}_3\right) = \max_{b\in\mathbb{F}_3^n}\left(\sum_{\substack{x\in\mathbb{F}_3^n \\ |x| \text{ Mod } 3=0}} \frac{\text{Re}\left(e^{i\frac{2\pi 2(|x|/3)}{3}}\prod_{i=1}^n e^{i\frac{2\pi x_i b_i}{3}}\right)}{\prod_{i=1}^n \Delta_*^p(x_i)}\right). \tag{4.6.30}$$

without yet completely defining a distribution over the delivered messages.
The correctness of the previous expression follows from the fact that $|x|/3$ Mod 3 and $\sum_{i\in\mathbb{F}_3^n} x_i b_i$ Mod 3 are up to sign changes respectively representative of the two functions f and l_a that we intended to compare. Also, we can simply drop the Mod 3 operation on f since this is inherently carried out by the arithmetic associated with the roots of unity under consideration.
With that, we will write the correlation from Equation (4.6.30) as,

$$\sum_{\substack{x\in\mathbb{F}_3^n \\ |x| \text{ Mod } 3=0}} \frac{\text{Re}\left(e^{i\frac{4\pi(|x|/3)}{3}}\prod_{i=1}^n e^{i\frac{2\pi x_i b_i}{3}}\right)}{\prod_{i=1}^n \Delta_*^p(x_i)} = \text{Re}\left(\sum_{\substack{x\in\mathbb{F}_3^n \\ |x| \text{ Mod } 3=0}}\left(\prod_{i=1}^n \frac{e^{i\frac{2\pi x_i(2+3b_i)}{9}}}{\Delta_*^p(x_i)}\right)\right) \tag{4.6.31}$$

$$= \text{Re}\left(\left(\left(\sum_{x_1\in\mathbb{F}_3}\frac{e^{i\frac{2\pi x_1(2+3b_1)}{9}}}{\Delta_*^p(x_i)}\right)\cdots\left(\sum_{x_{n-1}\in\mathbb{F}_3}\frac{e^{i\frac{2\pi x_{n-1}(2+3b_{n-1})}{9}}}{\Delta_*^p(x_i)}\right)\right)\bullet\frac{e^{i\frac{2\pi x_n(2+3b_n)}{9}}}{\Delta_*^p(x_n)}\right) \tag{4.6.32}$$

with $x_n = \left(\sum_{i=1}^{n-1} x_i\right)^{-1}$ and $\bullet$ being a conditional multiplication operation depending on the values of x_1 to x_{n-1} in the first expression, such that the total sum of the inputs fulfills the condition of being congruent with zero modulo 3.
Now that we have reduced the correlation expression to a sum over fixed and equivalent expressions, we can establish bounds for the maximum value of the entire expression to obtain an upper limit. In particular, we will now apply the method proposed for qubits in [293], and use the fact that the last term $e^{i\frac{2\pi x_n(2+3b_{n-1})}{9}}$ takes each value of x_n over $\mathbb{F}_3^n$,

$$\leq \max_{b\in\mathbb{F}_3^n}\left(\text{Re}\left(\left(\sum_{x_1\in\mathbb{F}_3}\frac{e^{i\frac{2\pi x_1(2+3b_1)}{9}}}{\Delta_*^p(x_1)}\right)\cdots\left(\sum_{x_{n-1}\in\mathbb{F}_3}\frac{\cdots}{\Delta_*^p(x_{n-1})}\right)\left(\sum_{x_n\in\mathbb{F}_3}\frac{e^{i\frac{2\pi x_n(2+3b_n)}{9}}}{\Delta_*^p(x_n)\cdot(p-1)}\right)\right)\right) \tag{4.6.33}$$

$$\leq \max_{b\in\mathbb{F}_3^n}\left(\mathsf{Re}\left(\frac{\left(\frac{1}{\Delta_*^p(1)}+\frac{e^{i\frac{2\pi(2+3b_i)}{9}}}{\Delta_*^p(1)}+\frac{e^{i\frac{2\pi(4+6b_i)}{9}}}{\Delta_*^p(2)}\right)^n}{p-1}\right)\right). \tag{4.6.34}$$

Considering the resulting expression and Δ_3^3 fully defining the input distribution, we can show that each term in Equation (4.6.34) has its norm bounded for the three different assignments of b_i by the following values,

$$\left|\frac{1+e^{i\frac{2\pi(2+3b_i)}{9}}+e^{i\frac{2\pi(4+6b_i)}{9}}}{3}\right| = \begin{cases} \frac{1}{\sqrt{3\left(\cos^2\left(\frac{\pi}{18}\right)+(-1+\sin\left(\frac{\pi}{18}\right))^2\right)}} < 0.45,\ b_i=0 \\ \frac{1}{\sqrt{3\left((-1-\cos\left(\frac{\pi}{9}\right))^2+\sin^2\left(\frac{\pi}{9}\right)\right)}} < 0.3,\ b_i=1 \\ \frac{1}{\sqrt{3\left((-1+\cos\left(\frac{2\pi}{9}\right))^2+\sin^2\left(\frac{2\pi}{9}\right)\right)}} < 0.85,\ b_i=2 \end{cases}. \tag{4.6.35}$$

Obtaining with that the bound $\mathsf{Corr}_{\mathcal{D}_3'}\left(w_{\mathcal{G}_3}^\star,\mathcal{G}_3\right)\leq\frac{(0.85)^n}{2}$, with $\mathcal{D}_3'$ being the uniform input distributions over $\mathbb{F}_3^n$ with Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)\ \mathsf{Mod}\ 3=0$. Additionally, the probability $\mathsf{Pr}_{x\sim\mathcal{D}_3'}[f(x)=g(x)]\leq\frac{1}{3}+\left(\frac{17}{20}\right)^n$, which asymptotically approaches $1/3$ and for $n\geq 12$ is already strictly smaller the $1/2$.

Equally, we will consider Δ_2^3 which defines the distribution $\mathcal{D}_3$ generated by considering a uniform distribution of strings over $\mathbb{F}_2^{2n}$ converted to strings in $\mathbb{F}_3^n$ with the Hamming encoding. In particular, in this case, we obtain that for the three different assignments of b_i, the term in Equation (4.6.34) is bounded by

$$\left|\frac{1}{4}+\frac{e^{i\frac{2\pi(2+3b_i)}{9}}}{2}+\frac{e^{i\frac{2\pi(4+6b_i)}{9}}}{4}\right| = \begin{cases} \frac{1}{2}\left(1+\sin\left(\frac{\pi}{18}\right)\right) < 0.59,\ b_i=0 \\ \sin^2\left(\frac{\pi}{18}\right) < 0.04,\ b_i=1 \\ \frac{1}{2}\sqrt{\frac{1}{2}\left(3+4\cos\left(\frac{2\pi}{9}\right)+\sin\left(\frac{\pi}{18}\right)\right)} < 0.89,\ b_i=2 \end{cases}. \tag{4.6.36}$$

In this case, we obtain the following bound, $\mathsf{Corr}_{\mathcal{D}_3}\left(w_{\mathcal{G}_3}^\star,\mathcal{G}_3\right)\leq\frac{(0.89)^n}{2}$. Additionally, the probability $\mathsf{Pr}_{x\sim\mathcal{D}_3}[f(x)=g(x)]\leq\frac{1}{3}+\left(\frac{9}{10}\right)^n$, which asymptotically approaches $1/3$ and for $n\geq 18$ is already smaller then $1/2$.

The effect of restricted dits in Equation (4.6.33) for the respective sum of that variables does make it have a constant normal equal to 1. Therefore, these cases do not contribute to increasing the correlation and the overall expression. Equally, the space of inputs is reduced with the same factor. This implies that restricting dits makes the overall problem equivalent to the non-restricted, non-local game with a smaller input. Also, as there is no structure of the game limiting the rearrangement of messages, we can reorganize the fixed variables such that they form a $r/(p-1)$ dits, and the previous analysis holds equally in this case.

Finally, there is no probabilistic classical strategy that surpasses the two upper bounds determined for the correlation measures by Yao's min-max principle. □

Higher order correlation bounds

Now, we will generalize the upper bound on the efficiency of any classical strategy, including the optimal classical strategy $w_{\mathcal{G}_p}^\star$, for all the XOR non-local games $\mathcal{G}_p$ defined over higher prime dimensions. Therefore, we now intend to constrain the following expression,

$$\mathsf{Corr}_{\mathcal{D}}\left(w_{\mathcal{G}_p}^\star,\mathcal{G}_p\right)=\mathop{\mathbb{E}}_{x\sim\mathcal{D}}\left(\mathsf{Re}\left(e^{i\frac{2\pi\left(\left|w_{\mathcal{G}_p}^\star(x)\right|-\sum_{i=1}^n x_i/p\right)}{p}}\right)\right). \tag{4.6.37}$$

Note that previously, for p equal to 2 and 3, due to the symmetry of the roots of unity, the correlations were still binary discriminators relating to the probability of two functions being exactly equal modulo these prime numbers. However, it is important to note that this measure no longer directly relates to the probability of the outcome being correct in binary discrimination; rather, it is a distance measure between the vectors formed by the outcome space of the two functions. Although a similar measure of distances can be created, it generates non-linear terms for which the lower bounding techniques no longer apply. Nevertheless, we can resort to this linear distance measure natural to higher dimensions and use it to prove a quantum-classical efficiency separation for the relation problems at each prime dimension.

Lemma 38. *Any local probabilistic classical strategy $w_{\mathcal{G}_p}$ that solves one of the* XOR *non-local games $\mathcal{G}_p$, with messages over $\mathbb{F}_p^n$, taken from a uniform distribution $\mathcal{D}_p$ over $\mathbb{F}_2^{n(p-1)}$ with Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)$ Mod $p = 0$ and a bit to base-p encoding, the maximal correlation with r restricted bits is bounded by,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(w^{\star}_{\mathcal{G}_p}, \mathcal{G}_p\right) \leq \mathsf{Re}\left(\left(\frac{2^{1-p} e^{\frac{2\pi i}{p^2}} \left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^{\frac{n-r}{p-1}} \Big/ (p-1)\right) \leq (c_p)^{\frac{n-r}{p-1}} \tag{4.6.38}$$

with $c_p \in (0,1)$. Simultaneously, for uniform input distribution $\mathcal{D}'_p$ over $\mathbb{F}_p^n$ with their Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)$ Mod $p = 0$ and r restricted dits, we have

$$\mathsf{Corr}_{\mathcal{D}'_p}\left(w^{\star}_{\mathcal{G}_p}, \mathcal{G}_p\right) \leq \frac{1}{p^{n-1}(p-1)} \mathsf{Re}\left(-\frac{e^{2\pi i/p^2}\left(-1+e^{2\pi i\frac{-1+p^2}{p}}\right)}{-1+e^{i\frac{2\pi}{p^2}}}\right)^{n-r} \tag{4.6.39}$$

$$\leq \frac{p}{p-1}\left(\mathsf{Re}\left(e^{\frac{2\pi i}{p^2}}\right)\right)^{n-r}. \tag{4.6.40}$$

Proof. To obtain this bound on the correlation, we again consider that classical strategies $w_{\mathcal{G}_p}$ allow each participant, identified as i within a set of n parties, to generate an output of either x_i or $x_i + b$, with b in $\mathbb{F}_p$ as defined by the generalized XOR non-local games. Thus, the capability of approximating the ideal outcome will depend on the distance of the closest linear function $l_a = \sum_{i=1}^n a_i \cdot x_i$ with a in $\mathbb{F}_p^n$, and the function defined by the respective non-local game $\mathcal{G}_p$ over $\mathbb{F}_p$, taking the form of a function of the type $f : \mathbb{F}_p^n \mapsto \mathbb{F}_p$, and specifically defined as $f(x) = (|x|/p)^{-1}$ Mod p.
Our generalized correlation measure will then be defined as follows for arbitrary p and distribution Δ_*^p,

$$\mathsf{Corr}_*\left(w^{\star}_{\mathcal{G}_p}, \mathcal{G}_p\right) = \max_{b\in\mathbb{F}_p^n}\left(\sum_{\substack{x\in\mathbb{F}_p^n \\ |x| \text{ Mod } p=0}} \mathsf{Re}\left(\frac{e^{i\frac{2\pi(p-1)(|x|/p)}{p}}\prod_{i=1}^n e^{i\frac{2\pi x_i b_i}{p}}}{\Delta_*^p(x)}\right)\right). \tag{4.6.41}$$

The previous expression is algebraically equivalent to the one in Equation (4.6.32), and the terms in the exponents are linear, in the sense that all have a degree of 1. Therefore, we can bound the previous expression similarly as follows,

$$\leq \max_{b\in\mathbb{F}_p^n}\left(\mathsf{Re}\left(\sum_{x_1\in\mathbb{F}_p}\left(\frac{e^{i\frac{2\pi x_1((p-1)+pb_1)}{p^2}}}{\Delta_*^p(x_1)}\right)\cdots\sum_{x_{n-1}\in\mathbb{F}_p}\left(\frac{\cdots}{\Delta_*^p(\ldots)}\right)\sum_{x_n\in\mathbb{F}_p}\left(\frac{e^{i\frac{2\pi x_n((p-1)+pb_n)}{p^2}}}{\Delta_*^p(x_n)\cdot(p-1)}\right)\right)\right) \tag{4.6.42}$$

$$\leq \max_{b\in\mathbb{F}_p^n}\left(\mathsf{Re}\left(\left(\sum_{x_j\in\mathbb{F}_p}\frac{e^{i\frac{2\pi x_j((p-1)+pb_j)}{p^2}}}{\Delta_*^p(x_j)}\right)^n \Big/ (p-1)\right)\right). \tag{4.6.43}$$

Unfortunately, at this point, it is no longer feasible to compute all the values for b in $\mathbb{F}_p$ to bind the expression for a general proof, as previously done for the qutrit case. Nevertheless, there is a simple geometric intuition for the values of the expression $\sum_{x_j\in\mathbb{F}_p} e^{i\frac{2\pi x_j((p-1)+pb_j)}{p^2}}$. For any value of b_j selected, the sequential product with the values of x_j moves the angle of each term involved in the sum by a step equivalent to a multiple of $2\pi i/p^2$, equal to $(p-1)+pb_j$. Ideally, the maximum for the expression would be reached with a step equal to zero, so we always have the maximum value of 1 and perfect correlation. Since this case with optimal correlation is not possible, as $(p-1)+pb_j$ is never zero for any value of b_j, we have that the subsequent best option is to have that expression equal to either 1 or p^2-1 to maximize either the real or the absolute values.
In addition, the same analysis holds equally for any of the distributions of the inputs $\Delta_*^p(x)$, as the same arguments apply for the selection of the value b_j. Therefore, by considering $b_j = p-1$ and Δ_p^p, we obtain the distribution $\mathcal{D}_p'$ and define the optimal correlation as follows,

$$\mathsf{Corr}_{\mathcal{D}_p'}\left(w_{\mathcal{G}_p}^\star, \mathcal{G}_p\right) \leq \frac{1}{p^{n-1}(p-1)}\mathsf{Re}\left(-\frac{e^{2\pi i/p^2}\left(-1+e^{2\pi i\frac{-1+p^2}{p}}\right)}{-1+e^{i\frac{2\pi}{p^2}}}\right)^n. \tag{4.6.44}$$

To obtain the exact value, one is only required to obtain a numerical bound for the previous expression depending on p. From an analytical point of view, we know that all the terms in the sum contained in Equation (4.6.43), except for the first term, are smaller than 1, even when we use the best possible value for b_i. Thus, we have that,

$$\mathsf{Corr}_{\mathcal{D}_p'}\left(w_{\mathcal{G}_p}^\star, \mathcal{G}_p\right) \leq \frac{p}{p-1}(c_p)^n, \text{ with } c_p \in (0,1). \tag{4.6.45}$$

Also, as c_p is no larger than $1-\mathsf{Re}\left(e^{\frac{2\pi i}{p^2}}\right)$, we have that it is guaranteed that it is infinitely close to 1. Thus, the correlation of any classical strategy goes exponentially fast close to zero with increasing input size.
Now we need to consider the case where we have $\mathcal{D}_p$ the uniform distribution over $\mathbb{F}_2$, whose distribution over $\mathbb{F}_p$ we capture with Δ_2^p. This case then defines the following bound on the correlation,

$$\mathsf{Corr}_{\mathcal{D}_p}\left(w_{\mathcal{G}_p}^\star, \mathcal{G}_p\right) \leq \mathsf{Re}\left(\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^n \Big/ (p-1)\right). \tag{4.6.46}$$

The previous expression should asymptotically approach zero with increasing input size n for arbitrary p. To verify that this holds, we need to show that the expression

$$\varkappa_p = \left|\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right|, \tag{4.6.47}$$

is smaller than 1 for arbitrary p. We will prove this inductively, informed by the limit of the expression $\varkappa_p$

with an increasing value of p being 1, given that

$$\lim_{p\to\infty}\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)=1. \tag{4.6.48}$$

Combining this limit with the fact that for $p=3$, the expression $\varkappa_3$ is smaller than 1, it is enough to prove that the same expression $\varkappa_p$ reaches from below its asymptotic bound. More precisely, we need to prove that this expression is monotonically increasing. For that, we only need to prove that for two values k and $k+1$, the following expression is negative,

$$\varkappa_k-\varkappa_{k+1}\propto\left(\frac{2e^{\frac{2\pi i}{k^2}}\left(1+e^{-\frac{2\pi i}{k^2}}\right)^k}{1+e^{\frac{2\pi i}{k^2}}}-\left(1+e^{-\frac{2\pi i}{(k+1)^2}}\right)^k\right). \tag{4.6.49}$$

With respect to this expression, we can show that the real part of $\left(1+e^{-\frac{2\pi i}{(k+1)^2}}\right)^k$ is larger than $\left(1+e^{-\frac{2\pi i}{k^2}}\right)^k$. For that, we consider the norms of these values without the exponent k, and as for the first expression, the corresponding angle $\theta_1=\frac{-2\pi i}{(k+1)^2}$ is larger than $\theta_2=\frac{-2\pi i}{k^2}$, the angle of the second expression, while both being in the fourth quadrant. Hence, the norm of the first term r_1 is larger than that of the second term r_2. Then, with the exponent k, we look into their respective polar representations and observe that the respective real parts of these are equal to $(r_1)^k\cos(k\theta_1)$ and $(r_2)^k\cos(k\theta_2)$. Now, as mentioned previously, r_1 is larger than r_2, and also both the angles $k\theta_1$ and $k\theta_2$ are in the fourth quadrant, with $k\theta_1$ being larger. Thus, having a larger value of $\cos(k\theta_1)$ than $\cos(k\theta_2)$, obtaining our initial assertion.
The final step is simply to consider that the real part of $\left(1+e^{\frac{-2\pi i}{k^2}}\right)^k$ does not get larger than $\left(1+e^{\frac{-2\pi i}{(k+1)^2}}\right)^k$ with the multiplication with $\frac{2e^{\frac{2\pi i}{k^2}}}{1+e^{\frac{2\pi i}{k^2}}}$. In particular, this value can be simplified to $1+i\tan\left(\frac{\pi}{k^2}\right)$, which, when multiplied with the previous value, makes the real part of $(1+e^{\frac{-2\pi i}{k^2}})^k$ equal to $(r_2)^k\cos(k\theta_2)(1-\tan\left(\frac{\pi}{k^2}\right))$, which is bounded by $2(r_2)^k\cos(k\theta_2)$ for $k\geq 2$. Also, the absolute value of the first term is smaller than the second for $k\geq 2$. Thus, for $k\geq 2$, this expression in Equation (4.6.49) is negative as we intended. Thus, we obtain that the value of $\mathsf{Corr}_{\mathcal{D}_p}\left(w^\star_{\mathcal{G}_p},\mathcal{G}_p\right)$ with the distributed messages being sampled from a uniform distribution over $\mathbb{F}_2^{n(p-1)}$ and encoded with the Hamming encoding to strings in $\mathbb{F}_p^n$ is bounded by $(c_p)^n$, with c_p being a constant strictly smaller than 1.
Finally, the effect of restricting bits or dits simply results in a version of the same XOR non-local game with fewer variables, analogous to the case in Lemma 37. Additionally, no probabilistic classical strategy can surpass the two upper bounds established for the correlation measures according to Yao's min-max principle. □

Average-case $\mathsf{bPTF}^0[k]$ **lower hardness**

Here, we will demonstrate that the classical circuit class $\mathsf{bPTF}^0[k]$ cannot solve the ISMR problems with a sufficiently significant correlation. In other words, exponentially large circuits will be required to achieve constant positive correlations with these problems. To accomplish this, we will utilize the classical correlations identified for the Modular XOR non-local games, along with the novel tools developed in Section 4.3.4.

NC^0 lower bounds

In the first step, we will use the previously determined correlation bounds for the XOR non-local games and lightcone arguments to demonstrate lower bounds on the efficiency of NC^0 circuits solving the class of ISMR problems. This can be done given that through the Hamming encoding, we can consider each set of $p-1$ bits as a representation of a base-p value in the non-local game for both the messages to the parties and their responses. Consequently, any valid solution to the $\mathcal{R}_p^{m(p-1)}$ problem would equally represent a valid solution to the respective XOR non-local game played over messages of the type $\mathbb{F}_p^n$, and vice versa.
In addition to the previous connection, we will also show that NC^0 circuits exhibit bounded locality in terms of their causal influence. This implies that each outcome depends on a fixed number of inputs that can never exceed n. More precisely, through some combinatorial arguments, it is possible to demonstrate that producing a correct outcome string with these circuits is equivalent to solving the aforementioned non-local games on a subset of the inputs with the same capacity as any of the optimal classical strategies considered before. This will allow us to show that unless the circuits have sufficient depth, they will fail with the same probability as the best classical strategies for the specific non-local games in question determined in Lemma 37 and Lemma 38, and this directly defines the efficiency with which they solve the ISMR problems on binary inputs and outputs.

Lemma 39. *Let C be a circuit with $n(p-1)$ binary inputs, arranged in n blocks of size $p-1$, and $m(p-1)$ binary outputs, arranged in m blocks of size $p-1$, with locality l for each outcome bit. Then, for any arrangement of the blocks, there exists a subset of input blocks S of size $\Omega(\min(n, \frac{n^2}{l^2 m\cdot(p-1)^4}))$ such that each output block depends on at most one block from S.*

Proof. To prove this we first consider that each of the output blocks does result from at most $(p-1)\cdot f_{in}$ bits from the previous layer, with gates of fan-in equal to f_{in}. For the subsequent layers, each bit depends on at most f_{in} elements of the previous layer, except for the last layer. Here as there are $p-1$ input bits per block, each element can depend on $(p-1)\cdot f_{in}$ input blocks. This means that each output block has a locality equal to $l' = (p-1)^2 \cdot f_{in}^d = (p-1)^2 \cdot l$, respective to the input blocks.
At this point, we can turn to the intersection graph that includes all the output blocks as vertices and connects vertices with an edge if they share an input block. This graph will have m vertices. Given that each output block can depend on $(p-1)^2 \cdot l$ input blocks, and each of these inputs can influence up to $(p-1)^2 \cdot l$ output blocks, it follows that each vertex is connected to at most $(p-1)^4 \cdot l^2$ other vertices. Therefore, this graph will have at most $\mathcal{O}(m \cdot (p-1)^4 \cdot l^2)$ edges.
The final step involves applying Turán's theorem, which provides an upper bound for the number of edges E in a graph G to avoid having with certainty a complete subgraph K_S, given by the inequality $\left(1-\frac{1}{S}\right)\frac{n^2}{2} \leq E$. Our goal, however, is to ascertain the minimum size of the independent set. To this end, we consider the complement graph $\overline{G}$, where an edge exists in G wherever it was absent in $\overline{G}$, and vice versa. This ensures that if the original graph contained a complete subgraph of size S, then the complement graph $\overline{G}$ will contain an independent set of the same size. Consequently, we can establish a lower bound for the size of this independent set as $\left(1-\frac{1}{S}\right)\frac{n^2}{2} \geq \binom{n}{2} - E$. By simplifying the previous expression, we find that $S \geq \frac{n^2}{E}$. Substituting in the number of edges we have, we obtain $S \geq \frac{n^2}{m(p-1)^4 l^2}$. □

Subsequently, by integrating the previous lemma with the upper bounds for the correlations of classical strategies with the Modular XOR non-local games, we demonstrate that the extent to which an NC^0 circuit successfully solves the ISMR problems is directly influenced by the circuit's locality, its depth, and the length of the desired output string.
In particular, we will determine these upper bounds on the correlation for uniform input distributions over $\mathbb{F}_2^n$ and, subsequently, for uniform distributions over $\mathbb{F}_p^n$.

Lemma 40. *Let C be a circuit in* $\mathsf{NC}^0/\mathsf{rpoly}$ *with locality l. Then, C fails to solve each of the $\mathcal{R}_p^m$ problems, with a uniform distribution $\mathcal{D}_p$ over $\mathbb{F}_2^n$ and Hamming weight satisfying $\left(\sum_{i=1}^n x_i\right)$ Mod $p = 0$, with a correlation larger than,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(C, \mathcal{R}_p^m\right) = \mathcal{O}\left(\mathsf{Re}\left(\frac{2^{1-p} e^{\frac{2\pi i}{p^2}} \left(e^{-\frac{2\pi i}{p^2}}\left(1 + e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1 + e^{\frac{2\pi i}{p^2}}}\right)^{min\left(n, \frac{n^2}{m(p-1)^4 l^2}\right)}\right). \tag{4.6.50}$$

In particular, for $p = 3$ we can bind the success probability by

$$\Pr_{\substack{x \sim \mathbb{F}_2^n \\ |x| \text{ Mod } 3=0}} [C(x) \in R_3^m(x)] \leq \frac{1}{3} + \left(\frac{9}{10}\right)^{min\left(n, \frac{n^2}{16ml^2}\right)}. \tag{4.6.51}$$

Proof. This follows directly from the combination of Lemma 39 with Lemma 37 and Lemma 38, provided that one considers all the blocks in the independent set S derived by Lemma 39 as outcome bits producing an outcome equivalent to linear functions in $\mathbb{F}_p$ over a set of input bits of equal size. Therefore, for this input distribution which translates as a uniform distribution over $\mathbb{F}_2^n$ to a biased distribution over $\mathbb{F}_p^n$ with the respective bias determined by Δ_2^p, the efficiency cannot exceed the respective probability or correlation determined in Lemma 37 and Lemma 38. Thereby establishing a lower bound on the efficiency of an NC^0 circuit based on its locality and size.
In addition, we can consider that a random advice string allows for the probabilistic mixture of NC^0 circuits. However, since all these circuits can be reduced to the efficiency of a single classical strategy for the ISMRP at hand, and since there is no probabilistic mixture with a larger correlation with the optimal outcomes, the same applies to $\mathsf{NC}^0/\mathsf{rpoly}$ circuits. □

Considering a biased binary input distribution, further reducing the respective classical correlation bound for NC^0 circuits is possible.

Corollary 7. *Let C be a circuit in* $\mathsf{NC}^0/\mathsf{rpoly}$ *with locality l. Then, C fails to solve each of the $\mathcal{R}_p^m$ problems with a biased binary input distribution $\mathcal{D}'_p$, which is defined by a uniform distribution over $\mathbb{F}_p^n$ that is converted into a binary distribution using one of the potential inverse mappings of the Hamming encoding, with an average correlation larger than,*

$$\mathsf{Corr}_{\mathcal{D}'_p}\left(C, \mathcal{R}_p^n\right) = \mathop{\mathbb{E}}_{\substack{x \sim H^{-1}(\mathbb{F}_p^{n/\log(p)}) \\ |x| \text{ Mod } p=0}} \left(\mathsf{Re}\left(e^{i\frac{2\pi\left(|C(x)| - \left|\mathcal{R}_p^n(x)\right|\right)}{p}}\right)\right) \tag{4.6.52}$$

$$\leq \frac{p}{p-1}\left(\mathsf{Re}\left(e^{\frac{2\pi i}{p^2}}\right)\right)^{min\left(n, \frac{n^2}{m(p-1)^4 l^2}\right)}. \tag{4.6.53}$$

In particular, for $p = 3$ we can with the same setting bind the success probability by

$$\Pr_{\substack{x \sim H^{-1}(\mathbb{F}_3^{n/\log(3)}) \\ |x| \text{ Mod } 3=0}} [C(x) \in R_3^m(x)] \leq \frac{1}{3} + \left(\frac{17}{20}\right)^{min\left(n, \frac{n^2}{16ml^2}\right)}. \tag{4.6.54}$$

Proof. The corollary follows exactly the same steps as the proof of Lemma 40 with the respective use of the bias consideration of Δ_p^p in Lemma 37 and Lemma 38. □

$\mathsf{bPTF}^0[k]$ **lower bounds**

The last lower bound to be determined for the higher-dimensional analysis pertains to the $\mathsf{bPTF}^0[k]$ circuit. This will follow, once again, from the use of our switching lemma developed in Lemma 31, in combination with the previous lower bounds for NC^0 circuits to which the initial circuit is reduced. Here again one of the main points for consideration is the goal to determine super-polylogarithmic values of k, for which no polynomial-size $\mathsf{bPTF}^0[k]$ circuit solves any of the $\mathcal{R}_p^m$ problems with fixed and constant correlation value.

Lemma 41. *For sufficiently large n and $q \in \mathbb{N}_{>0}$, any $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ circuit of depth $d \geq 4$, size $s \leq \exp\left(n^{1/(2d-2)}\right)$ and parameter $k \leq n^{1/(5d)}$ solves the $\mathcal{R}_p^m$ with correlation bounded by,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(|C|, |\mathcal{R}_p^m|\right) = \mathcal{O}\left(\mathsf{Re}\left(\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^{\frac{n^2}{2^{2q}m^{1+2/q}\mathcal{O}(\log(s)^{2d-1})k^{2d}}}\right)\right), \tag{4.6.55}$$

with $\mathcal{D}_p$ being the uniform distribution over string in $\mathbb{F}_2^n$ that satisfy $|x| \bmod p = 0$.

Proof. Beginning with a $\mathsf{bPTF}^0[k]$ circuit, we intend to reduce this object, enabling the application of appropriate lemmas to the problem at hand. The initial segment of our proof parallels the approach in Lemma 35, initiating with a random restriction denoted as ρ. The probability of this restriction is set to $p = \frac{1}{m^{1/q}} \cdot \mathcal{O}(\log(s)^{d-1} \cdot k^d)$, where t is chosen to be $\frac{pn}{8}$. Subsequently, a second random restriction, τ, is applied to the $2t$ variables within the global decision tree, which, after the initial random restriction, is reduced to $\mathsf{DT}(q-1)^m$. This second restriction occurs with a probability of $1 - \exp(-\Omega(pn))$.
Thus, using p-random restrictions we obtained that sampling $|x| \bmod p$ strings consistent with the restrictions does provide us with a uniform distribution over the $|x| \bmod p$ strings, given that over any non-total subset of the input bits these are uniform distributions in $\mathbb{F}_2$. Exactly, the same argument does work over the restriction τ given that the variables of the global decision tree are selected randomly in $\{0,1\}^{2t}$. This property of p-random restrictions, combined with the fact that we can reconstitute the surviving variables into a new string for which we need to solve the initial problem, allows us to apply Lemma 40 directly without any limitations on the specific locations of the variables that remain alive. Combining this with the fact that with a high probability more $pn/4$ variables are alive, and the circuit does have to evaluate the $\mathcal{R}_p^m$ problem with local decision trees of depth $q-1$, we obtain that these depend at most on 2^q of the variables alive. Consequently, we have that any $\mathsf{bPTF}^0[k]$ circuit does compute the ISMR problems with a correlation bounded by,

$$\begin{aligned}\mathsf{Corr}_{\mathcal{D}_p}\left(|C|, |\mathcal{R}_p^m|\right) &\leq \mathsf{Re}\left(\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^{\Omega\left(\frac{(pn)^2}{2^{2q}m}\right)} \Big/ (p-1)\right) \\ &= \mathcal{O}\left(\mathsf{Re}\left(\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^{\frac{n^2}{2^{2q}m^{1+2/q}\mathcal{O}(\log(s)^{2d-1})k^{2d}}}\right)\right),\end{aligned}$$

with $\mathcal{D}_p$ being the uniform distribution over string in $\mathbb{F}_2^n$ that satisfy $|x| \bmod p = 0$.
Finally, the same proof does hold for the case where the $\mathsf{bPTF}^0[k]$ circuit is given a random string as advice. Since any $\mathsf{bPTF}^0[k]$ reduces to the same object after the use of random restriction, and that

from Lemma 40 we do now that any probabilistic strategy defined by the random advice does fail, so do bPTF$^0[k]$/rpoly circuits. □

This finishes determining the classical upper bounds for the correlations achieved by the bPTF$^0[k]$ circuit for the entire set of ISMR problems. Finally, we will also enunciate a bound for the maximum probability with which the same circuit class can solve the specific case of $\mathcal{R}_3^m$, as in this case, it follows directly from the correlation bound.

Corollary 8. *For sufficiently large n and $q \in \mathbb{N}_{>0}$, any* bPTF$^0[k]$/rpoly *circuit of depth $d \geq 4$, size $s \leq \exp\left(n^{1/(2d-2)}\right)$ and parameter $k \leq n^{1/(5d)}$ solves the $\mathcal{R}_3^m$ with a probability bounded by,*

$$\Pr_{\substack{x\sim\mathbb{F}_2^n \\ |x| \text{ Mod } 3=0}} [C(x) \in R_3^m(x)] \leq \frac{1}{3} + \left(\frac{9}{10}\right)^{\Omega\left(\frac{n^2}{2^{2q}m^{1+2/q}\mathcal{O}(\log(s)^{2d-1})k^{2d}}\right)}. \tag{4.6.56}$$

Proof. The proof follows in the same manner as in Lemma 41 by executing all the steps up to the use of the lower bound for the success probability with NC^0/rpoly circuits. In particular, at this point, we consider the upper bound determined for the success probability in the specific case of $\mathcal{R}_3^m$ as outlined in Lemma 40. □

Qupit upper bound

In this section, we will describe QNC^0 circuits over qupits to solve probabilistically the ISMR problems. This will use the qupit generalizations of poor man's cat states from [96]. However, we apply some slight modifications such that it simplifies the following set of proofs as it also contributes to the self-containment of the document.

Definition 41 (Generalized poor-man's qudit state)**.** *Let p be a prime and $z^{+0} \in \mathbb{F}_p^n$. We define a generalized poor-mans qudit state $\left|\text{GPM}_p^n\right\rangle = \frac{1}{\sqrt{p^n}} \sum_{i\in\mathbb{F}_p} |z^{+i}\rangle$ with z^{+i} being the string that takes a initial defined string z^{+0} and applies the bitwise sum of i to each bit of the string obtaining with that,*

$$z^{+i} = \big((z_1 + i) \text{ Mod } p\big)||\big((z_2 + i) \text{ Mod } p\big)|| \ldots ||\big((z_n + i) \text{ Mod } p\big). \tag{4.6.57}$$

Now it will be important to prove these types of states can be created with a QNC^0 circuit.

Lemma 42. *For $n \in \mathbb{N}$, $\left|\text{GPM}_p^n\right\rangle$ can be prepared with* QNC^0 *circuit based on a fully connected and without cycles graph $G = (V, E)$ over the respective qudit prime dimension p. The respective string v defining the precise state is described by the following recurrence relation between vertex qudits and the measurement outcomes from the edge qudits,*

$$v_u = \left(\sum_{e_{i,j}\in Path(u,w)} (-1)^{|Path(u,j)|+|Path(u,w)|}\, e_{i,j} + v_w\right) \text{ Mod } p \tag{4.6.58}$$

for any two vertices $u, w \in V$, and the shortest path $Path(u, w)$ from u to w in the graph G.

Proof. We will provide proof by describing concretely the circuit that does create the $\left|\text{GPM}_p^n\right\rangle$ states. In particular, we will describe how the various stages of the circuit illustrated in Figure 22 compute the correct state.

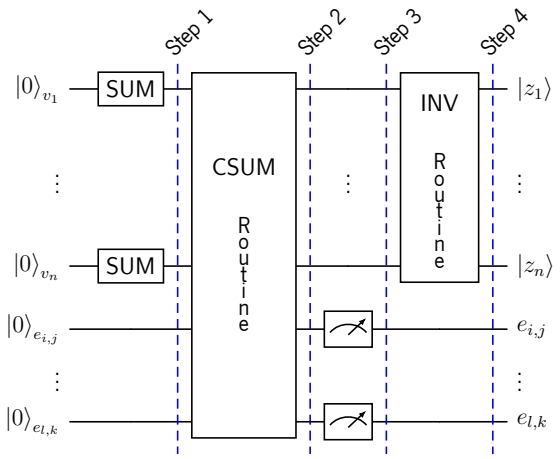


Figure 22: Illustration of a constant-depth quantum circuit for generating a generalized poor man's cat state, accompanied by the classical information needed to describe the specific instance created.

Step 1. The circuit does start with a qudit of prime dimension p for each vertex and edge in the graph $G = (V, E)$. Based on this labeling, we apply a Fourier gate or higher dimensional Hadamard gate F to each vertex qudit. This does create the state $|\Psi_1\rangle = \frac{1}{\sqrt{p^n}} \sum_{x \in \mathbb{F}_n^d} |x\rangle |0\rangle^{\otimes l}$.

Step 2. Then, the CSUM routine applies a CSUM gate between every edge e_i and vertex v_i qudit that is connected in the graph G, with the former being the target and the latter the control. Creating with that the state $|\Psi_2\rangle = \frac{1}{\sqrt{p^n}} \sum_{x \in \mathbb{F}_n^d} |x, e_{i,j}, \ldots, e_{l,k}\rangle$, with $e_{i,j} = v_j + v_i$ Mod p, given that $e_{i,j}$ is an edge between the vertexes v_j and v_i in G.

Step 3. Subsequently, one does measure all the qudits with the edge labeling. Previously, these were in a superposition of possible values for the sums of the computational bases of the respective vertexes. Now, these are fixed and this does fixes strictly that the quantum state over the vertex qudits fulfill the following relation,

$$v_u \equiv \left(\sum_{e_{i,j} \in Path(u,w)} (-1)^{|Path(u,j)|} \, e_{i,j} + (-1)^{|Path(u,w)|} v_w \right) \text{ Mod } p. \tag{4.6.59}$$

Considering a certain vertex in the graph, we immediately obtain that we have for this vertexes the superposition $\sum_{i \in \mathbb{F}_p} \frac{1}{\sqrt{p}} |i\rangle$, and that for each one of this computational basis states all the other computational basis states are fixed, such that, $|\Psi_3\rangle = \frac{1}{\sqrt{p}} \sum_{i \in \mathbb{F}^d} |x_i, e_{i,j}, \ldots, e_{l,k}\rangle$, with $x_1, \ldots, x_p \in \mathbb{F}_p^n$ and dictated by the relation described in Equation (4.6.59). In particular, if we select the first listed vertex v_0 as our reference, and with, we can describe the previous states as

$$|\Psi_3\rangle = \frac{1}{\sqrt{p}} \sum_{i \in \mathbb{F}_p} \left| v_0^{+i}, \ldots, v_j^{(-1)^{Path(0,j)} i}, \ldots, v_n^{(-1)^{Path(0,n)} i}, e_{i,j}, \ldots, e_{l,k} \right\rangle, \tag{4.6.60}$$

with $v^{+0} \in \mathbb{F}_p^n$.

Step 4. The state in the previous stage is very close to fulfilling the definition of a generalized qudit poor-mans cat state. The only issue is that for every vertex the path that is of odd size to the reference vertex, has the computational basis changing by the additive inverses of i of the reference qudit. In the final step, This can be simply solved by applying a INV gate to all these qudits in the "INV Routine". The final state, then is of the form $|\Psi_4\rangle = \frac{1}{\sqrt{p}} \sum_{i \in \mathbb{F}^d} |v^{+i}, \ldots, v_n^{+i}, e_{i,j}, \ldots, e_{l,k}\rangle$, with $v^{+0} \in \mathbb{F}_p^n$. This

exactly fulfills the definition of the poor-mans qudits states, $\left|\mathsf{GPM}_p^n\right\rangle$. Also, the individual z_i bits fulfill the following recurrence relation,

$$v_u \equiv \left(\sum_{e_{i,j} \in Path(u,w)} (-1)^{|Path(u,j)|+|Path(u,w)|} \, e_{i,j} + v_w \right) \text{ Mod } p. \tag{4.6.61}$$

□

Now we will show that there exists a QNC^0 circuit that can solve the probabilistic R_3^m ISMR problem, with a probability strictly larger than the upper bound determined for the classical solutions. We do the qutrit case first as it works as a specific instance and is easier to follow, and then we generalize for the arbitrary qudit prime dimension.

Lemma 43. *There exists a* QNC^0 *circuit that given a* $\left|\mathsf{GPM}_3^n\right\rangle$ *state, along with the corresponding graph* G *and a set of measured edges* $E = \{e_{i,j}, \ldots, e_{l,k}\}$*, is capable of solving the modular relation problem denoted as* $\mathcal{R}_3^m$*, over a uniform distribution of strings satisfying* $\sum_{i=1}^{n} x_i \text{ Mod } 3 = 0$ *in* $\mathbb{F}_2^n$*, with a success probability exceeding* $1/2$ *for any given input. Here,* $m = n \cdot g^2$*, the size of the outcome string, depends on the length* g *of the largest path in* G *from any vertex to a particular vertex that minimizes the maximum distance to any other vertex.*

Proof. For this proof, we consider the quantum circuit that archives this and is represented in Figure 23. Also, we consider the existence of a QNC^0 based on Lemma 59 that prepares the resource state, which will be in this case the $\left|\mathsf{GPM}_3^n\right\rangle$ state. Also, we consider the additional access to the corresponding graph G and string E for the measured edges in the state preparation.

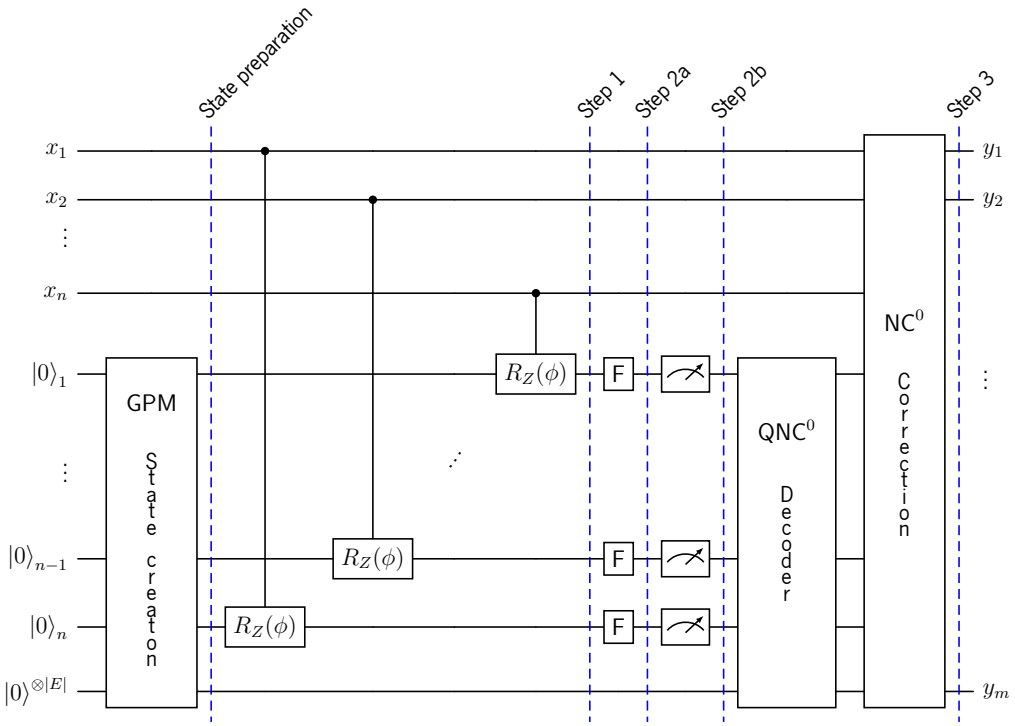


Figure 23: Illustration of a constant-depth quantum circuit solving all instances of the ISMR problems $\mathcal{R}_p^m$ with parameter p prime. Note that the last wire does carry the information about the edge measurements in the GPM state creation process $E = \{e_{i,j}, \ldots, e_{l,k}\}$.

Step 1. In this step, controlled qudit Z rotations will be applied to each one of the qudits of the resource state. The control of each one of these rotations will be a bit of input string x. Obtaining with that the following state,

$$|\Psi_1\rangle = \bigotimes_{i=1}^{n} R_z\Big(\frac{2\pi x_i}{9}\Big)\,|\mathsf{GPM}_3^n\rangle = \frac{1}{\sqrt{3}}\sum_{i=0}^{2} e^{\frac{2\pi i\langle x, z^{+i}\rangle}{9}}\,\big|z^{+i}\big\rangle\,. \tag{4.6.62}$$

Although the previous expression accurately represents the resulting state, it would be beneficial to reformulate the state in a manner more conducive to further analysis. Specifically, we aim to show that the resulting phases of these states, when rewritten as follows, more effectively capture the output strings obtained after applying step 2,

$$|\Psi_1\rangle = e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}}\frac{1}{\sqrt{3}}\sum_{i=0}^{2} e^{\frac{2\pi i\cdot\langle x,z^{+i}\rangle-\langle x,z\rangle}{9}}\,\big|z^{+i}\big\rangle \tag{4.6.63}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}}\frac{1}{\sqrt{3}}\Big(\,|z\rangle + \sum_{i=1}^{2} e^{\frac{2\pi i\cdot\left(2\langle z^{+i},x\rangle+(\sum_{j\in\{1,2\}\setminus i}\langle z^{+j},x\rangle)\right)}{9}}\,\big|z^{+i}\big\rangle\Big) \tag{4.6.64}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}}\frac{1}{\sqrt{3}}\Big(\,|z\rangle + e^{\frac{2\pi i\cdot(|x|/3+\langle x,(z^{+1})^2\rangle)}{3}}\,\big|z^{+1}\big\rangle + e^{\frac{2\pi i\cdot(2|x|/3-\langle x,z^2\rangle)}{3}}\,\big|z^{+2}\big\rangle\Big). \tag{4.6.65}$$

We begin by highlighting the initial phase component of z^{+0}, or equivalently z, by treating it as a universal phase factor. Then, the transition to the second expression is achieved by considering that $\sum_{i=0}^{2}\langle x, z^{+i}\rangle = 3|x|$, and rewriting the various phase terms. Likewise, given our assumption that $|x|$ Mod $3 = 0$, the term $\frac{3|x|}{9}$ will be matching to a multiple of 2π and can thus be disregarded.
For the last step, we need to examine the inner products present in the phases. For simplicity, we will analyze the ideal cases where z is uniformly equal,

Phases	$z = 0^{\otimes n}$	$z = 1^{\otimes n}$	$z = 2^{\otimes n}$
$2\langle z^{+1}, x\rangle + \sum_{j\in\{1,2\}\setminus 1}\langle z^{+j}, x\rangle$	$4\|x\|$	$4\|x\|$	$\|x\|$
$2\langle z^{+2}, x\rangle + \sum_{j\in\{1,2\}\setminus 2}\langle z^{+j}, x\rangle$	$5\|x\|$	$2\|x\|$	$2\|x\|$

.

Although z can assume any value in $\mathbb{F}_3^n$, we can divide the string into three sub-components corresponding to the aforementioned strings. Specifically, the first term contributes at least $|x|$ and the second term with $2|x|$, while one can introduce additional terms that determine the correct phases based on the specific random string z. This determination is straightforward due to an additional contribution of magnitude 3 in cases where the bits z_i are either 0 or 1 for the first phase term, which can be computed using the expression $\langle x, (z^{+1})^2\rangle$. The same principle applies to the second phase, with the additional contribution determinable via the expression $\langle x, z^2\rangle$, resulting in the final expression.

Step 2. In this step, we will first apply the respective Fourier gate F_3 to each one of the qudits of the state, and analyze the resulting support of the outcome state,

$$|\Psi_{2a}\rangle = \bigotimes_{l=1}^{n} F_3\left(e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}}\frac{1}{\sqrt{3}}\left(|z\rangle + e^{\frac{2\pi i\cdot(|x|/3+\langle x,(z^{+1})^2\rangle)}{3}}\,\big|z^{+1}\big\rangle + e^{\frac{2\pi i\cdot(2|x|/3+\langle x,z^2\rangle)}{3}}\,\big|z^{+2}\big\rangle\right)\right) \tag{4.6.66}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}}\frac{1}{\sqrt{3}}\Big(\sum_{y\in\mathbb{F}_3^n} e^{\frac{2\pi i\langle y,z\rangle}{3}}\,|y\rangle + e^{\frac{2\pi i\cdot(|x|/3+\langle x,(z^{+1})^2\rangle)}{3}+\frac{2\pi i\langle y,z^{+1}\rangle}{3}}\,|y\rangle \tag{4.6.67}$$

$$+ e^{\frac{2\pi i\cdot(2|x|/3+\langle x,z^2\rangle)}{3}+\frac{2\pi i\langle y,z^{+2}\rangle}{3}}|y\rangle\Big) \tag{4.6.68}$$

$$=e^{\frac{2\pi i\cdot\langle x,z\rangle}{9}+\frac{2\pi i\langle y,z\rangle}{3}}\frac{1}{\sqrt{3}}\sum_{y\in\mathbb{F}_3^n}|y\rangle+e^{\frac{2\pi i\cdot(|x|/3+|y|+\langle x,(z^{+1})^2\rangle)}{3}}|y\rangle+e^{\frac{2\pi i\cdot(2(|x|/3+|y|)+\langle x,z^2\rangle}{3}}|y\rangle\,. \tag{4.6.69}$$

Now, we can analyze the outcomes of the measurements after step 2b. In particular, let's consider the ideal case where the inner products $\langle x,(z^{+1})^2\rangle$ and $\langle x,z^2\rangle$ are both zero. In this scenario, we find that for $|x|/3+|y|$ Mod $p=0$, the three phases are equal to 1, and all the base states y meeting this criterion are measured with equal probability. Specifically, the strings y, observed after measurement, are all congruent with the additive inverse of $|x|/p$. Also, we do obtain that, whenever the $|x|/3+|y|$ Mod p equals either 1 or 2, it generates all three roots of unity as one can see in Equation (4.6.69). These basis states then have a zero probability of being observed. This proves that, for these ideal values of the inner products $\langle x,(z^{+1})^2\rangle$ and $\langle x,z^2\rangle$, we can perfectly solve the inverse strict modular relation problem $\mathcal{R}_3^n$. Unfortunately, these two additional terms in the phase can take any value in $\mathbb{F}_3$, and therefore scramble the outcomes that we intended to observe. Forcing us to analyze all possible cases.

Step 3. The final step, is intended to handle the randomness resulting from the inner products in the phases of Equation (4.6.69). In particular, we can process the outcome such that we increase the probability of having the correct outcomes to the desired relation problem at hand. Moreover, we will list all possible values, and divide their analysis into three cases.

Case 1	a_1	b_1	c_1
$\langle x,(z^{+1})^2\rangle$	1	1	0
$\langle x,z^2\rangle$	2	0	2

Case 2	a_2	b_2	c_2
$\langle x,(z^{+1})^2\rangle$	2	2	0
$\langle x,z^2\rangle$	1	0	1

Case 3	a_3	b_3
$\langle x,(z^{+1})^2\rangle$	2	1
$\langle x,z^2\rangle$	2	1

Case 1. For the assignment of values to the inner product displayed in a_1, the measured strings y are congruent with the additive inverse of $|x|/3+1$, due to the phases being related to $|x|/3+|y|+1$ Mod $3=0$. Furthermore, in cases b_1 and c_1, we obtain strings y that fulfill the same property with a probability of $1/2$.
In these scenarios, more than half of the subcases map to an incorrect string. However, more than half of the subcases also map to a string that deviates by the addition of two to its Hamming weight. This effect can be simply countered by adding the value 1 (it's additive inverse). Nevertheless, this addition must be informed by the specific cases under consideration. Therefore, we will use the term $\langle x,(z^{+1})^2\rangle$ for this purpose, and the method to construct this value will be described later. Ultimately, given that the probability of event a is non-zero, all these cases map to the correct string with a probability greater than $1/2$.

Case 2. For a_2, we encounter the inverse scenario of a_1. In this case, the states are projected onto strings that are equal to the additive inverse of $(|x|/3+2)^{-1}$, resulting in a shift of 1 from the correct Hamming weights. The subcases described in b_2 and c_2 exhibit a similar effect, mapping to the correct string half of the time, while in the other instances, the strings have their correct Hamming weights increased by one. It's noteworthy that, by applying the same correction as in case 1, specifically the term $\langle x,(z^2+1)^2\rangle$, the strings in a_2 which were all previously shifted, are now corrected. Furthermore, the two subcases b_2 and c_2 continue to map correctly half of the time. This, following the same line of reasoning presented previously, demonstrates that for all these cases, the input strings are mapped to the correct result with a probability greater than $1/2$.

Case 3. The last two cases, a_3 and b_3 are instances where both map with a probability of $1/2$ to the two incorrect strings. These strings are congruent to $(|x|/3+1)^{-1}$ and $(|x|/3+2)^{-1}$. Therefore, by applying our correction term $\langle x,(z^2+1)^2\rangle$, we ensure that one of them is correctly mapped in both scenarios, thus

correctly solving half of these instances. Additionally, by considering the ideal case discussed initially, which is unaffected by the corrections, more than half of the outcome strings are correct.

Correction string. Finally, as previously demonstrated, our correction is effective in all cases, ensuring consistency across all computed probabilities. As a result, we obtain final measured strings capable of solving the problem with a probability greater than $1/2$, given any $|\mathsf{GPM}\rangle$ characterized by arbitrary z and G. However, to fully substantiate this conclusion, one must demonstrate that the term $\langle x, (z^{+1})^2 \rangle$ can be computed using an NC^0 circuit. To achieve this, we can compute this value using the following expression,

$$\langle x, (z^{+1})^2 \rangle = \sum_{i \in [n]} x_i \cdot \left(\sum_{e_{i,j} \in Path(i,c)} (-1)^{|Path(i,j)|+|Path(i,c)|} \, e_{i,j} + z_c + 1 \right)^2 \text{ Mod } 3. \quad (4.6.70)$$

The inner sum $\sum_{e_{i,j} \in Path(i,c)} (-1)^{|Path(i,j)|+|Path(i,c)|} \, e_{i,j}$ can be constructed by enumerating the values of the edge measurements along the respective paths. Then, achieving the sum of the two values z_0 and 1 can be straightforwardly accomplished by appending these digits to the string. To realize the square, we only need to generate all the terms that would result from squaring this expression and perform the individual multiplication operations over $\mathbb{F}_3$ for these elements. Subsequently, all these bits are multiplied with the respective input bit x_i, and this process is repeated for all values of i in the outer sum. The total number of sums is n while using the largest path g we have that the exponentiation creates g^2 terms at most and we end up with ng^2 size correction string.
We have been describing these operations over $\mathbb{F}_3$, but one can perform these operations over $\mathbb{F}_2$ using an inverse of the Hamming encoding and the respective logical operations considering our encoding. This approach allows us to use an NC^0 circuit for this reduction after using a QNC^0 decoder to map the measured dits to bits. Finally, the binary version of the string $\langle x, (z^{+1})^2 \rangle$ is added to the measurement outcomes from stage 3, as previously described, thereby concluding the proof. □

The generalization to arbitrary qupits does follow the same ideas as previously. However, the states generated by the randomness of the qupit poor-mans cat states are more evolved. This will be encapsulated in the following lemmas, with the added consideration that one can solve the modular inverted strict relation problems closer to the correct outcomes within the Abelian domains and with a probability that exceeds mere random guessing.

Lemma 44. *A QNC^0 circuit, denoted as C_q, is given a state $\left|\mathsf{GPM}_p^n\right\rangle$, the corresponding graph G, and a set of measured edges $E = \{e_{i,j}, \ldots, e_{l,k}\}$. This circuit is capable of solving the modular relation problem, represented as $\mathcal{R}_p^m$, across a uniform distribution $\mathcal{D}_p$ of strings that satisfy the condition $\sum_{i=1}^n x_i$ Mod $p = 0$ within $\mathbb{F}_2^n$. It achieves this with a success probability of $\frac{2p-2}{p^2}$ for any given input. Moreover, it maintains an average correlation with the correct outcome within the Abelian domain,*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(C, \mathcal{R}_p^m\right) = \frac{p-1}{p^2}. \quad (4.6.71)$$

In this context, $m = n \cdot g^{p-1}$, where m represents the size of the outcome string. This size is influenced by g, which is the length of the longest path within G from any vertex to a designated vertex that minimizes the greatest distance to all other vertices.

Proof. To prove this, we will once again consider the circuit that effectively achieves this. In particular, we will revisit the circuit presented in Figure 23, as it possesses the structure necessary to solve all ISMR problems by simply adapting the correct prime-dimensional Hilbert space for each stage. Starting, with

the consideration that a resource state for the form of a $\left|\mathsf{GPM}_p^n\right\rangle$ state will be provided in addition to the corresponding graph G and string E for the measured edges in the state preparation.

Step 1. The first step applies controlled qudit Z rotations parameterized by the angle $\phi = \frac{2\pi}{p^2}$ to the resource state, while the control bits of each one of these rotations will refer to the input string x. Obtaining with that the following state,

$$|\Psi_1\rangle = \bigotimes_{i=1}^{n} R_z\left(\frac{2\pi x_i}{p^2}\right)\left|\mathsf{GPM}_p^n\right\rangle = \frac{1}{\sqrt{p}}\sum_{i=0}^{p-1} e^{\frac{2\pi i\langle x, z^{+i}\rangle}{p^2}}\left|z^{+i}\right\rangle \tag{4.6.72}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{p^2}}\frac{1}{\sqrt{p}}\left(|z\rangle + \sum_{i=1}^{p-1} e^{\frac{2\pi i\cdot\left(2\langle z^{+i},x\rangle + (\sum_{j\in\{1,\ldots,p-1\}\setminus i}\langle x, z^{+j}\rangle)\right)}{p^2}}\left|z^{+i}\right\rangle\right) \tag{4.6.73}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{p^2}}\frac{1}{\sqrt{p}}\left(|z\rangle + \sum_{i=1}^{p-1} e^{\frac{2\pi i\cdot\left(i|x|/p + (\sum_{j\in 1,\ldots,i}\left(\langle x, z^{+j}\rangle\right)^{p-1}\right)}{p}}\left|z^{+i}\right\rangle\right). \tag{4.6.74}$$

All the reductions until the last transition from Equation (4.6.73) to Equation (4.6.74) follow exactly equally as the reduction from Lemma 43. For the last, one does need to redo the analysis of the contribution to the resulting phases by each string of the type $z = 0^n, 1^n, \ldots, (p-1)^n$. In particular, one can determine that for these values of z, we have that,

Phases	$z = 0^{\otimes n}$	$\ldots$	$z = (p-2)^{\otimes n}$	$z = (p-1)^{\otimes n}$
$2\langle z^{+1}, x\rangle + \ldots$	$(p\Pi(p) - p + 1)\lvert x\rvert$	$\overset{r}{\leftrightarrow}$	$(p\Pi(p) - p + 1)\lvert x\rvert$	$(p\Pi(p) - 2p + 1)\lvert x\rvert$
$2\langle z^{+2}, x\rangle + \ldots$	$(p\Pi(p) - p + 2)\lvert x\rvert$	$\overset{r}{\rightarrow}$	$(p\Pi(p) - 2p + 2)\lvert x\rvert$	$(p\Pi(p) - 2p + 2)\lvert x\rvert$
$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
$2\langle z^{+(p-1)}, x\rangle + \ldots$	$(p\Pi(p) - 1)\lvert x\rvert$	$\overset{r}{\leftarrow}$	$(p\Pi(p) - p - 1)\lvert x\rvert$	$(p\Pi(p) - p - 1)\lvert x\rvert$

with $\Pi(p)$ representing the prime counting function, which determines the number of primes up to and including p. Additionally, the symbol $\overset{r}{\leftrightarrow}$ indicates that the value on the left and right repeats for all strings z in between. In contrast, $\overset{r}{\rightarrow}$ signifies that the value from the preceding string is repeated for all intermediate strings, while $\overset{r}{\leftarrow}$ implies that the value from the string on the right is repeated in between. Using these phases it is easy to account for all the terms contributions in the phase by the internal product of the input with the random string z in Equation (4.6.74).

Step 2. Subsequently, we will consider the effect of applying the respective Fourier gates F_p to the qudits of the resource state. More precisely we obtain the following states,

$$|\Psi_{2a}\rangle = \bigotimes_{l=1}^{n} F_p\left(e^{\frac{2\pi i\cdot\langle x,z\rangle}{p^2}}\frac{1}{\sqrt{p}}\left(|z\rangle + \sum_{i=1}^{p-1} e^{\frac{2\pi i\cdot\left(i|x|/p + (\sum_{j\in 1,\ldots,i}\langle x, (z^{+j})^{p-1}\rangle\right)}{p}}\left|z^{+i}\right\rangle\right)\right) \tag{4.6.75}$$

$$= e^{\frac{2\pi i\cdot\langle x,z\rangle}{p^2}}\frac{1}{\sqrt{p}}\left(\sum_{y\in\mathbb{F}_p^n} e^{\frac{2\pi i\langle y,z\rangle}{p}}|y\rangle + \sum_{i=1}^{p-1} e^{\frac{2\pi i\cdot\left(i|x|/p + (\sum_{j\in 1,\ldots,i}\langle x, (z^{+j})^{p-1}\rangle) + \frac{2\pi i\langle y, z^{+j}\rangle}{p}\right.}{p}}|y\rangle\right) \tag{4.6.76}$$

$$=e^{\frac{2\pi i\cdot\langle x,z\rangle}{p^2}+\frac{2\pi i\langle y,z\rangle}{p}}\frac{1}{\sqrt{p}}\left(\sum_{y\in\mathbb{F}_p^n}e^{\frac{2\pi i\langle y,z\rangle}{p}}|y\rangle+\sum_{i=1}^{p-1}e^{\frac{2\pi i\cdot\left(i|x|/p+|y|+(\sum_{j\in1,\ldots,i}\langle x,(z^{+j})^{p-1}\rangle\right)}{p}}|y\rangle\right). \tag{4.6.77}$$

The resulting state from step 2a will then be measured. Once again, if all the inner products of the type $\langle x,(z^{+i})^{p-1}\rangle$ are equal to zero, the final string y measured will fulfill $|x|/p+|y|$ Mod $p=0$. Therefore, it would produce the correct outcome for each of the inverse strict modular relation problems.
Regrettably, the values of these terms vary across the field $\mathbb{F}_p$. However, they assume uniformly random values within $\mathbb{F}_p$, facilitating the prediction of precise outcomes. More precisely, this allows us to consider the support of the outcome strings and their respective probabilities based on these values and the input strings.
In particular, for this, we will consider all possible vectors such as $(0,0,\ldots,0)$ and $(0,1,2,\ldots,p-1)$ to $(p-1,0,1,\ldots,p-2)$ with a simple inline shift, which represents the values of the terms $\langle x,(z^{+i})^{p-1}\rangle$ that induce a shift in the correct outcome by an increment of 0 and 1 to $p-1$, respectively. Following this representation, we can assert that it is possible to decompose any of the possible values that these terms might adopt into vectors of the form $(a_1,a_2,\ldots,a_{p-1})$. The degree of overlap between these vectors and the aforementioned basis vectors directly determines the probability of measuring a string offset related to the corresponding basis vector. Given this description and the uniform distribution of all a_i within $\mathbb{F}_p$, the probability of yielding each shift relative to the accurate outcome string is uniform. This ensures that a correct string is produced with a probability of $1/p$ and every incorrect string equally with a probability of $1/p$.

Step 3. This process now aims to use the information provided by the state's creation, which defines the random string z, to increase the probability of accurately computing the solutions to inverse strict modular relation problems.
To achieve this, we will begin by analyzing the vectorization that determines the probability of obtaining the outcome with a shift in its Hamming weight originating from the randomness in the creation of the generalized poor man's cat state. More specifically, we consider the use of a single term representative of one of the inner product terms $\langle x,(z^{+i})^{p-1}\rangle$. Thus, by selecting the first a_1 from $(a_1,a_2,\ldots,a_n)$ and then adding the inverse shift associated with $(a_1,a_1+1,a_1+2,\ldots,a_1+(p-1))$, we ensure that the outcome string is corrected for this component. In particular, we find that the outcome strings are shifted based on the inner product with a vector of the type $(0,a_2*,a_3*,\ldots,a_{p-1}*)$ and all the basis vectors, with all the values $a_2*,a_3*,\ldots,a_{p-1}*$ remaining uniformly random over $\mathbb{F}_p$.
This ensures that the overlap with a zero shift is $\frac{2p-2}{p^2}$, and all the other shifts from 1 to $p-1$ have probability $\frac{p-1}{p^2}$. Also, the average overlap with the correct outcomes within the Abelian domains can be determined as being

$$\mathrm{Corr}_{\mathcal{D}_p}\left(|C|,|\mathcal{R}_p^m|\right)=\frac{p}{2^n}\sum_{\substack{x\sim\mathbb{F}_2^n\\ |x|\ \mathrm{Mod}\ p=0}}\mathrm{Re}\left(e^{-i\frac{2\pi|C_q(x)|}{p}}e^{i\frac{2\pi|\mathcal{R}_p^m(x)|}{p}}\right) \tag{4.6.78}$$

$$=\frac{p-1}{p^2}, \tag{4.6.79}$$

with $\mathcal{D}_p$ being the uniform distribution over string in $\mathbb{F}_2^n$ that satisfy $|x|$ Mod $p=0$.
This value can be easily obtained by dividing the probabilities of having a shift as follows: with a probability of $\frac{p-1}{p^2}$, there is no shift at all, and with a probability of $1-\frac{p-1}{p^2}$, there is a uniform distribution of shifts. Therefore, the second fraction of shifts does not contribute to the value, as this represents a sum over all the values for the roots of unity with equal probability and equals zero. Simultaneously, a $\frac{p-1}{p^2}$ fraction of

the inputs contributes to the value 1.

Correction string. For the correction, we only need to guarantee that the term $\langle x,(z^{+1})^{p-1}\rangle$ is computationally possible to produce within the considered class such that we can sum it to the outcome string y. To achieve this, we consider that we can determine this value using the following expression,

$$\langle x,(z^{+1})^{p-1}\rangle=\sum_{i\in[n]}x_i\cdot\left(\sum_{e_{i,j}\in Path(i,c)}(-1)^{|Path(i,j)|+|Path(i,c)|}\,e_{i,j}+z_c+1\right)^{p-1}\ \text{Mod}\ p. \tag{4.6.80}$$

The inner sum $\sum_{e_{i,j}\in\text{Path}(i,c)}(-1)^{|\text{Path}(i,j)|+|\text{Path}(i,c)|}\,e_{i,j}$ can be constructed by enumerating the edge measurement results along the respective paths. Additionally, the sum of the values z and 1 can be straightforwardly achieved by appending these values to the string. To exponentiate, it is sufficient to create all the terms that result from this sequence of products, and since the exponent is finite, each term is a product involving at most $p-1$ terms, making them all efficiently computable. These terms are then multiplied in $\mathbb{F}_p$ with the respective input bit x_i. The total number of sums is n while using the largest path g we have that the exponentiation creates g^{p-1} terms at most and we end up with ng^{p-1} size correction string.
Similar to the qutrit case, one can use a QNC^0 decoder to map the measured dits to bits based on an inverse of the Hamming encoding. Therefore, all the previously described operations over $\mathbb{F}_p$ can be performed over $\mathbb{F}_2$ using the respective logical operations with an NC^0 circuit. In conclusion, the string $\langle x,(z^{+1})^{p-1}\rangle$ in bits is concatenated with the measurement outcomes from stage 3, thereby concluding the proof. □

Quantum vs. Classical circuit separations

We will now combine the lower bounds determined for NC^0 and $\mathsf{bPTF}^0[k]$ circuits in Section 4.6.2 with the upper bounds established for QNC^0 circuits in this section to derive new quantum-classical separations. We begin by stating explicit separations of qupit QNC^0 circuits against NC^0 circuits.

Theorem 4.6.5. *For every* ISMR *problem defined for a prime dimension p, denoted by $\mathcal{R}_p^m$, considering a uniform distribution $\mathcal{D}_p$ over $\mathbb{F}_2^n$ of strings with Hamming weight that satisfies $(\sum_{i=1}^n x_i)$ Mod $p=0$ as the input distribution, there exists a QNC^0 circuit with all-to-all connectivity that solves the problem for $m=n\cdot\log n^{p-1}$ archiving a correlation with the correct outcome of $\frac{p-1}{p^2}$. In contrast, any $\mathsf{NC}^0/\text{rpoly}$ circuit with locality l fails to solve this problem with a correlation larger than,*

$$\text{Corr}_{\mathcal{D}_p}\left(C,\mathcal{R}_p^m\right)=\mathcal{O}\left(\text{Re}\left(\frac{2^{1-p}e^{\frac{2\pi i}{p^2}}\left(e^{-\frac{2\pi i}{p^2}}\left(1+e^{\frac{2\pi i}{p^2}}\right)\right)^p}{1+e^{\frac{2\pi i}{p^2}}}\right)^{\frac{n}{\log n(p-1)4l^2}}\right). \tag{4.6.81}$$

In particular, for $p=3$ there exist a QNC^0 circuit with 3D connectivity that solves this specific ISMRP *instance for $m=n^{4/3}$ with an efficiency larger the $1/2$ on all inputs, while the success probability for any $\mathsf{NC}^0/\text{rpoly}$ is bound by*

$$\Pr_{\substack{x\sim\mathbb{F}_2^n\\ |x|\ \text{Mod}\ 3=0}}[C(x)\in R_3^m(x)]\leq\frac{1}{3}+\left(\frac{9}{10}\right)^{\frac{n^{2/3}}{16l^2}}. \tag{4.6.82}$$

Proof. We consider the upper bound determined in Lemma 44 to ascertain the correlation with which one can compute each of the ISMR problems using a QNC^0 circuit with all-to-all connectivity. Concurrently,

we refer to Lemma 40 for the classical lower bound of the same quantity, utilizing the values for $m = n \cdot \log n^{p-1}$ obtained by the quantum solution and the considered input distribution.
For the specific case where $p = 3$, we examine the lower bound on the success probability determined in Lemma 43 for a qutrit QNC^0 circuit with 3D connectivity, alongside the classical lower bound in Lemma 40 for the respective input distribution and output size of $m = n^{4/3}$. □

The previous separation can be extended using the techniques from the works of [33, 34] for parallel repetitions, such that the success probability decreases exponentially fast to zero. This results in an exponential separation between the quantum and classical efficiencies, specifically for QNC^0 and NC^0 circuits, across the entire class of inverted strict modular relation problems. Additionally, due to our findings in Corollary 7, it is possible to accelerate this convergence. A different input distribution allows for larger worst-case separations. We do not explicitly determine these results, as they are not the main focus of this document, and we intend to further explore the separation against the $\mathsf{bPTF}^0[k]$ class.
Finally, we combine the lower bounds for $\mathsf{bPTF}^0[k]$ and quantum upper bounds for the same problems. We determine that one can achieve separation between QNC^0 and $\mathsf{bPTF}^0[k]$ using any prime instance of the inverted strict modular relation problems with the asymptotically largest possible value of k.

Proof of Theorem 4.6.4. The quantum upper bound is derived by combining Lemma 59 and Lemma 44, which determines the precise correlation $\frac{p-1}{p^2}$. The output size m, which allows us to solve each of the problems using a QNC^0 circuit over qupits, follows from considering the all-to-all connectivity. Specifically, it is defined that the minimum-maximum path g in the graph will have a size of $\log n$, which in turn defines m to be $n \cdot \log n^{p-1}$.
Simultaneously, the classical lower bounds are derived by considering Lemma 41, with all parameters being defined by the quantum solution. Furthermore, all additional values for the expressions, as well as the optimal values of q, are determined in the same manner as in Theorem 4.6.1. □

Finally, as one can determine success probabilities for the ISMRP $\mathcal{R}_p^m$ where $p = 3$, just as one can for the binary case where $p = 2$, we can establish a separation in the success probabilities for this problem. Additionally, this case is particularly intriguing as it demonstrates that all instances of the $\mathcal{R}_3^m$ problem can be solved with a probability strictly greater than $1/2$ in the quantum case, whereas any classical circuit solves the problem with at most a probability asymptotically close to random guessing, which for this problem corresponds to a success probability of $1/3$. This indicates that QNC^0 effectively addresses the probabilistic version of $\mathcal{R}_3^m$, adhering to the standard definition for bounded-error probabilistic problems, while any classical circuit within the $\mathsf{bPTF}^0[k]$ class fails to solve the same problem. Furthermore, the quantum circuit can exhibit this advantage while maintaining geometric locality in a 3D geometry.

Proof of Corollary 6. We derive the lower bounds for the probabilities of the quantum solution using Lemma 43 and the parameters $m = n^{4/3}$ from the specific 3D connectivity.
Then, we consider the lower bounds for the probability of solving the $\mathcal{R}_3^{n^{4/3}}$ defined by Lemma 38. Finally, the determination of the values of q for both values of k follows equally as in Theorem 4.6.1. □

In conclusion, all the newly derived bounds, including those applicable to the qubit scenario, could potentially be extended with parallel repetition games, thereby enhancing the quantum advantage even further. We have not yet pursued this path, as our focus has been on expanding the problem set and tightening the bounds. Our priority is to establish unconditional separations against broader and more powerful classes of classical circuits, aiming to reduce the resources needed for a quantum advantage. This is because parallel repetition games only become valuable once the initial problem set, which is to be repeated, has achieved a quantum advantage.

4.7 Noise-resilient quantum advantage

In this section, we demonstrate that the separations we proved against $\mathsf{bPTF}^0[k]$ in the preceding section can be made noise-robust. Specifically, we show that a separation can still be achieved even when our quantum circuits are noisy and classical circuits are noiseless. As before, we present the qubit and qupit cases separately.
Both separations are proven for new noise-tolerant relations $\mathfrak{R}_p$, defined based on the ISMR problems as follows.

Definition 42. *Let $\mathcal{R}_p : \mathbb{F}_p^n \mapsto \mathbb{F}_p^{n'}$ be the original* ISMR *problem (Definition 38). Then, the noise tolerant extension $\mathfrak{R}_p : \mathbb{F}_p^n \times \mathbb{F}_p^{n'\cdot m} \mapsto \mathbb{F}_p$ is defined as follows,*

$$\mathfrak{R}_p(x, y) = \left(\left| \mathcal{R}_p^{n'}(x) \right| - \left| \mathsf{dec}^*(y) \right| \right). \tag{4.7.1}$$

Here dec* *is an arbitrary functions of the form* dec$^* : \mathbb{F}_p^{n'\cdot m} \mapsto \mathbb{F}_p^{n'}$ *which must be computable by an* AC^0 *circuit.*

We begin by presenting our noise-robust separation for qupits with arbitrary prime p, which is the more intricate case. Notably, for qupits with $p \geq 3$, the quantum circuits solving the candidate relation problems are not Clifford circuits. This necessity arises because Bell violations with stabilizer states alone do not generalize from qubits to qupits with $p \geq 3$ [121, 122, 123]. This indicates that quantum strategies can not have larger winning probabilities than classical ones if only Clifford circuits are considered for general qupits. Therefore, for this separation, we require quantum advice states, specifically magic states in their logical form, with the same code distance under which the remaining quantum circuit will operate. This requirement stems from the additional conjecture that any magic state factory that creates a logical magic state with at least poly$(\log n)$ code distance is not realizable within QNC^0. Simultaneously, we must resolve the incompatibilities between the presented quantum circuits, solving the ISMR problems from Section 4.6.2, and the quantum circuit architecture for non-Clifford circuits described and proven to be noise-resistance in Section 4.7.1, Section 4.7.1 and Section 4.7.1. Solving the previous incompatibility involves demonstrating that quantum circuits equipped with non-adaptive magic state injection gadgets can solve all the ISMR problems with equal effectiveness, as shown in Section 4.7.2.

Theorem 4.7.1. *Let x be an input drawn uniformly at random from the subset of n-bit strings with Hamming weight satisfying $(\sum_{i=1}^n x_i)$ Mod $p = 0$. Consider the local stochastic noise $\mathcal{E} \sim \mathcal{N}(\tau)$ with bounded probability $\tau < \tau_{th}$. Then, there is a $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuit that solves the relation $\mathfrak{R}_p(x, y)$ for output strings $y = o(n^2)$ with a constant positive correlation.*
Further, any circuit $C \in \mathsf{bPTF}^0[k]/\mathsf{rpoly}$ with fixed depth d and size s, for large enough $d' \in \mathbb{N}^+$ has exponentially small correlation with $\mathfrak{R}_p$ bounded by

	$k = \mathcal{O}(1)$ $(\equiv \mathsf{AC}^0/\mathsf{rpoly})$	$k = n^{1/(5d)}$
$\mathsf{Corr}_{\mathcal{D}_p}(\mathfrak{R}_p(x, C))$	$\exp\left(-\Omega\left(\frac{n^{1-\mathcal{O}(1)}}{(\log n)^{p-1}(\log s)^{2(d+d')-2}}\right)\right)$	$\exp\left(-\Omega\left(\frac{n^{3/5-\mathcal{O}(1)}}{(\log n)^{p-1}(\log s)^{2(d+d')-2}}\right)\right)$

We also obtain a separation for the qubit case. Later in this section, we highlight the distinctions between the two cases. However, the main difference is that the qubit case allows for an advice-free separation between the noisy QNC^0 and noiseless $\mathsf{bPTF}^0[k]$ circuits. Therefore, we will with our qubit separation lift the previously proven noise-resilient separations from NC^0 and AC^0 to $\mathsf{bPTF}^0[k]$ for optimal k [35, 37].

Corollary 9. *Let x be an input drawn uniformly at random from the set of binary even strings. Consider the local stochastic noise $\mathcal{E} \sim \mathcal{N}(\tau)$ with bounded probability $\tau < \tau_{th}$. Then, there is a* QNC^0 *circuit that solves the relation $\mathfrak{R}_2(x, y)$ for output string $y = o(n^2)$ with a constant probability bounded away from 1/2. However, any circuit $C \in$* $\mathsf{bPTF}^0[k]/\mathsf{rpoly}$ *with fixed depth d and size s, for parameter $k = 1/(5d)$ large enough $d' \in \mathbb{N}^+$ has exponentially small correlation with $\mathfrak{R}_2$ bounded by*

$$\Pr[\mathfrak{R}_p(x, C(x)) = 0] = \frac{1}{2} + \exp\left(-\Omega\left(\frac{n^{3/5-\mathcal{O}(1)}}{(\log s)^{2(d+d')-1}}\right)\right). \tag{4.7.2}$$

The correlation bounds previously described differ from the noise-free bounds in Theorem 4.6.4, particularly in the exponent of the $\log(s)$ term, which arises from the additional overhead for error correction. However, this does not compromise the exponential deviation between the classical and quantum correlations achieved by the respective circuit classes.

To demonstrate these separations with higher-dimensional systems, we show that the roadmap established in [35] for the qubit case can be adapted to extend this advantage in the presence of noise. We will proceed in several steps. First, we outline the requirements for the fault-tolerant construction in constant depth for the higher-dimensional versions of the code, including the additional state injection. Next, we introduce a new decoder for the surface code over qudits, highlighting the differences between the two cases. Then, we generalize the single-shot state preparation for the two-qudit states known as GHZ_2 (generalized qudit Bell states). Finally, we identify new non-adaptive qupit circuits based on the magic state injection gadget, solving the new noise-resilient relation problems to establish our bounds against $\mathsf{bPTF}^0[k]$.

4.7.1 Noise-resilient qupit Clifford circuits with quantum advice

To demonstrate our noise-robust separations, we introduce constant-depth quantum circuits that incorporate logical quantum states as advice, as illustrated in Figure 24. This model adopts the structure used in standard error-corrected quantum circuits with magic state injection, presenting broader interest. It is important to note, however, that while our circuit architecture is noise-robust, it does not support the error-resilient realization of adaptive quantum circuits that include gates controlled by prior measurement outcomes. Instead, it is limited to error-resilient, non-adaptive constant-depth Clifford circuits with magic state injection. Thus, we restrict our analysis to a subset of QNC^0 circuits for a noise-resilient version of these circuits.

The difficulty in realizing adaptive measurements in a noise-resilient manner arises because the code distance necessary for good error correction properties unfortunately also requires the execution of the decoding function, which is beyond the computational capabilities of QNC^0 circuits with the considered codes and decoder. This creates a conundrum where either the noise levels render the outcome useless or the information is recoverable but not by the circuit class QNC^0 itself. Therefore, we focus on non-adaptive constant-depth Clifford circuits with magic state injection, which we will show in Section 4.7.2 are sufficient to obtain noise-robustness for our qupit separations.

In particular, we will consider the conditions a quantum error correcting code must satisfy, so that we may use it to obtain a noise-resilient version of circuits with this structure. We will prove that these conditions are sufficient, and subsequently demonstrate that the qupit surface code, along with the hard normalization decoder, fulfills all the requirements for a noise-resilient execution of a non-adaptive Clifford circuit with advice states.

Definition 43 (Error-resilient constant-depth quantum code conditions)**.** *A CSS-type code Q_m exhibiting the following properties gives constant depth fault tolerant circuit constructions for qudit local stochastic noise $\mathcal{E} \sim \mathcal{N}(\varrho)$.*

1. *Any qudit advice state $|A\rangle$ required must be provided in its logical form $|\overline{\mathcal{A}}\rangle$, while being affected*

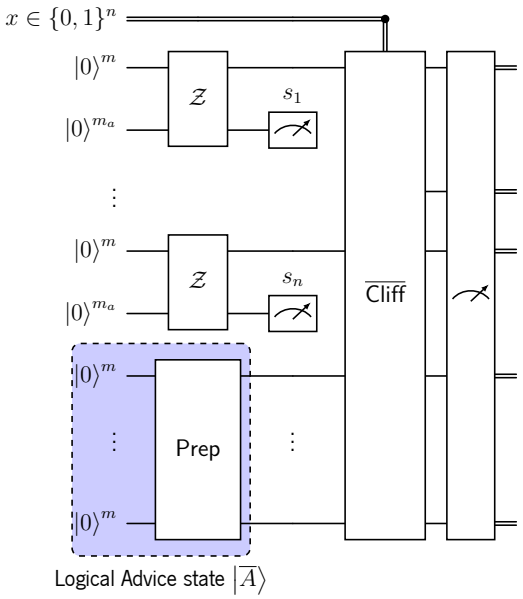


Figure 24: Constant-depth non-adaptive Clifford quantum circuit with quantum advice states.

by local stochastic noise $\mathcal{E}_A \sim \mathcal{N}(\tau)$ *at most as follows,*

$$\left|\overline{\mathcal{A}}\right\rangle \propto \mathcal{E}_A \, \mathsf{prep}(|A\rangle) \, |0\rangle^{\otimes(|A|\cdot m)} \, . \tag{4.7.3}$$

where $\mathsf{prep}(|\mathcal{A}\rangle)$ *represents the noise-free quantum circuit that prepares the logical state* $\left|\overline{\mathcal{A}}\right\rangle$ *over a code of distance* m*.*

2. *The logical qudit Clifford gates can be implemented in constant depth (as defined in Section 4.4.4).*

3. *There exist recovery and repair functions*

$$\begin{aligned} &\mathsf{rec} : \mathbb{F}_p^{m_a} \mapsto \mathsf{Pauli}(m) \\ &\mathsf{rep} : \mathsf{Pauli}(m + m_a) \mapsto \mathsf{Pauli}(m) \end{aligned}$$

respectively, where $\mathsf{Pauli}(n)$ *respresents a Pauli operator over* n *qupits, such that the application of a constant depth Clifford circuit* $\mathcal{Z}$ *on the state* $|0\rangle^{\otimes m} \otimes |0\rangle^{m_a}$ *followed by the* Z*-measurement of its ancillary qudits, leading to an outcome* $s \in \mathbb{F}_p^{m_a}$*, can be represented by the channel*

$$(|s\rangle\langle s| \otimes \mathsf{rec}(s)) \, \mathcal{Z} \left(|0\rangle^{\otimes m} \otimes |0\rangle^{\otimes m_a}\right) \propto |\bar{0}\rangle \otimes |s\rangle \, , \tag{4.7.4}$$

where $\left|\overline{0}\right\rangle$ *represents the logical zero state. In the case of local stochastic noise* $\mathcal{E} \sim \mathcal{N}(\varrho)$*, this becomes*

$$(|s\rangle\langle s| \otimes \mathsf{rec}(s)) \, \mathcal{E} \, \mathcal{Z} \left(|0\rangle^{\otimes m} \otimes |0\rangle^{\otimes m_a}\right) \propto \mathsf{rep}(\mathcal{E}) \, |\bar{0}\rangle \otimes |s\rangle \, , \tag{4.7.5}$$

for a local stochastic function $\mathsf{rep}(\mathcal{E}) \sim \mathcal{N}((c'\varrho)^{c''})$*.*

4. *There exists a decoding function* $\mathsf{dec} : \mathbb{F}_p^m \mapsto \mathbb{F}_p$ *such that*

$$\mathsf{dec}(y) := \mathsf{Mod}_p(y), \tag{4.7.6}$$

for $y \in \mathbb{F}_p^m$ *such that* y *belongs to the space of all basis vector obeying* $Z-$*type stabilizers. When subject to noise, for error rate* $q < q_{th}$ *i.e. a threshold parameter, and error string* $v \in \mathbb{F}_p^m$ *it is*

the case that

$$\mathsf{Pr}[\mathsf{dec}(x \oplus v) = \mathsf{Mod}_p(x)] \geq 1 - e^{(c' m^c)}, \tag{4.7.7}$$

for some c *and* $c' \in \mathbb{R}_{>0}$.

We have that any error correction code that fulfills the previous condition, will allow us to realize error resilient version of the any logical quantum circuit with the structure of Figure 24.

Lemma 45. *Let* $\mathcal{EC}_C$ *represent the error-corrected version of* C*, a logical classically controlled constant-depth Clifford circuit with advice states* $|A\rangle$*. Respecting the structure depicted in Figure 24 and utilizes a quantum error correction code featuring a decoding function* dec*, single-shot state preparation circuit* $\mathcal{Z}$*, a recovery function* rec*, and a repair function* rep*, all of which meet the conditions specified in Definition 43. Then, for an arbitrary input* $x \in \mathbb{F}_2^n$ *and denote the outputs of the noisy implementation* s *and* y *such that* $\mathcal{EC}_C(x) = (s, y)$*, with local stochastic noise* $\mathcal{E} \sim \mathcal{N}(\varrho)$ *affecting the circuit execution and* $\mathcal{E}_A \sim \mathcal{N}(\tau)$ *the advice state. Then, for* $m = \mathcal{O}(\mathsf{poly}\log n)$ *and all local stochastic noise with* $\varrho_{th} = 2^{-2^{\mathcal{O}(d)}}$*, where* d *is a constant representing the depth of the circuit, the following holds for all* τ *and* $\varrho < \varrho_{th}$*,*

$$\mathsf{Pr}[f(s, y) \equiv C(x)] > 0.99, \tag{4.7.8}$$

with $f(s, y) = \otimes_{i=0}^{n-1} \mathsf{dec}\,(y_i \oplus w_i)$ *for* w *such that* $\langle w|\, \overline{C}(x)\mathsf{rec}(s)^\dagger \overline{C}(x)^\dagger\, |0\rangle^{\otimes n \cdot m} = 1$.

Proof. We start by using the first condition of Definition 43, which guarantees that we can describe the advice states as $\left|\overline{\mathcal{A}}\right\rangle \propto \mathcal{E}_A\, \mathsf{prep}(A)\, |0\rangle^{\otimes |A| \cdot m}$, with $\mathcal{E}_A \sim \mathcal{N}(\tau)$. Additionally, given that the remainder of the logical circuit is a classically controlled Clifford circuit of constant depth $C(x)$, we have, according to condition 2 of Definition 43, that its logical version also is a Clifford circuit of constant depth. Therefore, all the Pauli errors that can occur due to the local stochastic noise under consideration can be commuted to the end using Lemma 24. Thus, the distribution $\mathcal{D}_{\mathcal{EC}_C}(x)$ over strings that result from the $\mathcal{EC}_C(x)$ circuit can be described as follows,

$$\mathcal{D}_{\mathcal{EC}_C}(x) = \langle s, y|\, \mathcal{E}_t\, \overline{C}(x)(\mathcal{Z} \otimes I)\, |0\rangle^{\otimes n \cdot m}\, |0\rangle^{\otimes n \cdot m_{anc}}\, \left|\overline{A}\right\rangle. \tag{4.7.9}$$

with $\mathcal{E}_t = \mathcal{N}(\tau')$ for $\tau' = \mathcal{O}(\max\left(\tau^{-2^d}, \varrho^{-2}\right))$ and ϱ being the noise level of the local stochastic noise affecting the state preparation circuit $\mathcal{Z}$ as well as $C(x)$ and the measurements.
Now, we need to describe the effect of the measurements in the single-shot state preparation to define the support of the circuit's outcome more precisely. For that, we commute the total error back through the Clifford circuit $\overline{C}(x)$, which we describe as $\mathcal{E}'_t = \overline{C}(x)\mathcal{E}_t\overline{C}(x)^\dagger$ and divide $\mathcal{E}'_t = \mathcal{E}'_{t_1} \otimes \mathcal{E}'_{t_2}$ with $\mathcal{E}'_{t_1}$ operating over the first $n \cdot (m + m_{anc})$ qudits, while $\mathcal{E}'_{t_2}$ operates over the last $|A|$ qudits of the advice state. Thus, we obtain that for each logical qupit that is affected by $\mathcal{E}'_{t_1}$ the following equality based on condition 3 of Definition 43,

$$\begin{aligned}
&(I \otimes |s\rangle\langle s|)\overline{C}(x)\, (\mathcal{E}'_{t_1} \otimes \mathcal{E}'_{t_2})(\mathcal{Z} \otimes I)\, |0\rangle^{\otimes n \cdot m}\, |0\rangle^{\otimes n \cdot m_{anc}}\, \left|\overline{A}\right\rangle && (4.7.10)\\
&\qquad \propto \overline{C}(x)(I \otimes \mathcal{E}'_{t_2})\mathsf{rec}(s)^\dagger \mathsf{rep}(\mathcal{E}'_{t_1})\, \left|\overline{0}\right\rangle^{\otimes n} |s\rangle\, \left|\overline{A}\right\rangle && (4.7.11)\\
&\qquad \propto \overline{C}(x)(\mathcal{E}''_{t_1} \otimes \mathcal{E}'_{t_2})\mathsf{rec}(s)^\dagger\, \left|\overline{0}\right\rangle^{\otimes n} |s\rangle\, \left|\overline{A}\right\rangle && (4.7.12)\\
&\qquad \propto \mathcal{E}'''_t \overline{C}(x)\mathsf{rec}(s)^\dagger\, \left|\overline{0}\right\rangle^{\otimes n} |s\rangle\, \left|\overline{A}\right\rangle, && (4.7.13)
\end{aligned}$$

with $\mathcal{E}''_{t_1} = \mathsf{rec}(s)^\dagger \mathsf{rep}(\mathcal{E}'_{t_1})\mathsf{rec}(s)$, and consequently using the proprieties of the local stochastic noise we obtain that $\mathcal{E}'''_t \sim \mathcal{N}(\tau'')$ with $\tau'' = \mathcal{O}\left(\max\left((c'\tau^{-2^d \cdot c''})^{-2^d}, (c'\varrho^{-2c''})^{-2^d}\right)\right)$.

Therefore, we now are able to rewrite the distribution $\mathcal{D}_{\mathcal{EC}_U}(x)$ as follows,

$$\mathcal{D}_{\mathcal{EC}_C}(x) = \langle s, y| \mathcal{E}_t''' \, \overline{C}(x)\mathsf{rec}(s)^\dagger \left|\overline{0}\right\rangle^{\otimes n} |s\rangle \left|\overline{A}\right\rangle . \tag{4.7.14}$$

Subsequently, we have that $\mathcal{E}_t''' = \omega \mathsf{X}(a)\mathsf{Z}(b)$. At this point, we can use condition 4 of Definition 43 to define a noise threshold $q < q_{th}$ based on our local stochastic noise to prove that for $p < p_{th}$, given $p_{th} = \Omega(q_{th}^{\mathcal{O}(1)})$, we have with exponentially high confidence,

$$\mathcal{D}_{\mathcal{EC}_C}(x) = \langle s, y \oplus a| \, \overline{C}(x)\mathsf{rec}(s)^\dagger \left|\overline{0}\right\rangle^{\otimes n} |s\rangle \left|\overline{A}\right\rangle , \tag{4.7.15}$$

where a is the vector that defines the $\mathsf{X}(a)$ operator, derived solely from the Pauli X component of $\mathcal{E}_t'''$. The previous assumption on the error threshold then implies that after the error correction process we obtain with exponential high confidence that $\mathsf{dec}\,(y \oplus a) = \mathsf{dec}(y)$ such that the resulting error corrected distribution is equal to,

$$\mathcal{D}'_{\mathcal{EC}_C}(x) = \langle s, y| \, \overline{C}(x) \, \mathsf{rec}(s)^\dagger \left|\overline{0}\right\rangle^{\otimes n} |s\rangle \left|\overline{A}\right\rangle \tag{4.7.16}$$

$$= \langle s, y \oplus w| \, \overline{C}(x) \left|\overline{0}\right\rangle^{\otimes n} |s\rangle \left|\overline{A}\right\rangle \tag{4.7.17}$$

for w such that $\overline{C}(x)\mathsf{rec}(s)^\dagger\overline{C}(x)^\dagger \left|0\right\rangle^{\otimes n\cdot m} = |w\rangle$.
From the last expression, after applying the function f defined previously—which corrects possible errors introduced during single-shot state preparations and decodes the error-corrected quantum circuit—we find that the outcome of our circuit C can be described by $\langle y| \, C(x) \left|0\right\rangle^{\otimes n} |A\rangle$. Finally, it is important to note that the bounds allow us to achieve this with a probability close to 1. □

We will demonstrate that the surface code over qupits fulfills all the necessary conditions of Definition 43. To begin, we consider the first two conditions. The first one follows trivially, as it is possible to create any quantum state $|A\rangle$ in its logical version within the surface code, while the local stochastic noise can be seen as a layer of errors occurring after the noise-free state preparation. For condition 2, we need to demonstrate that the Clifford operators are realizable in constant depth. Fortunately, this was shown previously for both the qubit and qudit cases in the work of [280].

Lemma 46 ([280])**.** *Any qudit Clifford operators C of prime dimension can be implemented in constant depth with the surface code.*

We turn to proving conditions 3 and 4 in the next two subsections. We first introduce the new decoder under consideration and prove that, when applied to the surface code, it fulfills conditions 3 and 4. Afterwards, we will show that the generalization of single-shot state preparation for qupits, as demonstrated for qubits in [35], holds equally well. We have changed the order of presentation for these two conditions for clarifying the details more coherently.

Qupit error threshold

As remarked in Definition 43, we need to show that the qupit surface code, along with the selected decoder, achieves the required noise resilience properties. In particular, the higher dimensional case requires several changes, the most important of which is the decoder used. Prior work [35] uses the *minimum weight perfect matching* (MWPM) decoder for the qubit case. We now discuss below why the qupit case needs a different decoder and present our choice, the hard decision renormalization group decoder.
A single physical (Pauli) error on the surface code over qubits generates two defects (i.e. X syndromes), appearing on the vertices v_1 and v_2 that share the edge (v_1, v_2) where the error was detected, which

form the boundary of the error. If we have another error on an adjacent edge (v_2, v_3), the defect on the common vertex v_2 cancels out, leaving only the defects on v_1 and v_3. Hence a succession of adjacent errors generates a path on the surface code with defects on its boundary. One method to tackle such errors is to find the most likely one and correct it. An efficient algorithm for this is to find minimum (Pauli) weight operators representing the errors that connect the defects. This problem can be mapped to a weighted graph where the goal is to find a perfect matching with minimum weight. The MWPM in this case can be solved efficiently [294, 295].
This error pattern happens because the Pauli operators in two dimensions (i.e. $p = 2$) are self-inverse, leading to cancellations. However, this is not the case for generalized Pauli operators with $p > 2$; these operators are in general neither self-inverse nor self-adjoint. Hence the intermediate defects do not cancel along a (edge-)path of errors, instead leaving scattered defects throughout the lattice. For $p > 2$, the equivalent problem would be to find the perfect matching of *hyperedges* in a p-uniform hypergraph, a well known NP-hard problem. The MWPM decoder is hence ill-suited to deal with qupits due to efficiency issues.

Hard-Decision Renormalization Group decoder

A more suitable choice for our case is the Hard-Decision Renormalization Group (HDRG) decoder [283]. This decoder operates iteratively. In each iteration or level, it considers all the defects (syndromes) in the lattice and groups them into subsets called clusters. The defects are clustered by comparing their pairwise distances with a threshold distance for each level, known as the search distance. All defects with pairwise distance below the threshold are merged into a cluster. Formally, at the lth level, the search distance is based on a function $D(l)$, chosen to be the Chebyshev distance $D(l) = 2^l$ in [283]. However, other choices and subsequent improvements have used different metrics or combinations thereof, such as the Manhattan distance and a linear search distance $D(l) = l + 1$.
In the first iteration, the decoder considers all the syndrome information and creates the first set of clusters, which then become the inputs for the second iteration. For each cluster at any level, it checks if the sum of the syndrome measurements equals zero modulo p, the dimension of the qupit system. If the sum is zero, the cluster is called neutral, indicating that a Pauli operator exists that can turn all the syndromes trivial (i.e., remove the defects). If the cluster is neutral, a valid correction operator is recorded, and the cluster is removed from the next iteration. If a cluster is charged (not neutral), it is added to the set of clusters to be grouped in the next iteration or level, where the search distance (defining which clusters are merged) will increase, and the process is repeated. If no cluster is left after the last iteration, which is defined by the largest possible value of the search distance that does not exceed the maximum distance between any two lattice points, the decoder returns the composition of all the Pauli operators determined for all levels so far. Otherwise, it returns failure.
Here, for simplicity, we use the original proposal for the HDRG due to [283], but the results can be improved by considering subsequent refinements [296, 297]. We start off with the essential definitions required to describe the operational behaviour of this decoder, which will help us prove the required error threshold.
First, we define the vertices and edges in the 2D lattice as sites, and consider the following distance measure between sites $x = (x_1, x_2)$ and $y = (y_1, y_2)$: $\mathsf{dist}(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$. Suppose the lattice is affected by a Pauli error $\mathcal{E}$. Then we define the diameter of a subset of $\mathcal{E}'$ errors (where the notion of subset means that $\mathsf{Supp}(\mathcal{E}') \subseteq \mathsf{Supp}(\mathcal{E})$) by the maximum pairwise distance between the sites within this subset, $\mathrm{diam}(\mathcal{E}') = \max\{\mathsf{dist}(a, b) \mid a, b \in \mathsf{Supp}(\mathcal{E}')\}$. To understand how the decoder merges error clusters, we define a notion of connectivity for subsets of errors. In particular, a subset of errors $\mathcal{E}'$ is r-connected if there does not exists subsets $A \in \mathsf{Supp}(\mathcal{E}')$ and $B = \mathsf{Supp}(\mathcal{E}' \setminus A)$ such that the distance between $a \in A$ and $b \in B$ is larger than r.
Next, we recursively define a notion of level-k errors in terms of subsets of a Pauli error $\mathcal{E}$ that can occur. These are distinct from the level-k *clusters*; the clusters over which the decoder operates are formed from

the defects generated by $\mathcal{E}$. Level-k errors ultimately determine the syndromes and influence the behavior and efficiency of the decoder, and will therefore be the object of analysis.

Definition 44. *A subset $\mathcal{E}' \subseteq \mathcal{E}$ of an error $\mathcal{E}$ is called a level-k error if it satisfies the following conditions, for a fixed integer $l >> 1$.*

- *It contains at least two disjoint level-$(k-1)$ errors.*
- *The maximum distance between single Pauli errors within $\mathcal{E}'$ is $\frac{l^k}{2}$.*
- *The level-0 errors are single Pauli errors.*

The union of all level-k errors of $\mathcal{E}$ is denoted by $\mathcal{E}_k$.

With this definition of level-k errors, any error $\mathcal{E}$ can be partitioned into a disjoint union of indexed subsets, where each subset consists of lattice sites in the support of $\mathcal{E}$ that appear only in errors of a level equal to the index of that subset. This implies that all subsets are mutually exclusive, denoted as $F_i = \mathcal{E}_i \backslash \mathcal{E}_{i+1}$. We refer to this as the decomposition of the error $\mathcal{E}$ and write

$$\mathcal{E} = \bigsqcup_{i=0}^{h} F_i, \tag{4.7.18}$$

where the length of the decomposition h is closely related to the structure of the error $\mathcal{E}$. This decomposition can uniquely determine whether the decoder will succeed or fail. In particular, a decoder that runs for less than h iterations will abort.

Lemma 47 ([283])**.** *Let $l \geq 10$ and m be the lattice size. For any Pauli error $\mathcal{E}$ with a decomposition of length h that satisfies $l^{h+1} < m$, $\mathcal{E}$ is corrected by the* HDRG *decoder.*

It is also of interest to understand the sizes of and distances between the errors in the last iteration of the decoder, which precisely determine the conditions under which the decoder fails. The following lemma addresses this issue.

Lemma 48 ([283])**.** *Let $l \geq 6$ and $\mathcal{E}'$ be l^k-connected subset of F_k. Then $\mathcal{E}'$ has diameter $\leq l^k$ and is separated from the set $\mathcal{E}_k \setminus \mathcal{E}'$ by distance $> \frac{1}{3} l^{k+1}$.*

Subsequently, considering the local stochastic nature of the errors, we must determine the threshold below which an error occurring according to our noise model has a sufficiently small decomposition with very high probability. Simultaneously, we obtain the error threshold below which the decoder operates with exponential accuracy. The proof follows as in [283], with some additional considerations for the local stochastic noise model that we use in this work.

Lemma 49. *There exists a constant threshold $\tau_{th} > 0$ such that for any $\tau < \tau_{th}$ the* HDRG *decoder over distance-m qupit surface code corrects local stochastic errors $\mathcal{E} \sim \mathcal{N}(\tau)$ with failure probability bounded by*

$$\Pr[\text{Fail}] \leq \exp\left(-\Omega(m^{\eta})\right), \tag{4.7.19}$$

for some constant $\eta > 0$.

Proof. The HDRG decoder continues to iterate as long as the area over which it attempts to merge clusters is still within the 2D lattice of the surface code. Therefore, the decoder fails if the decomposition of the error is longer than $h \approx \log(m)/\log(l)$ as given by Lemma 47. We need to bound the probability p that

an error $\mathcal{E} \sim \mathcal{N}(\tau)$ has a decomposition longer than h. This event implies the existence of at least one h-level error, which causes the decoder to fail.
The 2D lattice that defines our surface code and the associated distance measures imply that the HDRG decoder will consider squares over the lattice as the site-enclosing objects, to analyze k-level errors. For a level-h error to exist, it must overlap with a square $\square_h$ of side length l^h, which can be considered independently at any position on the 2D lattice, as the following properties are translation-invariant. Furthermore, from Lemma 48, we obtain that a larger square $\square'_{h+1}$ of size $9l^{h+1}$ centered on a box $\square_l$ must fully contain a level-k error if it overlaps with $\square_h$. This allows us to decompose the probability of failure as follows

$$\begin{aligned} \mathsf{Pr}[\text{HDRG fails}] &\leq \mathsf{Pr}[\mathcal{E}_h \in \square'] && (4.7.20)\\ &\leq \mathsf{Pr}\big[(\mathcal{E}_{h-1} \sqcup \mathcal{E}'_{h-1}) \in \square'\big] && (4.7.21)\end{aligned}$$

Since the failure probability of the HDRG decoder is bounded by the probability of $\square'$ containing a level-h error, we can bound the same probability in terms of the probability of the same square containing two disjoint level-$(h-1)$ errors by the definition of a level-h error.
Finally, the probability of having two disjoint level-$(h-1)$ errors in the same square can be further bounded. The events whose probabilities we are bounding here are known as *increasing events*, as defined in percolation theory. In simple terms, increasing events are those events that remain true when the state space leading to the event is extended or augmented. For instance, rain occurs whenever a certain threshold of water falls from the sky (precipitation rate), and it remains a true event if the precipitation rate increases and it rains more intensively. This implies that these events occur under varying conditions, as in our case. Most importantly, with that, we can apply the van den Berg and Kesten inequality [298] to bind the probability as follows,

$$\begin{aligned} \mathsf{Pr}[\mathcal{E}_{h-1} \in \square' \sqcup \mathcal{E}_{h-1} \in \square'] &\leq \mathsf{Pr}[\mathcal{E}_{h-1} \in \square'] \cdot \mathsf{Pr}[\mathcal{E}_{h-1} \in \square'] && (4.7.22)\\ &\leq \mathsf{Pr}[\mathcal{E}_{h-1} \in \square']^2. && (4.7.23)\end{aligned}$$

Subsequently, we have that there are $(3l)^2$ squares $\square''$ of size l^{h-1} in $\square'$, such that with a simple union bound we can write

$$\mathsf{Pr}[\mathcal{E}_{h-1} \in \square'] \leq (3l)^2 \cdot \mathsf{Pr}[\mathcal{E}_{h-1} \in \square'']. \tag{4.7.24}$$

This allows us to write the probability of the failure recursively in terms of smaller squares until we reach a single site, which coincides with the probability that a single qupit is affected by noise; this in turn is definitionally equal to τ for local stochastic errors[10]. Hence we have that

$$\begin{aligned} \mathsf{Pr}[\text{HDRG fails}] &\leq (3l)^{-4}(3l^4\tau)^{2^h} && (4.7.25)\\ &= (3l^4\tau)^{\Omega(m^\eta)}, && (4.7.26)\end{aligned}$$

for $\eta > 0$ using again that $h \approx \log(m)/\log(l)$. We thus obtain the stated asymptotic bound on the failure probability of the HDRG decoder. Conversely, it decodes correctly with high probability as required. □

The last element needed to prove the error threshold from our conditions in Definition 43 is to define the decoding function dec in terms of the string that the HDRG decoder outputs.

Corollary 10. *There is a decoding function* dec *based on the* HDRG *decoder that for a distance-m qupit*

[10] Notice that under this reduction, the HDRG decoder has the same error correction success probability for the local stochastic error model considered in this text as for the error model considered in [283].

surface code subjected to local stochastic noise $\mathcal{E} \sim \mathcal{N}(\tau)$ with $\tau < \tau_{th}$ satisfies

$$\mathsf{dec}(x) = \mathsf{Mod}_p(x) \tag{4.7.27}$$

and

$$\Pr[\mathsf{dec}(x \oplus v) = \mathsf{Mod}_p(x)] \geq 1 - e^{-\Omega(m^\eta)}, \tag{4.7.28}$$

for every $v \in \mathbb{F}_p^m$ defined by $|v\rangle = \mathcal{E}_X |0\rangle^{\otimes m}$, where $\mathcal{E}_X$ denotes the restriction of the error $\mathcal{E}$ to those qupits where it acts as a tensor of Pauli X operators.

Proof. Recall that the logical measurement operation of a logical qupit in the surface code is performed by single qupit Z measurements of the physical qupits along a diagonal of the lattice. The value of the logical measurement result is determined by the sum of the physical measurement results modulo p, with p being the dimension of the qupit system. In particular, in the noise-free case when $\tau = 0$, the result of this operation effectively produces the correct value for the logical qupit.
For the noisy case, we can model the error that occurs just before the measurement operation as $\mathcal{E} = \omega^i \mathsf{X}(a)\mathsf{Z}(b)$. This error introduces deviations to the measurement results through the $X(a)$ component. For convenience, we represent these deviations from the noise-free measurement results with an additional string $v \in \mathbb{F}_p^n$, which follows from the $\mathsf{X}(a)$ component. We then use the fact that the HDRG decoder, with exponentially high probability, can guess the string v from the error syndromes based on Lemma 49. Therefore, if we consider the decoding function dec as the modular remainder of the measured string combined with the outcome of the HDRG decoder v_{Guess}, we have, with exponential confidence, that

$$\mathsf{Mod}_p(x \oplus v \oplus v_{Guess}) = \mathsf{Mod}_p(x). \tag{4.7.29}$$

This finishes the proof of the corollary establishing the exponential confidence for our choice of the decoding function. □

With this proof of Corollary 10 we have completed the proof of the condition which was required to demonstrate the effectiveness of our selected decoder for the qupit surface code. We now proceed to the next crucial ingredient in our noise-resilient construction, single-shot state preparation.

Qupit single-shot state preparation

In this section, we will demonstrate how to achieve single-shot state preparation of the GHZ_2 state for qupit surface codes. Considering adaptive state preparation processes, we will outline the predetermined state preparation operations and the adaptive correction mechanisms. Specifically, the latter involves a designated recovery function used to correct errors from the fixed state preparation operations in creating our logical resource state and a repair function that addresses additional noise that may occur in the preceding circuit.
We then construct a 3D cubic lattice into which the logical GHZ_2 state will be encoded. To this end, we define all the required stabilizer generators on the 3D lattice and determine the measurement pattern that allows us to create our resource state. The additional Pauli operators that will be required, controlled by the measurement outcomes, define our recovery function. Lastly, we prove that there exists a valid repair function that takes as input any local stochastic error affecting single-shot state preparation and produces as output a global correction in the form of another local stochastic error. Recall that this type of correction implies that our decoder can recover the encoded information with a higher likelihood. This is achieved by demonstrating that the selected repair function satisfies the lifting property, as defined in [35]. Similar to the qubit case, it ensures the correct behavior of the repair function across all qupit cases. By combining these two demonstrations, we achieve a single-shot state preparation that fulfills the condition specified in Definition 43 required for fault-tolerant executions of shallow-depth quantum circuits.

Qupit recovery and repair functions

The construction of the resource state needed to demonstrate an advantage in the qupit case involves two main stages, each comprising several steps. In the first stage, a constant-depth Clifford circuit $\mathcal{Z}$ is applied to $2m$ physical qupits and m_{a} ancilla qupits. The former qupits will be designated for the two logical qupits in the $\left|\overline{\mathsf{GHZ}_2}\right\rangle$ state, while the latter are uniquely employed in the state preparation process. Specifically, the correction applied to the state after the measurement process in $\mathcal{Z}$ is described as the recovery operation $\mathsf{rec}(s)$, where s is a string resulting from measuring the m_{a} ancilla qupits. More precisely, after executing the constant-depth circuit $\mathcal{Z}$ and the recovery operation $\mathsf{rec}(s)$, we obtain the following state,

$$\left(|s\rangle\langle s| \otimes \mathsf{rec}(s)\right) \mathcal{Z} \left(|0\rangle^{\otimes 2m} \otimes |0\rangle^{\otimes m_{\mathsf{a}}}\right) \propto \left|\overline{\mathsf{GHZ}_2}\right\rangle \otimes |s\rangle \,. \tag{4.7.30}$$

However, when this process is subject to local stochastic errors, it changes the LHS of the above equation to

$$\left(|s\rangle\langle s| \otimes \mathsf{rec}(s)\right) \mathcal{E}\mathcal{Z} \left(|0\rangle^{\otimes 2m} \otimes |0\rangle^{\otimes m_{\mathsf{a}}}\right), \tag{4.7.31}$$

where $\mathcal{E}$ accounts for all the errors occurring in the circuit $\mathcal{Z}$. This is possible because $\mathcal{Z}$ is a constant-depth Clifford circuit, and we can commute the individual errors occurring in each layer to the end (using again Lemma 24) . However, this error will be carried to the output state; hence, in the presence of noise, Equation (4.7.30) would not be valid. To obtain the valid expression we need to resort to an additional degree of freedom that it will correct for noise, this will be expressed in the form of a repair function $\mathsf{rep}(\mathcal{E})$, and the updated condition for the single-shot state preparation,

$$\left(|s\rangle\langle s| \otimes \mathsf{rec}(s)\right) \mathcal{E}\mathcal{Z} \left(|0\rangle^{\otimes 2m} \otimes |0\rangle^{\otimes m_{\mathsf{a}}}\right) \propto \mathsf{rep}(\mathcal{E}) \left|\overline{\mathsf{GHZ}_2}\right\rangle \otimes |s\rangle \,. \tag{4.7.32}$$

Subsequently, we will describe these steps in more detail, showing how both the recovery and repair functions are going to be defined for the higher dimensional case. The arguments will follow similarly to those in the qubit cases in [35, 124].

Our process will utilize a 3D lattice, employing m qupits on two opposing faces for the logical qupits of the GHZ_2 state, with all m_a qupits in the bulk of this lattice serving as ancilla qupits. Specifically, we will define stabilizers over this lattice as $\mathcal{S}_0$ whenever they are uniquely associated with the ancilla qupits and $\mathcal{S}_1$ when they are defined over the entire lattice. These will be detailed further in the next subsection. However, to provide a high-level description, consider that elements of $\mathcal{S}_0$ are the qupits measured during the process of creating the $\left|\overline{\mathsf{GHZ}_2}\right\rangle$. Moreover, S_0^i will correspond to a trivial stabilizer of the $\left|\overline{\mathsf{GHZ}_2}\right\rangle$ state. Consequently, the syndrome of the error $\mathcal{E}$ associated with it will only depend on s. We will define this as a string $\mathsf{syn}_0(\mathcal{E}) \in \mathbb{F}_d^k$, where k represents the number of $\mathcal{S}_0$ generators. The syndrome elements will thus depend on the commutation between the stabilizer elements of $\mathcal{S}_0$ and $\mathcal{E}$. Since we are using qupits, this relationship can be defined through the Mod function, analogous to the parity function in the qubit case,

$$\mathsf{syn}_0(\mathcal{E})_i = \mathsf{Mod}(s, \mathsf{Supp}(S_\alpha^i)) = \sum_{u \in \mathsf{Supp}(S_\alpha^i)} s_u \ \mathsf{Mod}\ p, \tag{4.7.33}$$

with the index i used to describe the element of the stabilizer i.e one of $S_0^1, \ldots, S_0^k$ and the lower index α describes which subgroup of stabilizers we are referring to.

In contrast, elements of $\mathcal{S}_1$, will act with Z type Paulis on the ancilla qupits, but they will also act with both Z and X type Paulis on the qupits that will encode our resource state. Furthermore, all elements $\mathcal{S}_1^i$ will stabilize $\mathcal{Z}\,|0\rangle^{\otimes 2m} \otimes |0\rangle^{\otimes m_{\mathsf{a}}}$ such that we have,

$$\omega^{b(s)} \left(|s\rangle\langle s| \otimes \mathbb{1}_B\right) \mathcal{Z} \left(|0\rangle^{\otimes 2m} \otimes |0\rangle^{\otimes m_{\mathsf{a}}}\right), \tag{4.7.34}$$

with $b(s)$ depending only on the modular sum of the bulk measurements for each stabilizer. Therefore,

from s alone, one can infer the error that appears on the facets of the 3D structure with the logical qupits. If not for the existence of extra errors, one could easily obtain the intended logical state from the measurements of the auxiliary qupits. However, the actual phase that appears will be

$$\omega^{b(s)\oplus \mathsf{syn}_1(\mathcal{E})}\left(|s\rangle\langle s|\otimes \mathbb{1}_B\right))\,\mathcal{Z}\left(|0\rangle^{2m}\otimes|0\rangle^{m_a}\right), \tag{4.7.35}$$

with $\mathsf{syn}_1(\mathcal{E})$ accounting for the additional error that we cannot infer from s alone, but detected trough the stabilisers in the set $\mathcal{S}_1$. From here on, for simplicity, we will represent the phases by their powers alone. Additionally, in the ideal implementation, 0 will be considered the correct phase. Therefore, the phase accumulated, represented as $b(s)\oplus\mathsf{syn}_1(\mathcal{E})$, requires a correction, c, such that $c\oplus b(s)\oplus\mathsf{syn}_1(\mathcal{E})=0$. Unfortunately, the syndrome of the error on the surface cannot be determined from the measurement outcomes performed on the bulk (we do not have access to the values of the stabilizers S_1^i). This problem is addressed in [35] by defining an appropriate proxy P for the error, chosen to be a function of the syndrome measurements, $P=f(s)$, in such a way that it will have the same syndrome as the error on the bulk. Then, since the effects of the measurements on the bulk, which generate additional phases on the surfaces and $b(s)$, are known, one can consider the following correction $c\oplus b(s)\oplus\mathsf{syn}_1(P)=0$. Recognising c as the recovery function with syndrome $\mathsf{syn}_1(\mathsf{rec}(s))$ over the set $\mathcal{S}_1$ and using the remaining degrees of freedom that we have, in the form of a repair function, we correct for the difference between the recovery employed, based on the proxy, and the actual error present on the surface, such that

$$\mathsf{syn}_1(\mathsf{rec}(s))\oplus b(s)\oplus\mathsf{syn}_1(\mathcal{E})\oplus\mathsf{syn}_1(\mathsf{rep}(\mathcal{E}))=0, \tag{4.7.36}$$

leading us to a recovery of the form $\mathsf{syn}_1(\mathsf{rec}(s))=-b(s)\oplus\mathsf{syn}_1(P)$, and a repair function of the form $\mathsf{syn}_1(\mathsf{rep}(\mathcal{E}))=-\mathsf{syn}_1(P)\oplus-\mathsf{syn}_1(\mathcal{E})$.
This concludes the high-level description of the repair and recovery functions. Subsequently, we will show how this high-level behavior is concretized in our 3D constructions and demonstrate that these functions fulfill the required conditions from Definition 43 with their precise definitions.

Logical GHZ_2 states from the 3D block construction

In this section, we define the lattice used for single-shot state preparation of the logical GHZ_2 state. Our lattice retains the general structure described in [124] but replaces qubits with qupits at each site. Due to the asymmetries observed in the qupit surface code for $p>2$, which influence how plaquette and vertex stabilizers must be oriented (see Figure 16), we must carefully incorporate these asymmetries into the design of the constant-depth quantum circuit $\mathcal{Z}$. Specifically, we ensure the proper alignment of each stabilizer generator in the graph states produced by the quantum circuit. Moreover, we will meticulously verify the geometric placement of each stabilizer generator, ensuring that a valid measurement pattern is achieved for the preparation of a logical GHZ_2 state.
The three-dimensional lattice $\mathcal{C}$, used for the creation of the logical GHZ_2 states with a qupit surface code of distance d, has a side length of $r=2d+1$ and is defined as follows,

$$\mathcal{C}=\{(u_1,u_2,u_3):0\leq u_1,u_2\leq r-1,\,1\leq u_3\leq r\}. \tag{4.7.37}$$

In particular, the two opposite surfaces, corresponding to $u_3=1$ and $u_3=r$, are defined such that their first two coordinates u_1 and u_2 have different parities. We will refer to this set of qupits as the surface

$$\mathsf{Sur}=\{(e,o,u_3),(o,e,u_3)\in\mathcal{C}:u_3\in\{1,r\}\}^{11}. \tag{4.7.38}$$

When we use this lattice $\mathcal{C}$ to build the noise-resilient circuit, qupits that lie on Sur will encode the logical

[11] Throughout this section, we use the symbols o and e inside set notation to mean variables that range over all arbitrary odd and even indices respectively

GHZ_2 state, while we regard all other qupits as ancillary, contained in the *bulk* of this 3D lattice, defined by $\mathsf{Bulk} = \mathcal{C} \backslash \mathsf{Sur}$.
Similarly to the initial proposal in [124] for qubits, we will utilize four graphs—the odd graph, even graph, surface code graph, and dual surface code graph, respectively labeled T_o, T_e, T_{sc}, and T^*_{sc}. These graphs help us describe the measurement pattern required for preparing the logical GHZ_2 state. More specifically, the set of vertices of these graphs will help us in defining the stabilizers of $\mathcal{S}_0$ and $\mathcal{S}_1$ and are given by

$$\begin{aligned} V(T_e) &= \{(e,e,e) \in \mathcal{C}\}, & V(T_o) &= \{u = (o,o,o) \in \mathcal{C} : u_3 \notin \{1,r\}\}, \\ V(T_{sc}) &= \{u = (e,e,o) \in \mathcal{C} : u_3 \in \{1,r\}\}, & V(T^*_{sc}) &= \{u = (o,o,o) \in \mathcal{C} : u_3 \in \{1,r\}\}, \end{aligned}$$

while the edge sets of T_o and T_e will aid in defining the stabilizers of the resulting logical state, which are given by

$$\begin{aligned} E(T_e) &= \{(e,e,o),(e,o,e),(o,e,e) \in \mathcal{C}\}, \\ E(T_o) &= \{u = (o,o,e),(o,e,o),(e,o,o) \in \mathcal{C} : u_3 \notin \{1,r\}\}. \end{aligned}$$

Now, to address the previously mentioned asymmetry and achieve the intended logical GHZ_2 states, we define the constant-depth quantum circuit $\mathcal{Z}$ over the qupits on the lattice $\mathcal{C}$. For simplicity, we will divide the circuit into operations over the 2D layers running along u_1 and u_2, for even and odd values of u_3, designated as $\mathcal{Z}_{\mathsf{even}}$ and $\mathcal{Z}_{\mathsf{odd}}$, respectively. Additionally, as we will discuss, we will require stabilizers with four different directionalities. To keep track, we define the following function $dir(u) = (-u_1 + u_2 + u_3) \bmod 4$, which will help us position each of these four elements. Having considered all these details about the circuit, we obtain the following definitions,

$$\mathcal{Z}_{\mathsf{even}} = \left(\prod_{dir(u)=\{0,1\}} \mathsf{X}_u^\dagger \right) \mathsf{F}_u \left(\prod_{v \in n(u), dir(u)=\{2,3\}} \mathsf{CZ}_v^{-(v_1-u_1+v_2-u_2+v_3-u_3)} \right) \tag{4.7.39}$$

$$\left(\prod_{v \in n(u), dir=\{0,1\}} \mathsf{CZ}_v^{-(v_1-u_1-(v_2-u_2)+v_3-u_3)} \right) \prod_{u \in E(T_o) \cup E(T_e) \backslash \mathsf{Sur}} \mathsf{F}^{\otimes n} \tag{4.7.40}$$

forall $u = (\cdot,\cdot,e) \in \mathcal{C}$ with the nearest neighbors of each site u given by $n(u) = \{v \in \mathcal{C} : |u_1 - v_1| + |u_2 - v_2| + |u_3 - v_3| = 1\}$, and respectively

$$\mathcal{Z}_{\mathsf{odd}} = \left(\prod_{dir(u)=\{2,3\}} \mathsf{X}_u^\dagger \right) \mathsf{F}_u \left(\prod_{v \in n(u), dir(u)=\{0,1\}} \mathsf{CZ}_v^{-(v_1-u_1+v_2-u_2+v_3-u_3)} \right) \tag{4.7.41}$$

$$\left(\prod_{v \in n(u), dir=\{2,3\}} \mathsf{CZ}_v^{-(v_1-u_1-(v_2-u_2)+v_3-u_3)} \right) \prod_{u \in E(T_o) \cup E(T_e) \backslash \mathsf{Sur}} \mathsf{F}^{\otimes n} \tag{4.7.42}$$

forall $u = (\cdot,\cdot,o) \in \mathcal{C}$.
Then, the state $\mathcal{Z} |0\rangle^{\otimes |\mathcal{C}|}$, with $\mathcal{Z} = \mathcal{Z}_{\mathsf{even}} \mathcal{Z}_{\mathsf{odd}}$, resulting from our constant-depth circuit, will be measured in the Z basis on all the qupits of the lattice $\mathcal{C}$ such that $u \notin E(T_o) \cup E(T_e) \setminus \mathsf{Sur}$. These sites can be directly removed from our analysis for simplicity, as such measurements over graph states correspond to the simple removal of the vertex from the graph. Therefore, for the remaining qupits, the stabilizers of

the state K_u are transformed with the Fourier gate to the X basis, resulting in the stabilizers,

$$= S_0^a = \mathsf{Z}_u \left(\prod_{v \in n(u)} \mathsf{X}_v^{(v_1 - u_1 + v_2 - u_2 + v_3 - u_3)} \right) \text{ for } u = (\cdot, e), dir(u) = 3 \text{ or } u = (\cdot, o), dir(u) = 0; \tag{4.7.43}$$

$$= S_0^b = \mathsf{Z}_u \left(\prod_{v \in n(u)} \mathsf{X}_v^{(v_1 - u_1 - (v_2 - u_2) + v_3 - u_3)} \right) \quad \dots \quad dir(u) = 2 \quad \dots \quad dir(u) = 1; \tag{4.7.44}$$

$$= (S_0^a)^\dagger \quad \dots \quad dir(u) = 1 \quad \dots \quad dir(u) = 2; \tag{4.7.45}$$

$$= \left(S_0^b\right)^\dagger \quad \dots \quad dir(u) = 0 \quad \dots \quad dir(u) = 3; \tag{4.7.46}$$

with all the vertex are of the type $u \in E(T_o) \cup E(T_e) \setminus$ Sur. See Figure 25 for a spacial representation of the placement of the stabilisers over $\mathcal{C}$)

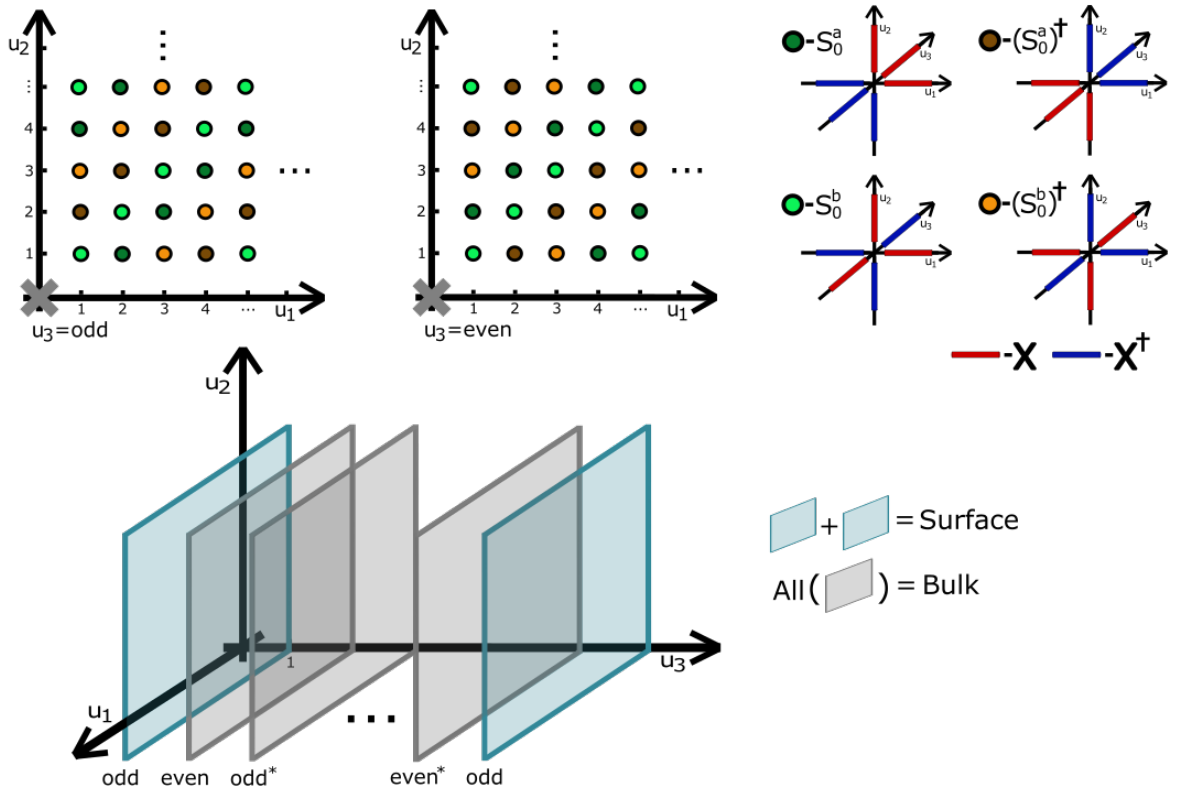


Figure 25: On the top left, we have a representation of the distribution of the stabilizer generators over layers (planes) parallel to the u_1 and u_2 axes, corresponding to even and odd values of u_3 in the lattice $\mathcal{C}$. The layers termed "even*" and "odd*" are equivalent, having the same pattern defined by the distribution of stabilizers, to the "even" and "odd" layers, respectively, except for the translation by two units in the same plane formed by u_1 and u_2. The cross (×) at the origin of the 2D graphs indicates that the u_3 axis is inward-directed. The top right panel displays a representation of the stabilizer generators K_u from Equation (4.7.43) with their respective spatial orientation. At the bottom, we have a depiction of the entire lattice $\mathcal{C}$, indicating the layers and defining stabilizer positions after the application of the constant-depth circuit $\mathcal{Z}$ to the lattice. Note that the first and last layers constitute the surface (Sur), where the logical qupits reside, while the remaining middle layers define the bulk (Bulk).

Having defined the stabilizer group $\mathcal{S}$ over the remaining qupits of the lattice $\mathcal{C}$, which arises from the generators K_u and the graphs T_o, T_e, T_{sc}, and T^*_{sc}, we can now specify the two special subgroups of this stabilizer group, $\mathcal{S}_0 < \mathcal{S}_1 < \mathcal{S}$. As discussed in the definition of the recovery and repair functions in the previous subsection, these subgroups facilitate the retrieval of the intended outcome state following the measurements of the ancillary qupits. More concretely, within our lattice $\mathcal{C}$, elements of $\mathcal{S}_0$ are designed to act trivially on Sur and to affect Bulk solely through Z operations. In contrast, elements of $\mathcal{S}_1$ perform X or Z operations on Sur and only Z operations on Bulk. These elements are specifically required to be stabilizers of the logical GHZ_2 state. Additionally, we intend to have a subset of the stabilizers of $\mathcal{S}_1$, which function as plaquette and vertex stabilizers of the surface qudit code.
Starting with the generators of $\mathcal{S}_0$, which are associated with the qupits corresponding to vertices of T_o and T_e, we can describe these based on our stabilizers referenced in Equation (4.7.43) as follows,

$$S_0^u = \prod_{v \in n(u)} K_v = \prod_{v \in n(u)} \mathsf{Z}_v^{(v_1 - u_1 + v_2 - u_2 - (v_3 + u_3))}, \text{ with } u \in V(T_e), \tag{4.7.47}$$

and

$$S_0^u = \prod_{v \in n(u)} K_v = \prod_{v \in n(u)} \mathsf{Z}_v^{(v_1 - u_1 - (v_2 - u_2) + v_3 - u_3)}, \text{ with } u \in V(T_o). \tag{4.7.48}$$

In both cases, all the X operators from the K_v stabilizers overlap with some other $\mathsf{X}^\dagger$ operators within the same stabilizer set and completely cancel out. Thus, we obtain exclusively Z operators for this stabilizer subgroup, as intended.
The second subgroup, $\mathcal{S}_1$, utilizes the graphs T_{sc} and T^*_{sc} to define two types of stabilizers; one functions as the vertex stabilizer and the other as the plaquette stabilizer of the qudit surface code, as we will see subsequently demonstrate.

$$S_1^u = K_u = \prod_{v \in n(u)} \mathsf{Z}_u^\dagger \left(\prod_{v \in n(u)} \mathsf{X}_v^{-(v_1 - u_1 - (v_2 - u_2) + v_3 - u_3)} \right), \text{ with } dir(u) = 1, u \in V(T_{sc}), \tag{4.7.49}$$

$$\prod_{v \in n(u)} \mathsf{Z}_u \left(\prod_{v \in n(u)} \mathsf{X}_v^{(v_1 - u_1 - (v_2 - u_2) + v_3 - u_3)} \right), \text{ with } dir(u) = 3, u \in V(T_{sc}). \tag{4.7.50}$$

These operators act exclusively on the qupits in the planes of the logical qupits where $u_3 = 1, r$, since operators in adjacent layers affect qupits that have already been measured out. Additionally, the vertices $u \in V(T_{sc})$ are not located on Sur, the 2D lattice containing the logical qupit. Consequently, these aforementioned operators precisely function as the vertex stabilizer operators within the qupit surface code for any logical states on Sur.
Subsequently, we will define the subgroup of $\mathcal{S}_1$ generating plaquette operators stabilizers,

$$S_1^u = K_{u \pm (0,0,1)} \prod_{v \in (n(u) \cap \mathsf{Sur})} K_v = \prod_{v \in n(u)} \mathsf{Z}_v^{-(-(v_1 - u_1) + v_2 - u_2 + v_3 - u_3)}, \text{ with } dir(u) = 1, u \in V(T^*_{sc}), \tag{4.7.51}$$

$$\prod_{v \in n(u)} \mathsf{Z}_v^{(-(v_1 - u_1) + v_2 - u_2 + v_3 - u_3)}, \text{ with } dir(u) = 3, u \in V(T^*_{sc}). \tag{4.7.52}$$

This equality follows because all the X operators for the selected elementary stabilizer generators K_u cancel each other out completely, leaving us with only the previously mentioned set of $\mathcal{Z}$ operators. In particular, these two stabilizers, each operating on half of the vertices defined by T^*_{sc}, generate the precisely the plaquette operators of the qupit surface code when restricted to Sur, with the correct orientations.
Finally, we have defined all the state-independent stabilizers for the logical qupits on the surface Sur. However, we still need to identify the final one for these states to determine the specific logical stabilizer state they represent. Therefore, given our goal to obtain the logical GHZ_2 state, ideally, the logical stabilizers $\overline{\mathsf{X}^\dagger} \otimes \overline{\mathsf{X}^\dagger}$ and $\overline{\mathsf{Z}} \otimes \overline{\mathsf{Z}^\dagger}$ would be generated from the graph state stabilizers in $\mathcal{S}_1$. Furthermore, the previous logical operators are defined over the lattice as follows,

$$\overline{\mathsf{X}^\dagger} \otimes \overline{\mathsf{X}^\dagger} = \prod_{u=(1,e,1)\in \mathsf{Sur}} \mathsf{X}^\dagger_u \prod_{u=(1,e,r)\in \mathsf{Sur}} \mathsf{X}^\dagger_u \text{ and } \overline{\mathsf{Z}} \otimes \overline{\mathsf{Z}^\dagger} = \prod_{u=(0,o,1)\in \mathsf{Sur}} \mathsf{Z}_u \prod_{u=(0,o,r)\in \mathsf{Sur}} \mathsf{Z}^\dagger_u. \quad (4.7.53)$$

We will demonstrate that our resulting logical state, up to some local Clifford operations, is equivalent to the intended GHZ_2 state. To begin, we obtain that for an even-dimensional code distance d, the stabilizer S_1^{X} is defined as follows,

$$S_1^{\mathsf{X}^\dagger} = \prod_{(u_1,u_1-1,e)\in\mathcal{C}} K_u \quad (4.7.54)$$

$$= \prod_{u=(u_1,u_1-1,1)\in\mathcal{C}} \mathsf{X}^\dagger_u \prod_{(u_1,u_1-1,r)\in\mathcal{C}} \mathsf{X}^\dagger_u \prod_{(u_1,u_1-1,e)\in\mathcal{C},dir(u)=3} \mathsf{Z}_u \prod_{(u_1,u_1-1,e)\in\mathcal{C},dir(u)=1} \mathsf{Z}^\dagger_u \quad (4.7.55)$$

$$\equiv \left(\overline{\mathsf{X}^\dagger} \otimes \overline{\mathsf{X}^\dagger}\right)_{\mathsf{Sur}} \prod_{(u_1,u_1-1,e)\in\mathcal{C},dir(u)=3} \mathsf{Z}_u \prod_{(u_1,u_1-1,e)\in\mathcal{C},dir(u)=1} \mathsf{Z}^\dagger_u, \quad (4.7.56)$$

with $(\cdot)_{\mathsf{Sur}}$ designating that this operator is applied on the qupits of the surface Sur, which encode our logical qupits. In particular, this ensures that the logical states we generate through our single-shot state preparation $|\phi\rangle$ are transformed by the application of the identified stabilizers as follows $\overline{\mathsf{X}^\dagger} \otimes \overline{\mathsf{X}^\dagger}$ to produce $|\phi\rangle = \lambda_{\mathsf{X}^\dagger\mathsf{X}^\dagger} |\phi\rangle$.
Similarly, we will also define another stabilizer for the resulting state, $S_1^{\mathsf{Z}^\dagger}$, as described below,

$$S_1^{\mathsf{Z}^\dagger} = \prod_{(u_1,u_1-1,o)\in\mathcal{C}} K_u = \prod_{u=(u_1,u_1-1,1)\in\mathcal{C}} \mathsf{Z}^\dagger_u \prod_{(u_1,u_1-1,r)\in\mathcal{C}} \mathsf{Z}^\dagger_u \equiv \left(\overline{\mathsf{Z}^\dagger} \otimes \overline{\mathsf{Z}^\dagger}\right)_{\mathsf{Sur}}. \quad (4.7.57)$$

This second stabilizer ensures that our state satisfies $\overline{\mathsf{Z}^\dagger} \otimes \overline{\mathsf{Z}^\dagger} |\phi\rangle = \lambda_{\mathsf{Z}^\dagger\mathsf{Z}^\dagger} |\phi\rangle$. Consequently, following a correction based on the measurement outcomes from elements in Bulk, the resultant state $|\phi'\rangle$ will, up to some known logical Clifford operations, be equivalent to the GHZ_2 state for two qupits, represented as $\left|\overline{\mathsf{GHZ}}_2\right\rangle = \frac{1}{\sqrt{p}} \sum_{i=0}^{p-1} \left|\overline{i}\overline{i}\right\rangle$.
We note that one could also consider the case where the code distance d is of odd dimension; however, the stabilizers would become slightly more complex. In addition, unlike the previous solutions and the qubit setting, the qupit surface code will exhibit directionality. Consequently, the copy of the surface code at $u_3 = r$ will serve as a mirror image of the surface code at $u_3 = 1$.

Properties of the Repair function

We will now show that the repair function continues to be manageable even in the higher dimensional case; in particular, it continues to output a local stochastic correction. For a local stochastic error $\mathcal{E} =$

$\mathcal{E}_{\text{Bulk}} \otimes \mathcal{E}_{\text{Sur}} \sim \mathcal{N}(\tau)$ with $\tau \leq 1$ acting on our lattice $\mathcal{C}$, this condition translates to

$$\Pr_{\mathcal{E}}[V \subseteq \text{Supp}(\text{rep}(\mathcal{E}_{\text{Bulk}}))] \leq \tau^{|V|}. \tag{4.7.58}$$

This can be inherited from the analysis of the qubit cases presented in [35]. However, we must define our recovery and repair functions more concretely. Unlike previous works that use the MWPM decoder, these functions will be based on the HDRG decoder. For example, the recovery function is defined as the Pauli correction operator produced by the HDRG decoder, parameterised on a fixed distance value k, which uses as input the syndromes generated by our proxy P and the additional Pauli corrections from the bulk measurements $b(s)$. Therefore, the repair function will be the Pauli operator derived from the difference between the two Pauli operators obtained from the HDRG, corresponding to the concrete error $\mathcal{E}_{\text{Bulk}}$ and our guessed proxy P,

$$\text{syn}_1(\text{rep}(\mathcal{E}_{\text{Bulk}})) = \text{syn}_1(\mathcal{E}_{\text{Bulk}}) \oplus \text{syn}_1(P). \tag{4.7.59}$$

Now, more importantly, we need to demonstrate that under this precise definition, the repair function retains the property expressed in Equation (4.7.58). To achieve this, we can leverage two properties, collectively introduced as the *lifting property* in [35], to substantiate that Equation (4.7.58) holds for our p-dimensional repair function. In particular, there must be a lifting function $f : V_K \to V_L \subseteq 2^{\mathcal{C}}$ that maps a subset V_K of the surface Sur to a set of subsets within $\mathcal{C}$. Then, using these functions, the first property stipulates that for each $V_K \subseteq \text{Supp}(\text{rep}(\mathcal{E}))$ there exists a set $V_{\mathcal{E}'} \in f(V_K)$ such that

$$|V_{\mathcal{E}'} \cap \text{Supp}(\mathcal{E})| \geq \alpha |V_{\mathcal{E}'}|. \tag{4.7.60}$$

Secondly, for every $\lambda > 0$, there must be a corresponding $\lambda_1 > 0$ such that

$$\sum_{V_L \in f(V_K)} \lambda^{|V_L|} \leq \lambda_1^{|V_K|}. \tag{4.7.61}$$

Previously, these properties were demonstrated for the qubit case using a repair and lifting function based on the MWPM decoder. Next, we will demonstrate that this property applies to a repair and lifting function constructed using the HDRG decoder. The lifting function based on the MWPM decoder was defined to output an edge-disjoint union of trees (also called a forest) with boundaries at the endpoints of a subset of the Bulk. In this paper, we define the lifting function to output the set of all subsets V_L, which are r-connected components that intersect V_K at each connected component. Recall that these connected components, according to our definition of r-connectivity, are subsets of the error at a distance bounded by r between each other.

Now for the first condition in Equation (4.7.60), consider that errors $\mathcal{E}_{\text{Bulk}}$ on the Bulk that anti-commute with the stabilisers of $\mathcal{S}_1$ and therefore, activate an S_1^i syndrome, are localised single errors on one of the vertexes of $V(T_{sc})$ or $V(T_{sc*})$. This follows from the fact that these stabilisers are related to single Z stabilisers when we are considering unique errors on the Bulk and, therefore, do not cause any degeneracy. Furthermore, we will observe that our repair function will have an error on the surface Sur, localised on one of the stabilisers S_1^i, to correct an error of our proxy P that generates a different syndrome from the error $\mathcal{E}_{\text{Bulk}}$ on the same vertex.

We consider the non-trivial case where $\text{syn}_1(\mathcal{E}_{\text{Bulk}})$ is a non-zero value for some stabilizer S_1^i, indicating the presence of an error at the corresponding vertex associated with S_1^i. If, in this situation, the Pauli operator resulting from our repair function, defined on the sites V_K, intersects the support of S_1^i, we can conclude that the lift of V_K certainly includes an error set with the connected component containing the error on the vertices on the bulk defined by S_1^i. Moreover, V_K is defined exclusively on the surface Sur and includes only the sites where the previous case repeats with the S_1^i type stabilizers. Specifically, the Pauli operators on the sites of V_K address errors at vertices in the Bulk, which are signaled by the syndromes

of S_1^i and were not accounted for by our proxy P. Therefore, the lift of V_K generates certainly a set of r-connected components of the error $\mathcal{E}_{\text{Bulk}}$. This and the fact that the size of the r-connected components is finite implies that the size of the set V_K correlates with the number of r-connected components of $\mathcal{E}_{\text{Bulk}}$, ensuring that the overlap between one of the elements in lift and the original error $\mathcal{E}_{\text{Bulk}}$ is proportional, thereby fulfilling the condition in Equation (4.7.60) for some constant α.
Secondly, we upper bound Equation (4.7.61) by upper bounding the number and size of the r-connected components that can be generated on the lattice $\mathcal{C}$ from a specific set V_K.

Lemma 50. *There exists a lifting function f defined using the HDRG decoder such that for any subset V_K of* Sur *and arbitrary value of $\lambda > 1$ we have that*

$$\sum_{V_L \in f(V_K)} \lambda^{|V_L|} \leq \lambda_1^{|V_K|}, \tag{4.7.62}$$

with $\lambda_1 > 1$.

Proof. We use the previously described lifting function f based on the HDRG decoder and employ two bounds to constrain the sum on the left-hand side of Equation (4.7.62). We start to bind the size of each connected component. To do this, we observe that at most $|V_K|$ connected components exist in $|V_L|$ since V_L must intersect all the vertices of V_K. Next, we bound the size of each component by taking each vertex in V_K as a root node and consider respectively the largest tree that can be spanned given the graph's connectivity D. Given that we are considering r-connected components, the maximum distance between our root element, a vertex in V_K, and the leaf nodes is r.
In the second step, we need to bind the number of possible sets V_l in the lift. Since these are simply all the possible sets one can form, we can consider the power set of the vertices in the largest set V_L previously determined. By sequentially applying these steps, we obtain that

$$\sum_{V_L \in f(V_K)} \lambda^{|V_L|} \leq \sum_{V_L \in f(V_K)} \lambda^{|V_K| \cdot D^r} \tag{4.7.63}$$

$$\leq \sum_{i=0}^{|V_K| \cdot D^r} \binom{|V_K| \cdot D^r}{i} \lambda^{|V_K| \cdot D^r} \leq \left((2\lambda)^{D^r}\right)^{|V_K|} \tag{4.7.64}$$

taking $(2\lambda)^{D^r}$ as λ_1 completes the proof. □

With this, we conclude that there exists a lifting function with the necessary properties to ensure that our the repair function for single-shot state preparation satisfies the local stochastic property defined in Equation (4.7.58).
Finally, we obtain that with the 3D construction described in the previous section, as well as the respective recovery and repair functions, we obtain a noise-resilient single-shot GHZ_2 state preparation process for the qupit surface code as required by our conditions in Definition 43.

Lemma 51. *Let $\mathcal{E} \sim \mathcal{N}(\tau)$ be a local stochastic error and consider a surface code with code distance d. Then, there exists a constant-depth Clifford circuit $\mathcal{Z}$, along with corresponding recovery and repair functions, such that the conditions in Definition 43 are satisfied, ensuring noise-resilient preparation of the logical* GHZ_2 *state.*

Proof. This culminates with all the descriptions in the proofs of this subsection. We consider for that construction a 3D lattice composed of qupits as described in Equation (4.7.37), along with the corresponding constant depth state preparation circuit $\mathcal{Z}$, defined by Equation (4.7.39) and Equation (4.7.41). After performing the respective measurements on the Bulk, we use the precise repair and recovery functions defined based on the stabilisers of $\mathcal{S}_0$ and $\mathcal{S}_1$ from Equation (4.7.47),Equation (4.7.48) and Equation (4.7.49),

Equation (4.7.51), Equation (4.7.54), Equation (4.7.57) respectively. The application of these functions, as shown previously, results in a logical GHZ_2 state on the qupit surface code, with a code distance of d. □

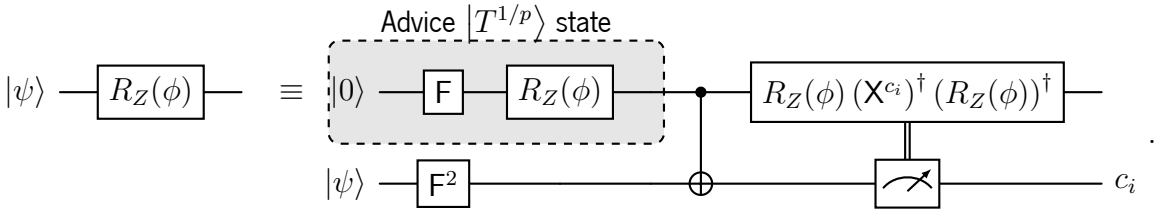
This provides the last ingredient for the noise-resilient implementation of any quantum circuit over qupits that has with the structure illustrated in Figure 24 using the qupit surface code. Next, we will discuss the specific quantum circuits that will resort to that circuit structure.

4.7.2 Noise-resilient quantum advantage with magic state injection

In this subsection, we address the incompatibilities between the logical quantum circuits proposed in section Section 4.6 and the circuit structure (Figure 24) required for proving noise resilience. This primarily involves managing non-Clifford operations in the quantum circuits, for which we have demonstrated computational separations for qupits with dimensions $p \geq 3$. These operations are essential and necessary because the separations observed with Clifford circuits using qubits do not generalize to qupits [121, 122]. Additionally, we face the constraint that the circuit cannot be adaptive, preventing the use of standard gate teleportation gadgets. To resolve this, we introduce a new quantum circuit that can solve the ISMRP family of problems using magic states and a non-adaptive constant-depth Clifford circuit.
We start off by demonstrating the existence of such a logical quantum circuit. We will then combine all the elements from the previous sections with our new logical quantum circuit, which has the necessary structure for noise-resilient execution. Finally, we will prove our noise-robust separations as outlined at the beginning of the section.

Non-adaptive Clifford circuits with magic state injection

To simplify the design of our new quantum circuits with the structure of Figure 24, we focus exclusively on the non-Clifford operators. In particular, we will address the implementation of the Z rotations for all dimensions $p \geq 3$. Our first step is to replace this gate with an equivalent realization using the following qupit gate teleportation gadget [299, 203], allowing us to offload the resource requirement for the execution of this gate into an advice state,



(4.7.65)

Obtaining advice states defined as follows,

$$\left|T^{1/p}\right\rangle := R_Z\left(\frac{2\pi}{p^2}\right)\mathsf{F}\,|0\rangle = \frac{1}{\sqrt{p}}\sum_{j=0}^{p-1} e^{\frac{2\pi i\cdot j}{p^2}}\,|j\rangle\,. \tag{4.7.66}$$

Understanding the state $\left|T^{1/p}\right\rangle$ in more detail is crucial, as it will define the type of corrections required in Equation (4.7.65). Therefore, defining the new quantum circuit we intend to execute, as well as the complexity of the advice states themselves. For instance, if $R_Z(\phi)$ is an operation from the third level of the Clifford hierarchy, the required correction within the gadget would be a Pauli operator. This correction is particularly convenient, as we can trivially commute it to the end of the circuit without significant complications.

To build intuition, notice that in the qubit case these advice states are related to the T operator. In particular, $\left|T^{1/2}\right\rangle = TH\left|0\right\rangle$. These states are called T magic states and are sufficient to perform universal quantum computation when combined with Clifford gates and adaptive measurements. Additionally, when provided as resource states in their logical version encoded with a quantum error correction code capable of realizing Clifford gates transversally, these states allow for universal *fault-tolerant* quantum computation. We now show that a similar type of relationship continues to hold between $\left|T^{1/p}\right\rangle$ states and the corresponding T-type magic states in higher dimensional qupit systems.

Lemma 52. *For all prime $p \geq 3$, the state $\left|T^{1/p}\right\rangle$ is derived from the pth root operator that creates the T-type qupit magic state.*

Proof. For universal quantum computation, it is known that any operator at the third level of the Clifford hierarchy or above, when combined with the Clifford group, allows for universal quantum computation [300]. The operation considered for implementing through-the-gate teleportation is a diagonal gate, which we intend to show is sufficient for universal computation along with Clifford operations and is related to standard magic states. As described in [301], the generalization of T gates for qupits, in combination with Clifford gates, allows for universal quantum computation over these higher-dimensional quantum systems. This generalization includes any operator in the family ($\omega = e^{\frac{2\pi i}{p}}$)

$$U(v_0, v_1, \ldots, v_{p-1}) = \sum_{k=0}^{p-1} \omega^{v_k} \left|k\right\rangle \left\langle k\right|, \tag{4.7.67}$$

parameterized by $\{v_0, v_1, \ldots, v_{p-1}\}$ satisfying $v_k = \frac{k}{12}(\beta + k(6\gamma + (2k+3)\beta)) + k\zeta$ for each $k \in [p-1]$ for some $\beta, \gamma, \zeta \in \mathbb{F}_p$. Furthermore, our advice state is created from the pth root of the operator that yields the following stat

$$\left|T\right\rangle = \frac{1}{\sqrt{p}} \sum_{k=0}^{p-1} w^k \left|k\right\rangle. \tag{4.7.68}$$

Now, we will prove the previously defined $\left|T\right\rangle$ states are magic states under the given conditions for the qupit generalised T gates. For that, we need to show that there exists a solution for the values of β, γ and ζ such that we obtain the vector $v_k = (0, 1, \ldots, p-1)$ for all primes. For instance, with $\beta = 0, \gamma = 2$, and $\zeta = 0$, we obtain the vector $v_k = (0, 1, 2)$ for the qutrit case. For $p = 5$, we obtain the vector $v_k = (0, 1, 2, 3, 4)$ with $\beta = 5, \gamma = 5$, and $\zeta = 6$. More generally, we have a computer-assisted proof[12], verifying analytically, that for all primes p, the following equation has solutions.

$$\frac{k}{12}(\beta + k(6\gamma + (2k+3)\beta)) + k\zeta \equiv k \ (\text{Mod } p). \tag{4.7.69}$$

This condition guarantees that $\left|T\right\rangle$ is indeed a T-magic state for all prime p as defined in [301, 302, 303], thus concluding the proof. □

We have characterized our advice states with respect to the T-type qupit magic state. On the other hand, we know that the latter is a magic state for a gate in the third level of the Clifford hierarchy, but any pth root of this gate must be of a higher level[13]. Thus we find that our advice state is not a magic state for the third level of the Clifford hierarchy, but for some gate of level $3 + p$. Consequently, the correction operator

[12] We resort to Wolfram Mathematica, specifically using the "Reduce" function, to analytically verify if there exist integer solutions for the free variables in all prime dimensions.

[13] By the recursive nature with which the Clifford hierarchy is defined (see Definition 34), we obtain, for instance, that the root of any operator at level l must be an element of level $l + 1$.

$R_Z(\phi)\left(\mathsf{X}^{c_i}\right)^\dagger\left(R_Z(\phi)\right)^\dagger$ for the associated gate teleportation gadget using the magic state is no longer a Pauli operator. This means we cannot simply commute it through to the end of the Clifford circuit. Since our pth root operator is in the $p+3$th level, this forces the correction operation to be an operator above the second level. Thus, it cannot be a Clifford operator either and cannot be implemented by any composition of the logical and transversal gates of the qupit surface code.

One potential solution is to use additional advice states to perform a cascade of gate teleportations, where each correction comes down one level in the Clifford hierarchy, ultimately making the final correction a Clifford operator [299, 203]. However, this approach would require a number of sequential adaptive steps which is beyond our reach in the shallow-depth QNC^0 model of interest to us. More precisely, we would need to decode the control qupit, which is a logical qupit, necessitating computations over a string of size $\Omega(\log n)$. Each output of a constant depth bounded fan-in circuit can only depend on a constant number of inputs, making such computations as required above infeasible in constant depth.

Nevertheless, whenever we intend to implement a rotation through a teleportation gadget as in Equation (4.7.65) but cannot realize the correction operations, we can overcome this issue by accounting for the effects on the output string due to the unrealized corrections, in addition to the rotations implemented by the respective gadget. More precisely, we will show that a circuit using the gadget from Equation (4.7.65) *without* the correction steps can still solve the target ISMRP with an exponentially larger correlation than any classical circuit. The proof for this statement will be based mainly on a different NC^0 reduction as the ones in Lemma 43 and Lemma 44, producing a different correction string from those analyzed so far. However, this is only possible if the correction operators have a specific form: they need to be a Z rotation followed by a Pauli operator. The Pauli operator can be commuted to the end and its effect inverted in a simple manner within the NC^0 reduction step. At the same time, the diagonal Z rotations will be accounted for with an additional correction string in the NC^0 reduction, as we will show. In the next lemma, we demonstrate that the correction operations have exactly the form that we require.

Lemma 53. *For any qupit input state $|\psi\rangle$, the output state in the first register of the teleportation gadget, as shown on the right-hand side of Equation* (4.7.65)*, can be written based on the measurement result $c \in \mathbb{F}_p$ in the second register as,*

$$\mathsf{X}^c G R_Z(\theta_1) G R_Z(\theta_2, S_c) R_Z(\phi)^c \left|\psi\right\rangle, \tag{4.7.70}$$

where $GR_Z(\theta)$ is the diagonal matrix with all diagonal elements equal to θ and $\theta_1 = -(2\pi c)/p^2$. The operator $GR_Z(\theta, S)$ is defined as follows

$$GR_Z(\theta, S) = \sum_{x \in S} e^{i\theta} \left|x\right\rangle\left\langle x\right| + \sum_{x \notin S} \left|x\right\rangle\left\langle x\right|, \tag{4.7.71}$$

and we have $\theta_2 = \frac{2\pi}{p}$ and the set S_c defined for each c as $[p-c, \ldots, p-1]$.

Proof. The output state of the teleportation gadget without applying the correction of the teleportation gadget shown in Equation (4.7.65), can be described based on the result of the measurement in the second register $c \in \mathbb{F}_p$ as,

$$\left(R_Z(\phi)\left(\mathsf{X}^c\right)^\dagger\left(R_Z(\phi)\right)^\dagger\right)^\dagger R_Z(\phi)\left|\psi\right\rangle. \tag{4.7.72}$$

Thus, we only need to show that the operator $\left(R_Z(\phi)\left(\mathsf{X}^c\right)^\dagger\left(R_Z(\phi)\right)^\dagger\right)^\dagger$ is equal to $\mathsf{X}^c G R_Z(\theta_1) G R_Z(\theta_2, S)$. This follows from a simple analysis of how the non-zero terms in these matrix operations are composed. First, observe that $\left(\mathsf{X}^c\right)^\dagger = \mathsf{X}^{p-c}$ and that $\left(R_Z(\phi)\right)^\dagger = R_Z(-\phi)$. From this, we know that if $c_i = 0$, then we have $R_Z(\phi)R_Z(-\phi) = I$. This follows simply from the fact that for all bases $|j\rangle$ with $j \in \mathbb{F}_p$,

the first operation introduces phase term $j \cdot \phi$, while the second introduced $-j \cdot \phi$, canceling each other out. However, if $c \neq 0$ the operator X^c is not the identity. This implies that the phases which operate are shifted according to the exponent $p - c$, and we get the following phases,

$$e^{j(c-j)\phi}, \text{if } j < p - c \text{ and } e^{j(-(p-c)-j)\phi}, \text{if } j \geq p - c. \tag{4.7.73}$$

Therefore, we have that c many bases state obtain a phase term with the value $-(p-c)\phi$ and $p - c$ many bases with a phase them having the value $c\phi$. The particular set of bases to which each one of these phases will be applied will be ignored for now as these follow from a simple analysis of the X^c operation for any $c \in \mathbb{F}_p \setminus 0$. Considering now the transposed conjugated operation, we obtain c many terms with the phase parameterized by $(p-c)\phi$ and $p - c$ terms with the phase $-c\phi$. Interestingly, one can write this operator simply as X^a operator with $a \in \mathbb{F}_p \setminus 0$ multiplied with a diagonal matrix defined by the previously determined phase terms.

$$D = \mathsf{diag}\Big(\underbrace{e^{-c\phi}, e^{-c\phi}, \ldots, e^{-c\phi}}_{(p-c) \text{ times}}, \underbrace{e^{(p-c)\phi}, e^{(p-c)\phi}, \ldots, e^{(p-c)\phi}}_{c \text{ times}} \Big). \tag{4.7.74}$$

Subsequently, we can factorize this diagonal matrix into two diagonal matrices as follows,

$$D_1 D_2 = \mathsf{diag}\Big(e^{-c\phi}, e^{-c\phi}, \ldots, e^{-c\phi}\Big)\mathsf{diag}\Big(\underbrace{1, 1, \ldots, 1}_{(p-c) \text{ times}}, \underbrace{e^{p\phi}, e^{p\phi}, \ldots, e^{p\phi}}_{c \text{ times}} \Big). \tag{4.7.75}$$

We obtain that D_1 is exactly $GR_Z(\theta_1)$, as $-c\phi$ is equal to $-(2\pi c)/p^2$. Similarly, D_2 is exactly equal to $GR_Z(\theta_2, S_c)$, since S_c corresponds to the elements over which D_2 operates with the phase $p\phi$, which is exactly $2\pi/p$. In our final step, we only need to determine the exponent of the operator X^a. Before the transposition, the phase terms were shifted by the operator X^{p-c}; thus, we obtain that $a = c$ after the transposition and complete the proof. □

Having defined the additional rotations that result from the teleportation gadget, we will now show that there exists a Clifford circuit which when provided with the quantum advice state $\left|T^{1/p}\right\rangle^{\otimes n}$ solves the ISMR problems with an exponentially larger correlation than $\mathsf{bPTF}^0[k]$ circuits with $k = n^{1/(5d)}$, as in Theorem 4.6.4.

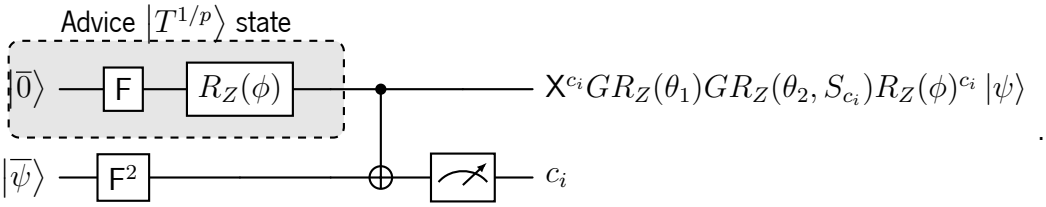
Lemma 54. *There exist a family of constant-depth Clifford circuits with the advice state $\left|T^{1/p}\right\rangle^{\otimes n}$, denoted as Cliff_{+T}, which can solve the ISMR problems $\mathcal{R}_p^m$, for a uniform distribution $\mathcal{D}_p$ of input strings that satisfy the condition $\sum_{i=1}^n x_i \ \mathsf{Mod}\ p = 0$ within $\mathbb{F}_2^n$. For any given input, these circuits achieve a success probability of $\frac{2p-2}{p^2}$. Additionally, they maintain an average correlation with the correct outcome within the Abelian domain, given by*

$$\mathsf{Corr}_{\mathcal{D}_p}\left(|\mathsf{Cliff}_{+T}|, |\mathcal{R}_p^m|\right) = \frac{p-1}{p^2}, \tag{4.7.76}$$

where $m = \mathcal{O}(n \cdot g^{p^3})$.

Proof. This circuit closely follows the quantum circuit considered in Figure 23. However, we substitute the classically controlled[14] rotations R_Z, which are non-Clifford operators, with the teleportation gadget as represented on the right-hand side of Equation (4.7.65) without the correction operations, obtaining the following classically controlled Clifford circuit,

[14] Whenever we refer to a classical controlled gate, we mean that the gate will be implemented or not based on a classical binary bit defining each possible case. This contrasts with standard controlled gates that can be controlled by a qubit or qupit, which might be in a superposition



Furthermore, we will consider the correction string $c \in \mathbb{F}_p^n$, which describes the measurement outcomes in all the teleportation gadgets and determines the additional rotations applied to the states. Note that $c_i = 0$ if $x_i = 0$, as no gadget is implemented for input bits, coming from the ISMR problems, with this value.

Having defined the setting, we need to redo the analysis of step 1 in Lemma 44, as this is the first step where these two quantum circuits attempting to solve the same problems differ intrinsically. Specifically, we must describe the effect of the additional rotations resulting from the inability to perform corrections from the teleportation gadget, as detailed in Lemma 53.

As we have three Z-rotations, they commute and we can implement these in any arbitrary order. Thus, we will consider the order most convenient for our analysis. First, we will analyze the rotations $GR_Z(\theta, S)$, which apply a Z rotation of magnitude θ to all the basis states in the set S. These rotations cannot be ignored, as they alter the measurement outcomes. To better describe this operation, we first define a vectorized version of the delta function $\overrightarrow{\delta}(x, i) : \mathbb{F}_p^n \mapsto \mathbb{F}_2^n$, by

$$\overrightarrow{\delta}(x, i) = \delta_i(x_1), \delta_i(x_2), \ldots, \delta_i(x_n), \tag{4.7.77}$$

with $\delta_i(x) = 1$ if $x = i$, and 0 otherwise.

Now, we can analyze the effect of these rotations applied to the generalized poor-mans qupit states independently, given that these effects will add to the rotations in the ideal quantum circuit,

$$\bigotimes_{i=0}^{n-1} GR_Z\left(\frac{2\pi}{p}, S_{c_i}\right) \left|\mathsf{GPM}_p^n\right\rangle = \frac{1}{\sqrt{p}} \sum_{i=0}^{p-1} e^{i\frac{2\pi \sum_{j=0}^{p-1}\sum_{k=0}^{j} \langle \overrightarrow{\delta}(c,j), \overrightarrow{\delta}(z^{+i}, n-k-1)\rangle}{p}} \left|z^{+i}\right\rangle \tag{4.7.78}$$

$$= \frac{\phi_g}{\sqrt{p}} \left(|z\rangle + \sum_{i=1}^{p-1} e^{i\frac{2\pi\left(\sum_{j=0}^{p-1}\sum_{k=0}^{j} \langle \overrightarrow{\delta}(c,j), \overrightarrow{\delta}(z^{+i}, n-k-1)\rangle - \langle \overrightarrow{\delta}(c,j), \overrightarrow{\delta}(z, n-k-1)\rangle\right)}{p}} \left|z^{+i}\right\rangle \right) \tag{4.7.79}$$

with the global phase $\phi_g = e^{\frac{2\pi \sum_{j=0}^{p-1}\sum_{k=0}^{j} \langle \overrightarrow{\delta}(c,j), \overrightarrow{\delta}(z, n-k-1)\rangle}{p}}$ and $S_{c_i} = [p - c_i, \ldots, p-1]$. Additionally, to simplify, we will express the terms in the phases as follows,

$$\varphi(z^{+i}, c) = \sum_{j=0}^{p-1} \sum_{k=0}^{j} \langle \overrightarrow{\delta}(c, j),\ \overrightarrow{\delta}(z^{+i}, n-k-1)\rangle. \tag{4.7.80}$$

This will allow us to rewrite the previous states and describing the state with all the rotations applied as follows,

$$\bigotimes_{i=0}^{n-1} R_Z\left(\frac{2\pi x_i}{p^2}\right) GR_Z\left(-\frac{2\pi \cdot c_i}{p^2}\right) GR_Z\left(\frac{2\pi}{p}, S_{c_i}\right) \left|\mathsf{GPM}_p^n\right\rangle \tag{4.7.81}$$

$$= \frac{e^{i\frac{2\pi\varphi(z,c)}{p}}}{\sqrt{p}} \left(\bigotimes_{i=0}^{n-1} R_Z\left(\frac{2\pi x_i}{p^2}\right) G R_Z\left(-\frac{2\pi \cdot c_i}{p^2}\right) \left(|z\rangle + \sum_{i=1}^{p-1} e^{i\frac{2\pi\left(\varphi(z^{+i},c)-\varphi(z,c)\right)}{p}} \left|z^{+i}\right\rangle \right) \right) \tag{4.7.82}$$

$$= \frac{\phi'_g}{\sqrt{p}} \left(|z\rangle + \sum_{i=1}^{p-1} e^{\frac{2\pi i\cdot\left(i|x|/p+\left(\sum_{j\in 1,\ldots,i}\left(\langle x,z^{+j}\rangle\right)^{p-1}+\varphi(z^{+i},c)-\varphi(z,c)\right)\right.}{p}} \left|z^{+i}\right\rangle \right), \tag{4.7.83}$$

with $\phi'_g = \exp\left(\frac{2\pi i\cdot\langle x,z\rangle+p\cdot\varphi(z,c)-|c|}{p^2}\right)$.
Subsequently, we repeat the same analysis as in Lemma 44. We know that the measurement outcome depends on the values of $\sum_{j\in[i]}\left(\langle x,z^{+j}\rangle\right)^{p-1}+\varphi(z^{+i},c)-\varphi(z,c)$, which once again are uniformly random values in $\mathbb{F}_p$ as both strings z and c are uniformly random in $\mathbb{F}_p^n$. Specifically, we can use the same analysis as before, given that we can focus uniquely on the overlap with the basis states of the following form $|\psi_p\rangle = |z\rangle + \sum_{j=1}^{p-1} e^{\frac{2\pi i\cdot(j|k|/p)}{p}} |z^{+j}\rangle$, as these define are the states that will project into strings that are congruent Mod p with $|k|$, and thus infer the probability that our state projects into a superposition of strings congruent with the correct outcome to the respective ISMR problem.
Once again, due to the random and uniformly distributed form of the terms $\sum_{j\in[i]}\left(\langle x,z^{+j}\rangle\right)^{p-1}+\varphi(z^{+i},c)-\varphi(z,c)$, the overlap with this state is equal to $1/p$, thus having zero correlation with ISMR problems. However, we can apply the same solution as in Lemma 44 and concatenate a correction string to the outcome in the form of one of these terms. We then obtain a positive correlation with the correct outcome string. In particular, we obtain the same correlation if we use the first term,

$$\langle x,(z^{+1})^{p-1}\rangle + \varphi(z^{+1},c) - \varphi(z,c). \tag{4.7.84}$$

The term $\langle x,(z^{+1})^{p-1}\rangle$ we already know how to compute with a NC^0 circuit, but we are missing to demonstrate that a string congruent modulo p with $\varphi(z^{+1},c)-\varphi(z,c)$ can be constructed with asymptotically small size using a NC^0 circuit. Given the string c, the values of the terms $\varphi(\cdot,\cdot)$ reduce to computing delta functions $\overrightarrow{\delta}(\cdot,\cdot)$. Moreover, the total number of such terms in the $\overrightarrow{\delta}(\cdot,\cdot)$ function does not pose a problem, as the number of these terms needing computation is upper-bounded by $\mathcal{O}(p^2)$.
Now we need only to show then how the $\overrightarrow{\delta}(a,b)$ for any input string a and dit b can be computed by an NC^0 circuit. For that we show that,

$$\overrightarrow{\delta}(a,b) = \left(\left(a\oplus(-b)^{\otimes n}\right)^{p-1}-1\right)^{p-1}. \tag{4.7.85}$$

Each of the terms $(a_i\oplus -b)^{p-1}=1$ if $a_i\neq b$, and equal to 0 otherwise. This means that we already have an inverse of the desired string. Now, we can subtract 1 so that all the 1s turn into 0s and the 0s to $p-1$. By applying the exponentiation of $p-1$, we ensure that all the 0s stay as zeros and the $p-1$ indices turn 1 as desired, effectively inverting the 0s and 1s.
These terms must be efficiently computable for both the strings c and z^{+i}. For c, this is trivially true as we have access to the individual bits, and thus, each one of these operations is over a single dit and can be computed trivially with an NC^0 circuit. For z^{+i}, this again is efficiently computable as the $\oplus$ operation can be realized by adding these values to the string, and the same method can be applied for the -1 operation. For the exponentiation, as described before, we can simply list all the products of terms, which are still bounded in number by $\mathcal{O}(n\cdot g^{p^2})$ and perform these operations dit-wise. Finally, the inner product between these terms $\langle\overrightarrow{\delta}(c,\cdot),\overrightarrow{\delta}(z^{+i},\cdot)\rangle$ can be evaluated with the same approach as we did for $\langle x,(z^{+1})^{p-1}\rangle$ in Lemma 44. With this correction string of size $\mathcal{O}(n\cdot g^{p^3})$ computed by an NC^0 circuit, we solve the respective ISMR problems with the intended correlation.
The last element to consider is that we can decompose the previously considered circuit into a constant-

depth Clifford circuit if we are given the magic states $\left|T^{1/p}\right\rangle$ (as proven in Lemma 52). □

With this, we have completed our construction of a family of quantum circuits over qupit gate sets composed of qupit Clifford gates with an additional $\mathsf{T}^{1/p}$ gate that can solve the ISMR problem for the respective prime dimension. This demonstrates that all the problems can be solved with minimal, finite, and standard univeral gate sets. Additionally, we will subsequently show that the ability to rewrite these circuits as non-adaptive constant-depth Clifford circuits with access to magic states allows for error resilience while maintaining constant depth.

Noisy quantum circuits retain a computational advantage against $\mathsf{bPTF}^0[k]$.

Armed with all these ingredients, we are now finally ready to combine the conditions for the fault-tolerant realization of constant-depth quantum circuits with the alternative family of non-adaptive quantum circuits proposed in Section 4.7.2 to demonstrate that noisy qupit $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuits can solve a family of relation problems beyond the capabilities of $\mathsf{bPTF}^0[k]$ circuits.
For this purpose, we use the relational problems from Definition 42, which are closely related to the original ISMR problems, and the noise-resilient implementation of $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuits solving the these problems from Section 4.7.2. More precisely, we will show that QNC^0 circuits with $\left|\overline{T^{1/p}}\right\rangle$ provided as advice states, despite suffering from local stochastic noise, can solve a family of relational problems that are AC^0-reducible to the ISMR problems (see Figure 26). To this end, we will exploit the fact that the functions associated with the noise-resilient implementation—dec and rec—are computable in AC^0, a necessary requirement for proving classical hardness against $\mathsf{bPTF}^0[k]$.
We first prove that there exist $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuits that solve the respective $\mathfrak{R}_p$ problem for all prime p with a constant positive correlation, bounded away from zero.

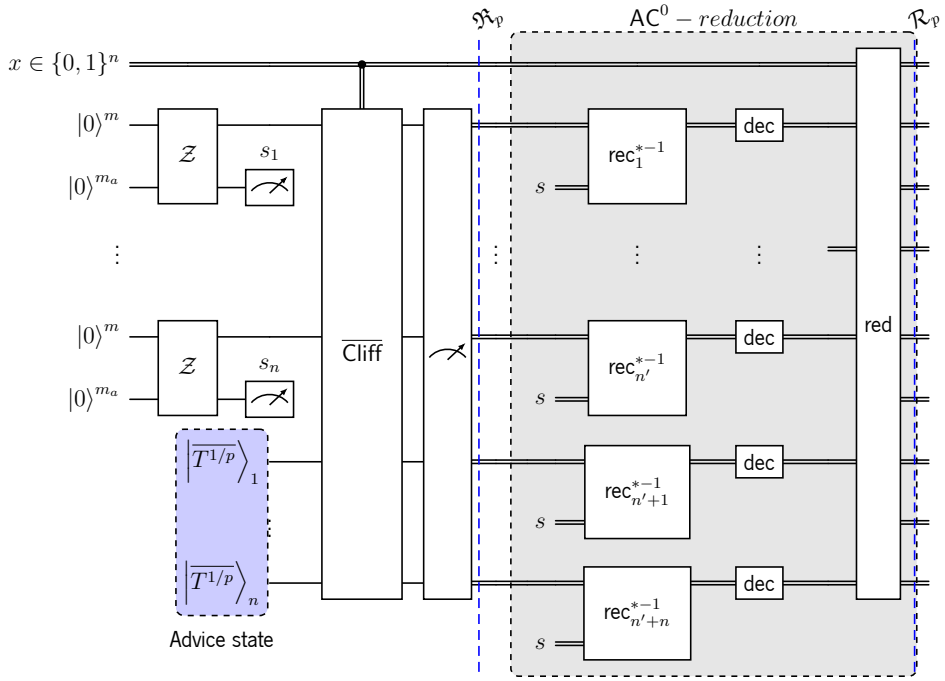


Figure 26: noise-resilient constant-depth Clifford quantum circuit with quantum advice states.

Lemma 55. *For all primes p, there exists a family of classically controlled constant-depth Clifford circuits with all-to-all connectivity in the class $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$, denoted as QC, that for any input x drawn uni-*

formly at random from the subset of n-bit strings with Hamming weight satisfying $(\sum_{i=1}^{n} x_i)$ Mod $p = 0$, and affected by noise $\mathcal{E} \sim \mathcal{N}(\tau)$ with a fixed probability $\tau < \tau_{th}$, QC solves the respective relation $\mathfrak{R}_p$ for output string of length $y = o(n^2)$ with a constant positive correlation,

$$\mathsf{Corr}_{\mathcal{D}_p}\Big(\mathfrak{R}_p\big(x, QC(x)\big)\Big) = 0.99 \cdot \left(\frac{p-1}{p^2}\right). \tag{4.7.86}$$

Proof. We will prove this claim by considering the noiseless versions of the circuit family described in Lemma 54, which solve the ISMR problems with a circuit architecture that can be noise-resilient based on Lemma 45 and Definition 43. Subsequently, we consider their logical version in the qudit surface code and show that the implementation of any element of this circuit family is noise-resilient with this code. Finally, we will demonstrate that decoding the logical outcome of any of these quantum circuits can be done efficiently by polynomial-size AC^0 circuits, and thus their outcomes are AC^0-reducible to solutions to the ISMR problems (see Figure 26 for a representation of the described circuit).
Now we will repeat the previous steps with all the necessary details. For that, we consider non-adaptive constant-depth Clifford circuits described in Lemma 54, and translating them into their logical version within the qupit surface code for a fixed distance $d = \text{poly}(\log n)$. In addition, we consider n repetitions of $\left|\overline{T^{1/p}}\right\rangle$, the logical version of the magic state $T^{1/p}$ in the surface code with distance d, as our advice state, which can be affected at most by local stochastic noise $\mathcal{E} \sim \mathcal{N}(\tau)$ below a fixed threshold τ_{th}. Simultaneously, we consider the single-shot state preparation process in Section 4.7.1 for creating the logical $\left|\overline{\mathsf{GHZ}_2}\right\rangle$ states within the surface code with the same distance d, without applying the recovery function rec but still producing the string s. Next, we apply the logical version of the classically controlled Clifford circuit described in Lemma 54, which we designate $\overline{\mathsf{Cliff}}(x)$, except we omit the execution of certain logical Sum gates that are unnecessary because we start with the $\left|\overline{\mathsf{GHZ}_2}\right\rangle$ states and not logical $\left|\overline{0}\right\rangle$ states. The quantum circuit then concludes with a layer of Z measurements, as displayed in Figure 24. Given the outcome of this circuit, we will consider the string s from the single-shot state preparation and the outcome of the logical circuit y. We can define can then define the functions $\mathsf{rec}_i^{*\dagger}$ (displayed in Figure 26) as the inverse of the recovery function $\overline{\mathsf{Cliff}}(x)\mathsf{rec}(s)^{\dagger}\overline{\mathsf{Cliff}}(x)^{\dagger}$, restricted to the single logical qupit i. This undoes the effects of the errors occurring during the state preparation. Additionally, we compose to this operation the inversion of the logical X^a gate from the teleportation gadget introduced in Lemma 54, completing the definition of each $\mathsf{rec}_i^{*\dagger}$. Furthermore, each one of these functions rec_i^{*-1} can be computed by an AC^0 circuit using the arguments presented in [35, 36].
Now, given the assumption that the entire circuit is affected by local stochastic noise $\mathcal{E} \sim \mathcal{N}(\tau)$ bounded below τ_{th}, the noise-resilient single-shot GHZ_2 state preparation from Section 4.7.1, and choosing the decoding function dec to be a function implemented by the HDRG decoder, which can be computed by an AC^0 circuit for our code distance, we derive from the combination of Lemma 51, Corollary 10, and Lemma 45 that after the decoding function, we have the exact outcome of the quantum circuit described in Lemma 54 with a probability of 0.99. By the same lemma, we have an NC^0 reduction, (represented as red in Figure 26), which provides outcome strings that correlate as described with the $\mathfrak{R}_p$ problems, completing the proof since it is trivially computed by an AC^0 circuit. □

Simultaneously, we can prove that any $\mathsf{bPTF}^0[k]$ circuit has close to zero correlation with the $\mathfrak{R}_p^{n'}$ problems for the size of the outcome strings that the $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ under consideration produces.

Proof of Theorem 4.7.1. From Lemma 55, we obtain that there exists a $\mathsf{QNC}^0/\left|\overline{T^{1/p}}\right\rangle$ circuit that solves the $\mathfrak{R}_p^{n'}$ for all prime p with the specified outcome strings. Now, we intend to demonstrate that, for the same parameters, $\mathsf{bPTF}^0[k]$ circuits have close to zero correlation with the same relation problems.

To link the correlation of solving $\mathfrak{R}_p^{n'}$ with $\mathcal{R}_p^{n'}$, we use that by definition they are AC^0-reducible to each other. Thus, a solution with a high correlation to $\mathfrak{R}_p^{n'}$ by a $\mathsf{bPTF}^0[k]$ implies the existence of a slightly larger $\mathsf{bPTF}^0[k]$ with the same correlation for $\mathcal{R}_p^{n'}$. Therefore, by considering a constant depth d' for this AC^0 reduction, we obtain the bounds described in Theorem 4.7.1 based on the lower bound proven in Theorem 4.6.4, concluding the proof. □

We conclude by considering the isolated qubit case, which has two main distinguishing features compared to the previously proven qudit cases. First, it does not require magic states; unlike those for all other primes $p \geq 3$, the qubit circuit will be a Clifford circuit. Second, it allows for separation with a fixed probability due to the one-to-one relationship between the correlation function and the likelihood of producing a correct or incorrect outcome for the $\mathfrak{R}_2^{n'}$ problem.

Proof of Corollary 9. This corollary follows more easily, as we consider exactly the same ideas for the qupit circuit in Lemma 55, but we do use as the underlying logical Clifford circuit the circuit from Lemma 44 for qubits, without additional advice state, nor teleportation gadgets, given that the necessary rotation is a Clifford operator.
For the $\mathsf{bPTF}^0[k]$ lower bound, we use the same lower bound as in Theorem 4.6.2, which allows us to bound the success probability of any $\mathsf{bPTF}^0[k]$ solving the $\mathfrak{R}_2$ problem. □

4.8 Discussion

In this paper we have advanced the nascent line of work showing unconditional separations between the computational power of classical and quantum shallow-depth circuits. We have shown that linear-size local QNC^0 circuits can solve multi-output search problems (relation problems) that polynomial-size circuits of k-biased polynomial threshold gates fail to solve with appreciable probability even on average. We have also developed a family of non-local games over qupits of each prime dimension and used the separation in winning probabilities between classical and quantum strategies to show that the computational separation we establish for qubit QNC^0 circuits extends to QNC^0 circuits over higher dimensional quantum systems. This is especially significant in the light of our belief that under finite universal gate sets, the corresponding QNC^0 circuit classes over p and q dimensional systems are incomparable for $p \neq q$ (see Conjecture 3). Finally, we have also shown that the computational separations we establish are noise-robust, in that even quantum circuits affected by local stochastic noise continue to exhibit computational advantages over noiseless classical circuits, using techniques that draw on magic state injection and hard decision renormalization group decoders.
We conclude this article exploring the applicability of our results to practical developments in quantum computing and machine learning. As we rapidly progress through the NISQ era, with several noisy intermediate scale quantum devices with different underlying architectures and technologies on the horizon, precise estimates of the resource requirements for demonstrations of quantum computational advantage become increasingly relevant. Such resource estimation and optimization is of particular significance to work such as ours that focuses on shallow-depth circuits that may be within the reach of both theoretical optimization and practical implementation. In Section 4.8.1, we present the best known order-of-magnitude circuit width, depth and size estimates for witnessing unconditionally provable separations of computational power between QNC^0 and the largest known classical circuit classes to date.
In learning theory, the expressivity of hypothesis classes such as neural network architectures commands as much attention as computational complexity. Decades of work have approached this problem with tools from circuit complexity. Proving that a forward pass through a neural network architecture can be simulated by a circuit class such as AC^0 establishes upper bounds that allow us to carry over known

results about circuit complexity lower bounds to understand the expressivity and generality of the neural network. In Section 4.8.2, we discuss some relations between our findings on circuit complexity and recent results on the simulation of popular neural network models by shallow depth classical circuits.

4.8.1 Resource estimation

In physical implementations that test the kind of unconditional separations that we present, it is key to pin down at what values of circuit depth d and input size n (i.e. number of input qubits) we observe a transition in the circuit size. That is, at what depths and input sizes do the quantum advantages kick in? To answer this question, we can make preliminary estimates by solving for the parameter values at which the asymptotic lower bounds for the size of the best classical circuit become equal to the quantum upper bounds. Using the best estimates for the constant factors hidden in the asymptotics, we solve this equality for values of the input size n and depth d at which the quantum circuit size becomes smaller than the classical circuit size.

We can leverage our exact-case hardness bound in Theorem 4.2.1 to derive tighter estimates compared to what is possible using prior work, bringing theoretical predictions closer to the capabilities of current quantum devices. For context, the transition point for Shor's factoring algorithm is estimated to be $\sim$1,700 qubits, 10^{36} Toffoli gates, and a circuit depth of 10^{25} [125], while for the HHL quantum matrix inversion algorithm it is roughly 10^8 qubits and a depth of 10^{29} [11]. Recent advancements in quantum hardware favor larger devices (i.e., more qubits) over those with prolonged coherence, and so there is a push towards shallower circuits [126, 127]. Thus, separations against NC^0 represent a promising avenue for near-term quantum devices, requiring only hundreds to thousands of qubits to demonstrate classical non-realizability results, such as Bell violations [128, 129, 18]. While much work in this direction has focused on such information theoretic demonstrations, a *computational* example of this is solving the 2D Hidden Linear Function problem using a noise-free quantum circuit. In this case, the classical model being compared—circuits of bounded fan-in—require a depth of at least $d = \Omega(\log n)$ [18]. This provides a demonstrable quantum advantage when the quantum circuit resolves an instance with a depth strictly smaller than the minimum possible depth for any classical circuit. In a more realistic scenario that includes noise, the depth of the quantum circuit may increase by a constant factor, but the minimal classical circuit depth will still be at least $d = \Omega(\frac{\log n}{\log\log n})$ [35]. This adjustment, due to the presence of noise, slightly increases the previous estimates of the input sizes required for the problem to witness a quantum advantage. Additionally, imposing connectivity restrictions on the classical circuit—which are not part of the NC^0 definition—the size of the input strings at which the crossover to the quantum advantage regime happens may be effectively reduced [304].

Moving up the ladder of potential demonstrations of unconditional quantum advantage, the challenge is to outperform larger classical constant-depth circuit classes such as AC^0. Once again we may ask what the minimal size input size is at which which classical circuits require more gates to solve the instance than quantum circuits. As the quantum circuits for the Parity Halving Problem (PHP) have a linear number of gates $\Theta(n)$, our main tasks are to bound the constant factor hidden by the Θ notation, and to determine the minimum classical circuit size. As before we equate the PHP lower bound on the size of the AC^0 circuits and the upper bound on the size of the QNC^0 circuits, and solve for the input size. It is crucial that we do this after fixing the minimal allowed depth, because otherwise the subexponential dependence of the classical circuit size on the depth results in astronomically large input sizes. Specifically, using the local 2D QNC^0 circuit proposed in [27] that achieves a separation against AC^0 under average-case scenarios, avoiding asymptotic simplifications and using the precise bounds and parameters ($q = \sqrt{\log n}$) provided by the authors, and using the depth $d = 5$ version of the circuit rewritten in the MBQC setting, we find that even with these choices we would require approximately 10^{97} qubits to observe a quantum advantage.

$\mathsf{bPTF}^0[k]/\mathsf{rpoly}$	All-to-all
Exact, $k = \mathcal{O}(1)$	$\log_2(s) = \Omega\left(e^{\frac{1}{1-d}}\left(\frac{n}{\sqrt{n\log n}\log(n\log n)}\right)^{\frac{1}{-1+d}}\right)$
Exact, $k = n^{1/(5d)}$	$\log_2(s) = \Omega\left(e^{\frac{1}{1-d}}\left(\frac{n^{\frac{3}{10}}}{\sqrt{\log n}\log(n\log n)}\right)^{\frac{1}{-1+d}}\right)$
Average, $k = n^{1/(5d)}$	$\Pr[\text{Success}] \leq \frac{1}{2} + 2^{-\Omega\left(\frac{n^{8/5}(n\log n)^{-1-\frac{2}{\sqrt{\log n+\log\log n}}}}{2^{\left(2\sqrt{\log n+\log\log n}\right)}(\log s)^{2d-2}}\right)}$

Table 6: Size lower bounds for the circuit classes $\mathsf{bPTF}^0[k]$ without simplifications and optimal values for the parameter q.

However, by moving from local 2D circuits to all-to-all qubit connectivity and optimizing the parameters in the random restrictions technique used to lower bound the classical circuit size specifically for this connectivity, we can drop this requirement to 10^{21} qubits with depth-3 quantum circuits, a considerably lower value than the previous rough estimates. Similar estimates can be obtained for increasingly larger classical circuit classes by leveraging our average-case hardness bound in Theorem 4.2.1. For instance, for $\mathsf{bPTF}^0[n^{1/(5d)}]$ circuits, we obtain that one would require 10^{40} qubits for a demonstration of quantum advantage, once again using suitably optimized quantum circuits and parameter values (see Table 6).
Towards this goal, with our exact-case hardness bound from Theorem 4.2.1, without any simplification as shown in Table 6, we are able to reduce by several orders of magnitude the estaimted minimal quantum resources required for demonstrations of quantum advantage. Specifically, we find that depth-3 circuits with 10^{11} and 10^{22} qubits respectively can exhibit quantum advantages over AC^0 and $\mathsf{bPTF}^0[n^{1/(5d)}]$ circuits. This substantial reduction in resource requirements marks a step forward in the progression of quantum advantage experiment design.
To conclude, we remark that our estimates do not by any means represent lower bounds on the threshold of input size necessary to establish a quantum advantage regime, as improvements in proof methodologies and parameter optimization are still possible.

4.8.2 Neural networks as biased threshold circuits

In this concluding section, we turn our attention to the connections between our main results and active fields of research in classical and quantum machine learning. We especially intend to (loosely) interpret how our techniques and conclusions relate to classical neural networks. We shall also identify potential implications of computational separations between shallow depth circuit classes for studying quantum over classical advantages in machine learning tasks.

Connection to classical neural networks

The class of classical constant depth circuits TC^0 was first defined in the quest for greater computational power beyond Boolean AC^0 circuits [259]. Furthermore, it is also well motivated as a model for neural networks composed of Boolean neurons, inspired by biological neural networks [305]. This class notably includes the threshold function $\mathsf{Th}_w : x \in \{0,1\}^n \mapsto 1$ iff $|x| > w$ and 0 otherwise, which captures the most basic activation behavior that can be exhibited by a neuron. Interestingly, this theoretical relationship has typically been analyzed retrospectively: new neural network models first prove themselves as practically

useful, and only subsequently is their standing relative to the circuit classes studied in complexity theory understood, both interms o expressivity and computational capability. For instance, neural networks with fan-in bounded by the logarithm of the input size fall within the AC^0 class [306], and thus they cannot even compute the parity function. On the other hand, neural networks whose nodes can have unbounded fan-out are known to be simulatable in TC^0 for a wide variety of weight types associated with the neurons' activation function [307]. Based on the widely held conjecture that $\mathsf{TC}^0 \subsetneq \mathsf{NC}^1$, these neural networks are therefore unlikely to be able to solve problems considered 'simple,' such as iterated matrix multiplication and solving linear equations, which are known to be solvable in NC^1 and P respectively [308, 309].
More recently the transformer architectures of neural networks [310], well-known for their spectacularly successful and widespread use in large language models (LLMs) such as ChatGPT, have been analyzed through the lens of constant-depth circuit classes. For instance, depending on the type of attention function used, the expressivity of transformer models may be limited; one example is the upper bound on hard attention in terms of languages computable by AC^0 circuits [311]. This suggests that any such model is generally weak, as we know of various relevant and simple computational functions, such as parity and majority, that are not included in this circuit class. Nevertheless, it is also worth noting that even this potentially weak attention model still finds powerful applications in computer vision models [312], and continues to be an area of ative development [313]. Transformers with the less restrictive saturated attention exhibit greater expressivity, and in a sense their power coincides with TC^0 circuits [3, 314]. This matches up with our empirical understanding that transformers employing more complex attention functions demonstrate better performance. The transformer architecture can be further extended to be able to capture computations in P, and indeed even transformers restricted to hard attention are known to be Turing-complete if they are allowed to perform computations with arbitrary precision and run for an unbounded number of time-steps [315]. However, this is not a realistic setting in practice, and it has been pointed out that practical cases with the precision of their internal weights and variables restricted to be logarithmic in the input length (a generous allowance) correspond exactly to uniform TC^0 [6]. We refer the reader to [2] for a survey of studies on the expressivity of neural network models in terms of circuit complexity.

We now proceed to explore the connection between the analysis and techniques in our work and the existing literature. We will demonstrate how one can compute commonly used activation functions using $\mathsf{bPTF}^0[k]$ circuits, and examine how the bias k influences this. Notably, this shows that a substantial subclass of neural networks, including transformers with non-trivial attention functions which extend beyond AC^0 circuits [3] can be analyzed using our techniques and switching lemmas. Furthermore, we remark that this analysis is applicable not only to decision problems but also to relation problems. This is of significance since neural networks, especially LLMs, should primarily be considered sequence-to-sequence models, that take (bit) strings as input and produce (bit) strings as output.
Consider discretizing a function of the type $f : \{0,1\}^n \mapsto [0,b] \subseteq \mathbb{R}$ by choosing parameter $w \in \mathbb{Z}$ and discretizing the range $[0,b]$ into steps of size w, to obtain $\hat{f} : \{0,1\}^n \mapsto \mathbb{Z}_B$ where $B = w \cdot \lfloor \frac{b}{w} \rfloor$. Discretizing an activation function allows us to implement it by $\mathsf{bPTF}^0[k]$ circuits, generating a k-biased function. This means that the function exhibits a 'tail' that always outputs 1 for inputs with a Hamming weight greater than k, akin to an AND type activation function. In Algorithm 3, we demonstrate that a discretized version of any activation function can be implemented accurately by $\mathsf{bPTF}^0[k]$ circuits for input bitstrings of Hamming weight less than k.
It is easy to implement the discretized version for input bitstrings with a Hamming weight greater than $n-k$ by a related circuit. It includes an initial segment that always outputs 0 for all inputs with a Hamming weight less than $n-k$, resembling an OR type activation function. Specifically, our algorithm requires only a single layer of PTF gates in parallel, thereby achieving a depth-1 $\mathsf{bPTF}^0[k]$ circuit that implements $\hat{f}$. Figure 27a presents a notable example of this discretization for a specific parameter k of the $\mathsf{bPTF}^0[k]$ circuits for the widely used ReLU activation function, which is defined as $f(x) = \max\{0, x-c\}$ (where

we have shifted the centre from 0 to c). Our scheme considers ReLU : $\{0,1\}^n \to [0, n-c]$ taking n-bit strings as input and interpreting their Hamming weight as the input x.

Remark. *Note that functions with larger activation regions (i.e., those with higher values of the bias k) can be achieved by composing deeper* bPTF$^0[k]$ *circuits. Also, by combining* AND *and* OR *type decompositions within a single depth-1 layer, one can obtain non-trivial activation regions for both Hamming weights smaller than k and larger than $n-k$. This approach results in only a central region of fixed Hamming weight rather than the entire tail as in each case,* AND *and* OR *type decompositions.*

Algorithm 3 Hamming-Weight Bounded Activation Function Decomposer

procedure Decomp(Bias parameter k, activation function $f : \{0,1\}^n \mapsto \mathbb{R}$, resolution parameter $w \in \mathbb{N}$)
  $gates \leftarrow [], i \leftarrow 0$.
  $l \leftarrow \lfloor \max_x (f(x)/w) \rfloor$ for $|x| \leq k$. ▷ Number discrete steps/gates in the decomposition.
  while $\exists x \; s.t \; f(x) - i \cdot w > 0$ with $|x| \leq k$ **do**
    Find $S_i \subseteq [n]$ s.t. $\forall j \in S_i \; \lfloor (f(x_j) - i \cdot w) \rfloor = 0$.
    Define

$$\text{bPTF}_i[k](x) = \begin{cases} 1, \text{ for all } x_j \text{ with } j \in S_i \text{ and } |x_j| \leq k \\ 0, \text{ otherwise} \end{cases} .$$

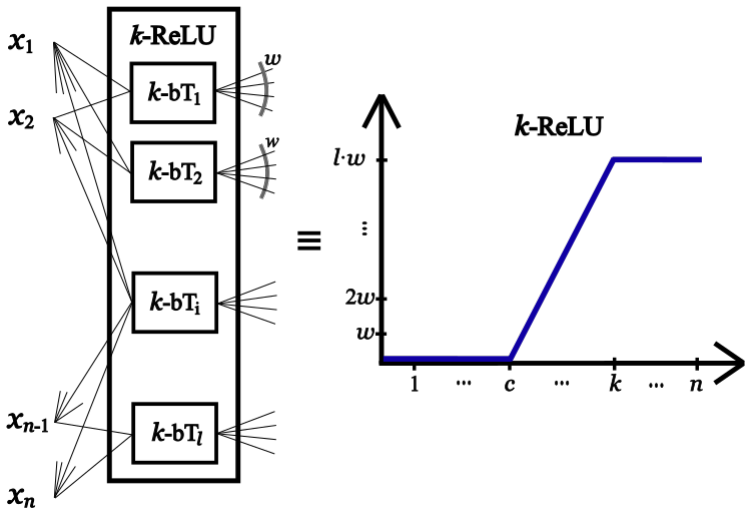
    Concatenate $(gates, \text{bPTF}_i(k))$.
    $i = i + 1$.
  end while ▷ The loop ends when $w \cdot i > \max_x f(x)$.
  return $(l, gates)$
end procedure

High-level description: The algorithm sequentially calculates the set of points of the activation function at discrete intervals defined by $w \cdot i$, where i is an integer. It then groups these points into sets of indices corresponding to where the discretized function matches the value $w \cdot i$. This process continues iteratively until the final discretized step at $w \cdot k'$ closely approximates the maximum value of f.

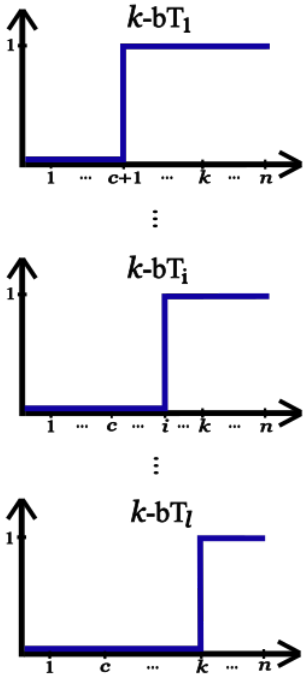
Connection to quantum neural networks

As we have seen, our computational separations demonstrate a quantum advantage over a model that encompasses many interesting classical machine learning models. From another perspective, this contributes to the extensively studied field of quantum machine learning. In particular, several studies have identified learning tasks that might benefit from a quantum advantage. Much work has also focused on the fundamental problem of learning a robust quantum solution, with a significant discussion based on the premise that models which are easy to train tend to be classically simulatable and thus are unlikely to offer quantum advantage. Conversely, models that are very general and potentially offer a quantum advantage tend to be difficult to train, from a theoretical perspective.

It is of great interest to understand which models emerge as good candidates for quantum advantage in the first place. In this setting, some notable classical-quantum separations rely on encoding a problem known to be hard into a valid quantum solution [316]. Others resort to quantum phenomena to explain enhanced expressivity and performance by explicitly pointing out where non-locality or contextuality clearly distinguishes between quantum and classical models at the operational level [317, 318, 319, 320].



(a) Representation of a k-ReLU gate, which is equivalent to a ReLU gate, up to precision w, for every input string with a Hamming weight bounded by k. The k-ReLU gate is contained in bPTF$^0[k]$; the precise circuit can be obtained by Algorithm 3, which defines k' the number and the particular bPTF$[k]$ gates used in the k-ReLU layer along with the respective fan-out wires. These elements are depicted on the left side of the figure, and the composite gate, which is the outcome of the circuit, is displayed on the right.



(b) Illustration of the unbiased region of each one of the k' bPTF$[k]$ gates employed in the k-ReLU gate. These functions are obtained from Algorithm 3, based on the parameter k, the activation function f and w the size of the discretization steps.

Our computational separation resembles the latter, where non-classical phenomena are translated into computational complexity. Furthermore, separations like ours have been utilized in a manner that combines both of the previously mentioned ideas: they demonstrate that certain problems that intrinsically exhibit quantum phenomena can be efficiently solved by quantum models, while remaining intractable for classical models cannot handle. Interestingly, these separations enable us to explore potential quantum advantages in machine learning problems and the benefits of quantum models over classical ones. These analyses have extended to include standard figures of merit used in machine learning, such as the Kullback-Leibler divergence, which was shown to differ between classical and quantum models [321]. This opens avenues for further research.

Finally, we note that the line of work we advance demonstrates that even simple and potentially easy-to-learn quantum circuits can outperform their classical counterparts. This supports other findings that even efficiently learnable quantum circuits could provide a quantum advantage in practical information processing tasks [317, 322, 323, 318, 321, 319, 324, 325].

Chapter 5

"These are the other Potentials."

— The Matrix

The power of constant-depth Toffoli and qudit quantum circuits

Overview

Significant progress has recently been made in realizing multi-qubit gates as native operations within constant-depth circuits, achieving execution in fixed time steps regardless of the number of qubits involved [79, 80, 81, 82, 326]. These advancements inspire further exploration of shallow-depth quantum circuits incorporating such gate sets. Within this framework, another prominent goal of quantum circuit complexity is to identify problems that shallow quantum circuits composed from these gate sets can efficiently solve, while their classical counterparts would require substantially more computational resources.

Building on this context, our first contribution in this chapter is to prove new separations between classical and quantum constant-depth circuits. Firstly, we show a separation between constant-depth quantum circuits with quantum advice $\mathsf{QNC}^0/\mathsf{qpoly}$, and $\mathsf{AC}^0[p]$, which is the class of classical constant-depth circuits with unbounded-fan in and $(\mathrm{mod}\ p)$ gates. In addition, we show a separation between QAC^0, which additionally have Toffoli gates with unbounded control as well as classical fanout, and $\mathsf{AC}^0[p]$. This establishes the first such separation for a shallow-depth quantum class that does not involve quantum fanout gates.

Secondly, we consider QNC^0 circuits with infinite-size gate sets. We show that these circuits, along with (classical or quantum) prime modular gates, can implement threshold gates, showing that $\mathsf{QNC}^0[p] = \mathsf{QTC}^0$. Finally, we also show that in the infinite-size gateset case, these quantum circuit classes for higher-dimensional Hilbert spaces do not offer any advantage to standard qubit implementations. The paper providing the content for this section can be found in [96].

5.2 Introduction

In the current landscape dominated by NISQ (Noisy Intermediate-Scale Quantum) devices [327], understanding the computational power of shallow-depth quantum circuits has both practical and theoretical relevance. From a practical perspective, given the noisy aspect of near-term devices, focusing on tasks that achieve quantum advantage in low depth is fundamental. From a theoretical perspective, (classical) constant-depth circuits have been crucial in the development of lower bound techniques, e.g. [328]. Thus, the combination of these techniques, along with the computational capabilities of constant-depth quantum circuits, enables the discovery of new quantum-classical separations and, furthermore, allow a more comprehensive understanding of quantum computational complexity classes [51**?** , 74].

The holy-grail in quantum circuit complexity is to find problems that can be solved by shallow-depth quantum circuits and that cannot be solved by polynomial-time classical computation. Unfortunately, there are several caveats in the current approaches. Firstly, the current super-polynomial quantum advantage proposals rely on computational assumptions, offering no unconditional proof of quantum superiority. Secondly, another set of results requires quantum gates with arbitrary accuracy, which cannot be implemented in practice in real-world devices at a constant depth. This problem arises when infinite gate sets are assumed [78, 263], as it has been shown that even if precise quantum operations were feasible, the computational overhead for classical control could negate any quantum advantage [119].

More recently, there is a new line of research that has addressed some of these issues, at the cost of achieving weaker statements. More concretely, the seminal work of [18] has shown a problem that can be solved by constant-depth quantum circuits (QNC^0), and which cannot be solved by constant-depth classical circuits with bounded fan-in (NC^0). We would like to highlight that while this separation is weaker (since the lower bound is only against constant depth circuits), it is *unconditional*, which is rather rare in complexity theory. Moreover, these proposals also have practical interest with their extension to include fault-tolerance [329].

The result of [18] has been also extended to achieve better quantum/classical separations: We currently have separations between quantum constant-depth circuits and constant-depth classical circuits with *unbounded fan-in* (AC^0), even if the classical circuit has access to $\pmod 2$ gates ($\mathsf{AC}^0[2]$) [27].[1] In these results, we are required to show a quantum upper and a classical lower bound for the same problem. The main challenge lies in finding new techniques for classical lower and quantum upper bounds that allow us to answer the question:

What is the largest class of classical circuits that we can separate from constant-depth quantum circuits?

The main contribution of this work is to prove new separations between constant-depth quantum circuits and constant-depth classical circuits. Additionally, we revisit quantum circuits with infinite gateset and prove new collapses of different quantum circuit classes.

5.2.1 Our contributions

Our first contribution is showing that a set of relation problems for which we have constant-depth quantum circuits on qudits with quantum advice and is not contained in $\mathsf{AC}^0[p]$.

Result 1. *(Informal; see Theorem 5.4.3) For all primes p, there exists a finite gate set on qudits of dimension p such that constant-depth quantum circuits with quantum advice can solve relation problems that are intractable for any polynomial-size $\mathsf{AC}^0[p]$ circuit, i.e., $\mathsf{QNC}^0_p/\mathsf{qpoly} \not\subseteq \mathsf{AC}^0[p]$.*

[1] Note that QNC^0 has access to quantum advice in separations against classical circuits with parity gates.

This result extends the $\mathsf{AC}^0[2]$ separation from [27] to general primes. To prove it, we generalize the relation problems on modular operations that have been used in [18, 27]. Roughly, our modular relation problems with parameters $p, q \in \mathbb{N}$ consist of a pair of bit strings (x, y) such that $|x| \pmod p = 0$ iff $|y| \pmod q = 0$. Here, $|x|$ and $|y|$ represent the Hamming weights of the respective bit strings. We show that by exploiting higher-dimensional gate sets, we can solve these problems in constant-depth quantum circuits with a copy of a high-dimensional GHZ state. At the same time, we expand the classical lower bound to $\mathsf{AC}^0[p]$. However, this separation achieves only a quantum/classical distinction regarding error probabilities, which differ by a constant value. In order to improve these separations, we then consider the parallel repetition of these modular problems as proposed in [27, 34]. In our results, we simplify the proof of [27] by considering the one-sided error of the quantum circuits and using "seminal results" in circuit complexity such as the Vazirani XOR lemma [134] and Razborov-Smolensky separations [19].
We follow by showing that QAC^0 circuits, which allow classical but not quantum fanout, can generate high-dimensional GHZ states. For this, we employ optimal depth and size methods for GHZ state creation, which is based on poor-man's cat states and balanced binary trees as the resource state entanglement structure. This gadget, along with Result 1, allows us to show that for every prime p, there exists a constant-depth quantum circuit employing Toffoli gates of dimension p that computes a relation unattainable by any polynomial-size $\mathsf{AC}^0[p]$ circuit. Moreover, when Toffoli gates of arbitrary arity can implement any qubit operation, we attain an unconditional separation between this complexity class and all $\mathsf{AC}^0[p]$ classes simultaneously. It is worth noting that this aligns with the standard definition of QAC^0 as outlined in the existing literature [137, 135, 138, 139] while including classical fan out and excludes quantum fanout, thereby removing one of the technically difficult multi-qubit quantum operations to achieve such separations. Furthermore, recent advances and practical demonstrations have suggested that some quantum device platforms inherently support multi-qubit operations, enabling the construction of multi-qubit Toffoli gates [79, 80, 81, 82].

Result 2. *(Informal; see Theorem 5.4.4) For every prime p, there is a relation problem solvable by QAC^0 circuits without quantum fanout, whereas any $\mathsf{AC}^0[p]$ circuit of polynomial size solves it with a negligible probability. Therefore, $\mathsf{QAC}^0 \not\subseteq \mathsf{AC}^0[p]$ holds for all prime p.*

This establishes the first separation between QAC^0 and $\mathsf{AC}^0[p]$ for any prime p with the single addition of classical fan-out, the weakest resource so far [136, 76, 99], pushing forward the question of the minimal resource, which was recently asked in [135]. We notice that in order to achieve such a result with our techniques, it is crucial to consider qudits of higher dimensions. Additionally,
we remark that QAC^0 can generate a GHZ state without the need of unbounded parity gates, as proposed in [135]. Our QAC^0 circuits use quantum measurements and adaptive operations to generate the states. Surprisingly, the capability to produce this state does not imply the capability to compute the parity function. We conjecture that generating the GHZ state through a unitary procedure (i.e. without measurements) would require quantum fanout or parity gates. Thus, the question of whether the parity function can be computed in QAC^0 without fanout gates remains open for further investigation [52].

Next, we provide some limitations on our techniques by showing a classical upper bound on the modular relation problems that we consider.

Result 3. *(Informal; see Lemma 63) For every primes p, q, the modular relation problem on strings (x, y) such that $|x| \pmod p = 0$ iff $|y| \pmod q = 0$ can be solved by $\mathsf{NC}^0[q]$ circuits.* [2]

We notice that this result was unknown even for $p = 2$, and it characterizes $\mathsf{AC}^0[p]$ as the largest constant-depth classical circuit class for which a quantum-classical separation can be identified in the plain model using our proposed relation problems. Indeed, it is conjectured that this particular set of modular relation

[2] These circuits are composed of bounded fan-in gates, as in NC^0, but additionally contain a single unbounded fan-in Mod_p gate.

problems might be sufficient to achieve unconditional separations between the aforementioned $\mathsf{AC}^0[p]$ classical circuit classes and less powerful quantum circuit classes, such as QNC^0. In parallel, separations against classes like TC^0 (or larger classes) have been achieved with different classes of problems but require the interactive setting [74].

We then switch gears and consider the class of constant-depth quantum circuits with an infinite size gateset and multi-qudit fanout.[3] In this setting, we show that for any prime p, we have the collapse of the constant-depth quantum hierarchy for quantum circuits with p-dimensional qudits, extending the qubit case [77].

Result 4. *(Informal; see Theorem 5.5.3) For all prime p, there exists an gate set on qudits of dimension p such that constant depth quantum circuits with auxiliary unbounded fan-in* Mod_p *gates implement quantum threshold gates, i.e.,* $\text{i-}\mathsf{QNC}^0_p[p] = \text{i-}\mathsf{QTC}^0_p$.[4]

This result establishes computational equivalence across shallow-depth quantum circuits when implemented on systems over prime-dimensional Hilbert spaces. Specifically, our quantum circuits are constructed using Fourier transforms over Abelian groups and this approach allows the creation of circuits for functions over finite fields of the form $\mathbb{F}_p^n \mapsto \mathbb{F}_p$. These circuits allow the execution of multi-qubit controlled logical operations.
Subsequently, we show that constant-depth quantum circuits implemented in higher prime dimensions can be implemented with qubits with a multiplicative overhead. This highlights that for any algorithm falling within these classes, employing qudits instead of qubits offers no substantial computational benefits—except perhaps for more straightforward hardware implementations and better error-correcting schemes in Hilbert spaces of varying dimensions.

Result 5. *(Informal; see Theorem 5.5.4). For all primes p and q, any constant depth quantum circuit with gates on qudits of dimension p and quantum threshold gates can be replicated by a constant-depth quantum circuit over qubits with auxiliary* Mod_q *gates,* $\text{i-}\mathsf{QTC}^0_p \subseteq \text{i-}\mathsf{QNC}^0_2[q]_\mathsf{c}$.

This result is achieved by mapping qudit operations onto tensor products of qubits. The qudit operations are produced using exact unitary decomposition methods, as previously demonstrated [107, 108]. Specifically, this theorem implies that the qubit-specific hierarchy can be collapsed using any arbitrary classical modular prime gate — a capability previously believed to require either quantum fanout or multi-qubit parity gates. Consequently, we have that $\text{i-}\mathsf{QNC}^0[p]_\mathsf{c} = \mathsf{QACC}^0$ holds for all primes p. This significantly contrasts with classical analog circuit classes, where varying modular gates results in different complexity classes.
From a practical standpoint, it also demonstrates that quantum circuits with fixed-depth qubits, when augmented with unbounded fan-in classical prime modular operations Mod_p, can implement the quantum subroutines of algorithms like factoring, which are conjectured to convey exponential quantum advantage [53, 99]. Moreover, constant-depth circuits with these gate sets can be realized with standard gate sets in logarithmic depth, generating still fairly shallow quantum circuits. In addition, using classical modular gates makes implementation practically easier, as these gates can be efficiently realized, while their quantum analogs are technically more difficult to implement.
Finally, as future work, we propose extending Result 5 to show that any $\text{i-}\mathsf{QTC}^0_p$ circuit can be simulated by $\text{i-}\mathsf{QNC}^0_q$ circuits with additional classical modular gates, for arbitrary primes p and q. This would establish the equivalence between all constant-depth quantum circuits over qudits of different prime dimensions, which could be advantageous for hardware implementations that favor a particular gate sets over a prime-dimensional Hilbert space.

[3] For any prime p, multi-qudit fanout for p-dimensional qudits is equivalent to Mod_p gates.

[4] As we describe later, in our notation, the prefix $i-$ denotes that the gate set is infinite, and the subscript p denotes that the qudits have dimension p. Additionally, $[p]$ means access to a quantum and $[p]_c$ to a classical unbounded fan-in modular p gate.

5.2.2 Chapter organization

Section 5.4 delves into the study of shallow-depth quantum circuits with finite-size gate sets. In this section, we elucidate how quantum circuits can solve modular relation problems, which we introduce in Section 5.3. Additionally, we establish the $\mathsf{AC}^0[p]$ lower bounds for the circuit classes addressing these problems and conclude the section by defining the $\mathsf{NC}^0[p]$ classical upper bounds for the relation problems under consideration. In Section 5.5, we shift our focus to constant-depth quantum circuits over qudits. This section explores the implications of employing infinite-size gate sets and unbounded fan-in modular gates, demonstrating both hierarchical collapses and computational equivalences between qubit and qudit formulations.

We assume readers have a basic understanding of qudit quantum computation, types of computational problems (such as decision, search, and sampling), and classical constant-depth circuits. For additional information, refer to Section 5.3 for an overview.

5.3 Preliminaries

5.3.1 Classical circuit classes

Circuit classes form a crucial aspect of computational theory, helping to categorize and understand the computational capabilities of various models. Certain classical circuit classes emerge as especially relevant when studying constant-depth quantum circuits. This subsection will list and define the main classes of interest while also summarizing pertinent theorems associated with them.

Definition 45 (NC^k class)**.** *The* NC^k *class is defined as the computational problems that can be solved in poly-logarithmic depth* $log(n)^k$ *using polynomially-many gates. The gate set for this class comprises the following bounded fan-in gates,*

$$\left\{ \boxed{\text{AND}} \ , \ \boxed{\text{OR}} \ , \ \boxed{\text{NOT}} \right\} .$$

Definition 46 (AC^k class)**.** *The* AC^k *class is defined as the computational problems that can be solved in poly-logarithmic depth* $log(n)^k$ *using polynomial-many gates. The gate set for this class comprises the following unbounded fan-in gates,*

$$\left\{ \boxed{\text{AND}} \ , \ \boxed{\text{OR}} \ , \ \boxed{\text{NOT}} \right\} .$$

Definition 47. ($\mathsf{AC}^k[p]$ class)*. The* $\mathsf{AC}^k[p]$ *class is defined as the computational problems that can be solved in poly-logarithmic depth* $log(n)^k$ *using polynomial-many gates. The gate set for this class comprises the following unbounded fan-in gates,*

$$\left\{ \boxed{\text{AND}} \ , \ \boxed{\text{OR}} \ , \ \boxed{\text{Mod}_p} \ , \ \boxed{\text{NOT}} \right\} ,$$

with Mod_p *gate defined as follows,*

$$\mathsf{MOD_p}(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i \mod p = 0 \\ 0 & \text{if } \sum_{i=1}^{n} x_i \mod p \neq 0 \end{cases} . \tag{5.3.1}$$

The Razborov-Smolensky separations between all the $\mathsf{AC}^0[p]$ classes will be fundamental to establishing new separations between these and quantum constant depth circuit classes.

Theorem 5.3.1. *[19] Let p be a prime number, and q is not a power of p, then computing Mod_p with a depth d, $\mathsf{AC}^0[q]$ circuit, requires $\exp\big(\Omega(n^{1/2d})\big)$ size.*

Furthermore, it should be noted that this theorem implies that when we construct a graph using the $\mathsf{AC}^0[q]$ classes as vertices and draw edges between pairs of classes that are not subsets of each other, we obtain a complete graph. Thus, for each edge, at least two problems distinguish one class from the other (see Figure 28).

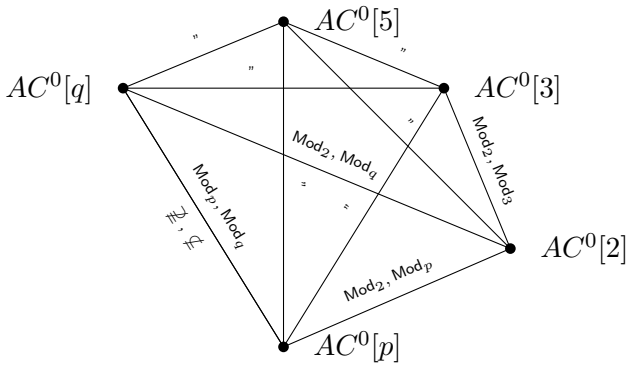


Figure 28: Graph representation of $\mathsf{AC}^0[p]$ classes with p prime. Each edge denotes that the connected classes are distinct, without one being a subset of the other. Along the edges, examples highlight problems that are in one class but excluded from the other.

Definition 48 (TC^k class)**.** *The TC^k class is defined as the computational problems that can be solved in poly-logarithmic depth $log(n)^k$ using polynomial-many gates. The gate set for this class comprises the following unbounded fan-in gates,*

$$\Big\{ \boxed{\text{AND}} \ , \ \boxed{\text{OR}} \ , \ \boxed{\mathsf{TH}_t} \ , \ \boxed{\text{NOT}} \Big\} \ ,$$

with TH_k gate defined as follows,

$$\mathsf{TH_t}(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i \geq t \\ 0 & \text{if } \sum_{i=1}^{n} x_i < t \end{cases}. \tag{5.3.2}$$

We introduce a class that is not standard in circuit complexity literature but holds significance for the discussions and results presented in this text.

Definition 49. ($\mathsf{NC}^0[p]$ *class*). *The $\mathsf{NC}^k[p]$ class is defined as the computational problems that can be solved in poly-logarithmic depth $log(n)^k$ using polynomial-many gates. The gate set for this class comprises the bounded fan-in* AND, OR, NOT *gates, and a single unbounded fan-in* Mod p *gates.*

A considerable number of important results in computational complexity pertain to the circuit classes we have defined, as well as their interrelations [330, 19, 328, 43]. Consequently, we will present the established containments for these classes,

$$\mathsf{NC}^0 \subsetneq \mathsf{AC}^0, \mathsf{NC}^0[p] \subsetneq \mathsf{AC}^0[p] \subsetneq \mathsf{TC}^0 \subseteq \mathsf{NC}^1. \tag{5.3.3}$$

This knowledge is crucial for comprehending the capabilities of each circuit class and discerning the position of quantum circuit classes relative to them.
Lastly, we will present the Vazirani-XOR lemma, which is of great interest to the study of parallel repeated games. This lemma is of particular interest for the analysis of constant-depth circuits and will be used to improve our classical hardness results.

Lemma 56. *(Vazirani's XOR Lemma [134]). Let D be a distribution on $\mathbb{F}_2^m$ and χ_S denote the parity function on the set $S \subseteq [m]$, defined as $\chi_S(x) = \oplus_{i \in S} x_i$. If $|\mathbb{E}_{x \in D}[(-1)^{\chi_S(x)}]| \leq \epsilon$ for every nonempty subset $S \subseteq [m]$, then D is $\epsilon \cdot 2^{m/2}$-close in statistical distance to the uniform distribution over $\mathbb{F}_2^m$.*

5.3.2 Quantum computation with qudits

We first review the basic operations on qudits, and then explain the qudit MBQC model.

Qudit operations

In this preliminary section, we present the qudit gates that will be employed in subsequent sections of the text to construct the quantum circuits essential for our main proofs.
The first qudit operation that we will define will be a generalized version of the qubit X gate.

Definition 50 (X qudit gate)**.** *The* X *gate for qudits is defined as follows,*

$$\mathsf{X}_d |m\rangle := |m+1 \mod d\rangle \ . \tag{5.3.4}$$

Then a rotation about the Z axis will be defined as follows,

Definition 51 (Rotation Z qudit gate)**.** *The rotation Z operator for qudits is defined as follows,*

$$\mathsf{R}_Z^d(\phi) := \sum_{j=0}^{d-1} e^{i(1-sgn(d-1-j)\phi)} |j\rangle \langle j| \ . \tag{5.3.5}$$

In addition, we will define and make use of a more complex rotation Z operator.

Definition 52 (Generalized rotation Z qudit gate)**.** *The generalized rotation Z operator for qudits is defined as follows,*

$$\mathsf{GR}_Z^d(\phi) := \sum_{j=0}^{d-1} e^{i(\phi j)} |j\rangle \langle j| \ . \tag{5.3.6}$$

The last single qudit operation will be the qudit Hadamard gate.

Definition 53 (Fourier gate)**.** *The Fourier gate is defined as follows*

$$\mathsf{F}_d |m\rangle = \frac{1}{\sqrt{d}} \sum_{n \in \mathbb{Z}_d} w^{mn} |n\rangle \,, \tag{5.3.7}$$

with $w = e^{i\frac{2\pi}{d}}$.

Now the respective two qudit gates of interest are the qudit CNOT gate and the controlled rotation Z gate.

Definition 54 (SUM gate)**.** *The* SUM *gate is defined as follows,*

$$\mathsf{SUM} |n\rangle |m\rangle = |n\rangle |n+m\rangle \,. \tag{5.3.8}$$

Definition 55 (Controlled generalized rotation Z qudit gate)**.** *The* CGR_Z^d *gate is defined as follows,*

$$\mathsf{CGR}_Z^d \left|n\right\rangle \left|m\right\rangle = \left|n\right\rangle \left(\mathsf{GR}_Z^d\right)^n \left|m\right\rangle . \tag{5.3.9}$$

The last two multi-qudit operations are the following.

Definition 56 (fanout_p^{-1})**.** *The* fanout_p^{-1} *gate is defined as follows,*

$$\mathsf{fanout}_p^{-1} \left|x_0, x_1, \ldots, x_n\right\rangle := \left|x_0, (p - x_0) + x_1 \mod p, \ldots, (p - x_0) + x_n \mod p\right\rangle , \tag{5.3.10}$$

and can be implemented with $(\mathsf{fanout}_p)^{p-1}$*.*

Definition 57 (qMOD_p^{-1})**.** *The* qMOD_p^{-1} *gate is defined as follows,*

$$\mathsf{qMOD}_p^{-1} \left|x_0, x_1, \ldots, x_n\right\rangle := \left|x_0 + (p - x_1 + \ldots + x_n) \mod p, x_1, \ldots, x_n \mod p\right\rangle , \tag{5.3.11}$$

and can be implemented with $(\mathsf{qMOD}_p)^{p-1}$*.*

Qudit MBQC

The measurement-based quantum computation (MBQC) model was first introduced in [331, 332]. Unlike the standard circuit model, MBQC operates by performing measurements on universal quantum resource states [333]. Although these two models have a very different nature, they are computationally equivalent modulo a polynomial overhead. In this work, the generalization of MBQC to qudits [334] is especially relevant, since the circuits arising from this model exactly characterize the circuit classes, which we will analyze in detail.

In qudit MBQC, quantum algorithms are commonly referred to as *measurement patterns*. These patterns are defined by both a resource state and a distinct set of measurement operators for each qudit in the resource state. A noteworthy aspect is the existence of adaptivity, defined by the interdependencies between the measurement operators and their resulting outcomes. These interdependencies guide the information flow within the quantum state, which experiences decoherence due to the measurements; this guiding principle is termed 'flow.' This flow can be succinctly represented as a directed graph, denoted as $G = (V, E)$, where the vertices are associated with the measurement operators, and the update conditions between operators and results are represented by the edges of the graph [156]. The steps for executing a *measurement pattern* are detailed below,

1. Initialize a set of qudits, the number of which matches the count of vertices $|V|$ in the graph, all set to the state $F_p \left|0\right\rangle$;
2. Apply the entangling gates E_{v_i,v_j} between the necessary qudits to produce the resource state.
3. For each temporal layer t_i in the flow, sequentially:
   (a) Implement the Pauli corrections for each vertex in t_i, based on prior measurement outcomes $M_{j<i}$ as dictated by the edges of graph G.
   (b) Conduct the qudit measurement $M_{t_i}^b(\theta)$ within the same temporal layer t_i of the flow graph, for every time step i.

This model has been both the focus and inspiration for various works in studying the complexity of constant-depth quantum circuits [74, 335, 51]. However, a detailed analysis exploring the computational differences within the model, especially concerning the utilization of higher-dimensional resource states and measurement operators, remained to be delved into [334].

5.3.3 Quantum circuit classes

The initial circuit class under examination comprises a gate set defined by a finite collection of gates and is restricted to single qudit control gates, encompassing standard universal gate sets like the Clifford+T gate set.

Definition 58 (QNC^0_d)**.** *For $d \in \mathbb{N}$, let $\mathcal{H}^d$ denote a d-dimensional Hilbert space. $\mathcal{G}_d$ being a finite set of quantum gates that operate on $\mathcal{H}^d$, while $\mathcal{C}_d$ represents a fixed number of gates that act on one qudit, controlled by another qudit. We define QNC^0_d as the class of quantum circuits with constant depth and polynomial size, employing a gate set compiled from the combination of $\mathcal{G}_d$ and $\mathcal{C}_d$.*

The subsequent class to be analyzed will supplement the finite number of qudit gates with their multi-qudit controlled variants. This introduces an alternative characterization for the quantum counterpart of AC^0 that has not been addressed in prior literature and guarantees a fairer comparison between the quantum and classical constant-depth circuit classes.

Definition 59 (QAC^0_d)**.** *For $d \in \mathbb{N}$, let $\mathcal{H}^d$ denote a d-dimensional Hilbert space. Let $\mathcal{G}_d$ represent a finite set of gates acting on $\mathcal{H}^d$, which are permitted unbounded classical fanout—defined as the ability to replicate a gate's output indefinitely, allowing its use as input for numerous other gates—when these gates are characterized as maps from basis states to basis states with classical inputs. Let $\mathcal{T}_d$ denote a fixed number of multi-qudit controlled gates acting on a single qudit. We define QAC^0_d as the class of quantum circuits with a constant depth and polynomial size, using a gate set drawn from the combination of $\mathcal{G}_d$ and $\mathcal{T}_d$.*

In this work, we also consider the complexity classes i-QNC^0_d and i-QAC^0_d. These variants of QNC^0_d and QAC^0_d incorporate the full set of one-qudit unitaries.

Definition 60 (i-QNC^0_d)**.** *For a given positive integer d, let $\mathcal{U}_d$ represent the set of all unitary operators on a d-dimensional Hilbert space $\mathcal{H}^d$, and let $\mathcal{C}_d$ denote the set of single qudit controlled gates from $\mathcal{U}_d$. We define i-QNC^0_d as the class of quantum circuits with constant depth and polynomial size, using a gate set drawn from the combination of $\mathcal{U}_d$ and $\mathcal{C}_d$.*

Definition 61 (i-QAC^0_d)**.** *For a given positive integer d, let $\mathcal{U}_d$ represent all unitary operators on a d-dimensional Hilbert space $\mathcal{H}^d$, and let $\mathcal{T}^k_d$ denote a multi-qudit controlled gate selected from $\mathcal{U}_d$. We define i-QAC^0_d as the class of quantum circuits with a constant depth and polynomial size, using a gate set drawn from the combination of $\mathcal{U}_d$ and $\mathcal{T}^k_d$ with unbounded classical fanout.*

Additionally, we consider access to quantum advice, which is a quantum state that depends on the input size but not on the input. In particular, we consider $\mathsf{QNC}^0_d/\mathsf{qpoly}$, which consists of QNC^0_d with a polynomial-size quantum advice.
We also examine extensions of certain constant-depth circuit classes, including unbounded fan-in modular operations. For example, the class i-$\mathsf{QNC}^0_d[p]$ integrates i-QNC^0_d circuits with an additional quantum modular gate defined as follows,

$$\mathsf{qMOD}_p\left|x_0, x_1, \ldots, x_n\right\rangle := \left|x_0 + (x_1 + \ldots + x_n) \mod p, x_1, \ldots, x_n \mod p\right\rangle. \tag{5.3.12}$$

Finally, we use the suffix $[p]_c$ to denote the inclusion of a classical unbounded fan-in modular gate Mod_p in these circuit classes.

5.3.4 Modular Relation Problems and Supporting Materials

We now define the notion of modular relation problems.

Definition 62 (Modular relation problem)**.** *The modular relations problem* $\mathcal{R}^m_{q,p} : \{0,1\}^n \mapsto \{0,1\}^m$ *is defined as*

$$\mathcal{R}^m_{q,p}(x) = \left\{ y \;\middle|\; y \in \mathbb{F}_2^m,\ |y| \pmod q = 0 \textit{ iff } |x| \pmod p = 0 \right\}. \tag{5.3.13}$$

This class of problems subsumes the relation problem used in [18, 27], which uses $\mathcal{R}^{o(n^2)}_{2,4}$ for the NC^0 and the AC^0 separations, and [27], that uses $\mathcal{R}^{o(n^2)}_{2,3}$ for their $\mathsf{AC}^0[2]$ separation.
In this work, we also consider the parallel repetitions of these relation problems.

Definition 63 (Parallel-k modular relation problem)**.** *A modular relation problems* $k\text{-}\mathcal{R}^m_{q,p} : \{0,1\}^{n\cdot k} \mapsto \{0,1\}^{m\cdot k}$ *are defined as,*

$$k\text{-}\mathcal{R}^m_{q,p}(x_1,...,x_k) = \left\{ (y_1,...,y_k) \;\middle|\; \forall x_i \in \mathbb{F}_2^n,\ y_i \in \mathbb{F}_2^m,\ |y_i| \pmod q = 0 \textit{ iff } |x_i| \pmod p = 0 \right\}.$$

Finally, we also define two key orthogonal bases for qudit states that will be used in our works. The first one is an extension of the qubit X-basis for qudits.

Definition 64 (Qudit orthogonal X-basis)**.** *The qudit orthogonal X-basis can be described by the set of states,*

$$|X^m_d\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \omega^{j\cdot m} |j\rangle \tag{5.3.14}$$

with $\omega = e^{\frac{i2\pi}{d}}$ *and* $m \in \{0,1,...,d-1\}$.

The second orthogonal basis is tailored specifically for qudit GHZ states and will be necessary for the analysis of the quantum circuits we will present.

Definition 65 (Qudit-GHZ orthogonal X-basis)**.** *The qudit-GHZ orthogonal X-basis can be defined by the set of states,*

$$\left|\mathsf{GHZ}^m_{d,n}\right\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \omega^{j\cdot m} |j\rangle^{\otimes n} \tag{5.3.15}$$

with $\omega = e^{\frac{i2\pi}{d}}$ *and* $m \in \{0,1,...,d-1\}$.

Finally, we state two technical lemmas regarding these bases that are needed in our results.

Lemma 57. $\langle X^m_d | X^n_d \rangle = \langle \mathsf{GHZ}^m_d | \mathsf{GHZ}^n_d \rangle = \delta_{m,n}$.

Proof. All the elements of the qudit orthogonal X-basis can be represented as,

$$|X^m_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \omega^{i\cdot m} |i\rangle \ = \mathsf{F}_d |m\rangle \ . \tag{5.3.16}$$

This yields that,

$$\langle X^m_d | X^n_d \rangle = \langle m | \mathsf{F}_d \mathsf{F}^\dagger_d |n\rangle . \tag{5.3.17}$$

Combining this with the fact that the $\mathsf{F}^\dagger_d \mathsf{F}_d |m\rangle = |m\rangle$, we obtain that $\langle X^m_d | X^n_d \rangle = \delta_{m,n}$. The same result holds for the qudit-GHZ orthogonal X-basis, given that the bases are simply a tensor product over n elements of the same bases as in the qudit X-basis. □

Lemma 58. $\mathsf{F}^{\otimes n}_d \left|\mathsf{GHZ}^m_{d,n}\right\rangle = \frac{1}{\sqrt{d^{n-1}}} \sum_{\substack{x\in\mathbb{F}^n_d \\ |x| \bmod d = -m}} |x\rangle$.

Proof. By considering the effect of the operator on the basis states we determine that,

$$\mathsf{F}_d^{\otimes n} \left|\mathsf{GHZ}_d^m\right\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \left(\omega^{i\cdot m} \mathsf{F}_d^{\otimes n} \left|i\right\rangle^{\otimes n}\right) \tag{5.3.18}$$

$$= \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \omega^{i\cdot m} \Big(\frac{1}{\sqrt{d^n}} \sum_{x\in\{0,1,\ldots,d-1\}^n} \omega^{i\cdot|x|} \left|x\right\rangle\Big) \tag{5.3.19}$$

$$= \frac{1}{\sqrt{d^{n+1}}} \sum_{x\in\{0,1,\ldots,d-1\}^n} \sum_{i=0}^{d-1} \omega^{i\cdot(m+|x|)} \left|x\right\rangle \tag{5.3.20}$$

with $|x| = \sum_{i=0}^{n-1} x_i \mod d$. This implies that for all the inputs x for which $m + |x| \mod d = 0$ the value of $\omega^{i\cdot(m+|x|)} = 1$. In contrast, for inputs x for which $m + |x| \mod d \neq 0$ we will use the geometric progression to demonstrate first that the sum of all roots of unity is equal to zero,

$$\sum_{i=0}^{d-1} \omega^i = \frac{1-\omega^d}{1-\omega} = 0\,. \tag{5.3.21}$$

Then, with one more step, we can show that the same is true for the amplitudes of the inputs considered,

$$\sum_{i=0}^{d-1} \omega^{i\cdot(m+|x|)} = \frac{1-\omega^{d\cdot(m+|x|)}}{1-\omega^{m+|x|}} = 0\,. \tag{5.3.22}$$

Finally, this implies that only for values that $m + |x| \mod d = 0$ the amplitude of the respective basis states $\left|x\right\rangle$ are non-zero. This produces exactly the states,

$$\frac{1}{\sqrt{d^{n-1}}} \sum_{x\in X} \left|x\right\rangle,\ with\ X = \left\{y \mid y \in \{0,1,\ldots,d-1\}^n,\ |y| \mod d = -m\right\}. \tag{5.3.23}$$

□

5.4 Shallow depth quantum circuits with finite-size gate sets

In this section, we show new separations between quantum shallow circuits and classical ones. More precisely in Section 5.4.1, we prove the quantum upper bounds, and in Section 5.4.2, we prove the classical lower bounds. We combine these results in Section 5.4.3 in order to achieve our new separations. Finally, in Section 5.4.4, we discuss the limitations of our techniques.

5.4.1 Quantum upper bounds

In this subsection, we introduce lower bounds for the probability that QNC^0 circuits with quantum advice and QAC^0 circuits can solve the modular relation problems.

Shallow depth quantum circuits with advice

In this subsection, we will introduce quantum circuits with a quantum advice that compute $\mathcal{R}_{q,p}^{q\cdot n}$ with a one-sided error.

Theorem 5.4.1. *For all fixed and distinct primes p and q, $\mathcal{R}_{q,p}^{q\cdot n}$ can be solved in* $\mathsf{QNC}_q^0/\mathsf{qpoly}$ *with one-sided error at most* $\frac{1-\cos\left(\frac{2\pi}{p}\right)}{q}$.

Proof. The circuit for $\mathcal{R}_{q,p}^{q\cdot n}$ problem is described in Figure 29.
The initial state is $|x\rangle \left|\mathsf{GHZ}_{q,n}^0\right\rangle$, where the GHZ state is the quantum advice of the circuit. Then, we apply a layer of n control-Z rotations on the n qubits with a phase equal to $\frac{2\pi}{p}$. The i-th bit of the input x acts as the control qubit to the control-Z with the target being the i-th qubit of the GHZ state.
The first register is then

$$\left|\psi^{(1)}\right\rangle = \frac{1}{\sqrt{q}}\Big(\sum_{i=0}^{q-2}|i\rangle^{\otimes n} + e^{i\frac{2\pi|x|}{p}}|q-1\rangle^{\otimes n}\Big) = \sum_{i=0}^{q-1} c_i \left|\mathsf{GHZ}_{q,n}^i\right\rangle, \tag{5.4.1}$$

with the second equality being the decomposition of $\left|\psi^{(1)}\right\rangle$ in the qudit-GHZ orthogonal X-basis.
If we apply $F_q^{\otimes n}$ to $\left|\psi^{(1)}\right\rangle$, we have

$$\left|\psi^{(2)}\right\rangle = \mathsf{F}_q^{\otimes n}\left|\psi^{(1)}\right\rangle = \sum_{i=0}^{q-1} c_i\,(\mathsf{F}_q^{\otimes n}\left|\mathsf{GHZ}_{q,n}^i\right\rangle) = \sum_{i=0}^{q-1} c_i \sum_{\substack{z\in[q]^n \\ |z| \mod q=-i}} |z\rangle,$$

where the last equality comes from Lemma 58.
Our goal is to compute the probability of measuring this state and get an output y such that $|y| \pmod p = 0$. For that, notice that c_0 is:

$$\left\langle\mathsf{GHZ}_{q,n}^0\middle|\psi^{(1)}\right\rangle = \frac{1}{\sqrt{q}}\sum_{i=0}^{q-1}\langle i|^{\otimes n}\,\frac{1}{\sqrt{q}}\Big(\sum_{i=0}^{q-2}|i\rangle^{\otimes n} + e^{i\frac{2\pi|x|}{p}}|q-1\rangle^{\otimes n}\Big) \tag{5.4.2}$$

$$= \frac{q-1+\cos\left(\frac{2\pi|x|}{p}\right)}{q}. \tag{5.4.3}$$

We then divide that analysis into two cases:

1. If $|x| \mod p = 0$, $(c_0)^2$ equals 1. Consequently, measuring $|\psi_2\rangle$ leads to a string y such that $|y| \mod q = 0$.

2. If $|x| \mod p \neq 0$, we have $c_0^2 \leq \frac{q-1+\cos\left(\frac{2\pi|x|}{p}\right)}{q}$. This implies that with a probability of least $\frac{1-\cos\left(\frac{2\pi}{p}\right)}{q}$, it holds that $|y| \mod q \neq 0$.

The remaining step consists of mapping $y \in [q]^n$ to bit-strings $y' \in \{0,1\}^{q\cdot n}$ such that the Hamming weight is preserved. This can be done with extra ancilla qubits by mapping each $y_i \in [q]$ to the bit-string $1^{y_i}0^{q-y_i}$. We denote the circuit that performs such an operation as $C(q,2)$. □

Shallow depth quantum circuits with Toffoli gates

We show now that quantum circuits of constant depth built with multi-qudit controlled operations along with intermediate measurements can generate qudit-GHZ states, demonstrating that QAC_q^0 (without advice) can solve $\mathcal{R}_{q,p}^{q\cdot n}$ with a one-sided error.
To achieve these qudit GHZ states, we use a specialized entanglement structure built upon "poor-man's cat states," as referenced in [27, 263, 336]. Specifically, poor-man's cat states are formulated using graphs

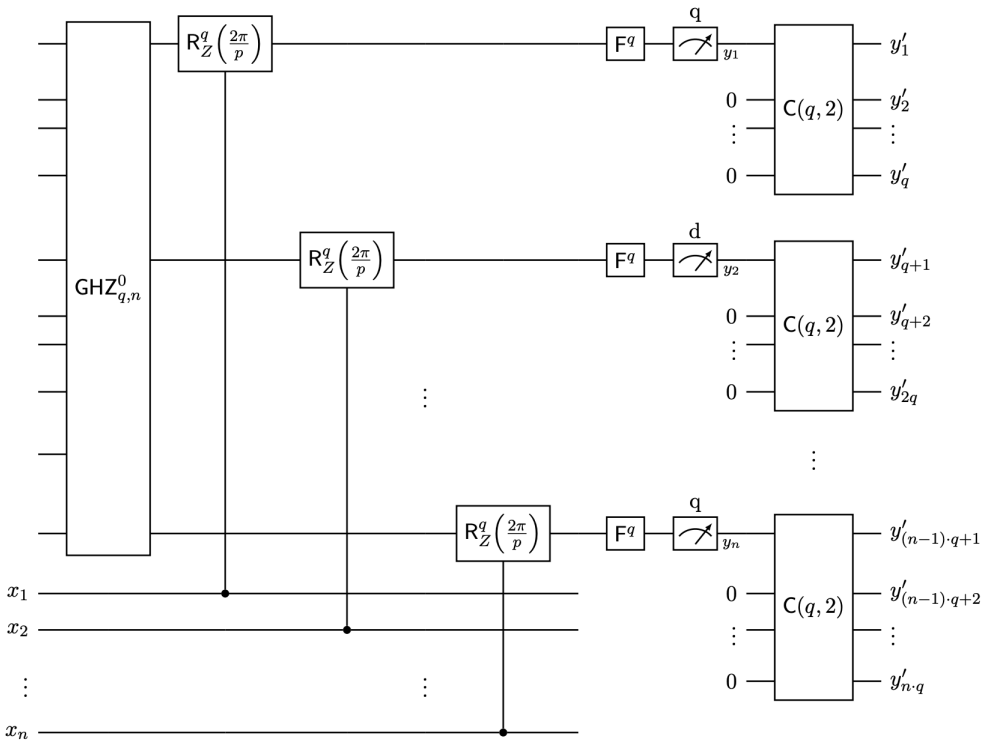


Figure 29: Parameterized quantum circuit class for values q, p, and n, incorporating the advice quantum state $\left|\mathsf{GHZ}^0_{q,n}\right\rangle$, which solves the $\mathcal{R}^{q\cdot n}_{q,p}$ with bounded one-sided error.

where the number of vertices is equal to the state's qubit count, and the edge configuration guarantees a unique path between all vertices. In our work, we fix this graph to balanced binary trees $\mathcal{B}_n$ and extend these quantum states to higher dimensions.

Definition 66 (Binary tree-structured poor-man's qudit state)**.** *Let q be a prime. We consider the balanced binary tree $\mathcal{B}_n = (V, E)$, where $V = \{1, ..., n\}$ and $E = \{\{i, \lfloor \frac{i}{2} \rfloor\} : i \in \{2, ..., n\}\}$. As an example, we depict $\mathcal{B}_6$ in Figure 30.*
We define a binary tree-structured poor-mans qudit state $|\mathsf{BPM}_{q,n}\rangle$ as follows

$$|\mathsf{BPM}_{q,n}\rangle = \frac{1}{\sqrt{q^n}} \sum_{e_2,...,e_n,v_1 \in \mathbb{F}_q} |e_2, ..., e_n\rangle |v_1, v_2, v_3, \ldots, v_n\rangle , \tag{5.4.4}$$

where each v_i, $i \in \{2, ..., n\}$ is recursively defined as

$$v_i = e_i - v_{\lfloor \frac{i}{2} \rfloor}. \tag{5.4.5}$$

We now show that these states can be created using constant-depth quantum circuits.

Lemma 59. *For $n \in \mathbb{N}$, $|\mathsf{BPM}_{q,n}\rangle$ can be prepared in QNC^0_q.*

Proof. For creating this state, we start with a single qudit for each vertex and edge in $\mathcal{B}_n$, i.e., we have $|V| + |E|$ qudits, where the i-th qudit, for $1 \leq i \leq n-1$ corresponds to the edge $\{i+1, \lfloor \frac{i+1}{2} \rfloor\}$ and the i-th qudit, for $n \leq i \leq 2n-1$ corresponds to the vertex $i - n + 1$. We label qudits by their corresponding vertices/edges.
We start by applying the Fourier gate F_q to each of the vertex qudits and then, for edge $e = \{u, v\} \in E$

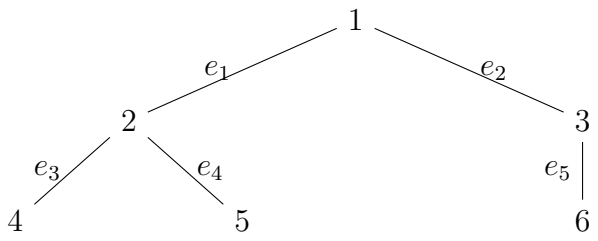


Figure 30: Representation of $\mathcal{B}_6$.

we apply a $\mathsf{SUM}_{u,e}$[5] and $\mathsf{SUM}_{v,e}$. The overall state after these operations is:

$$\sum_{v_1,\ldots,v_n\in\mathbb{F}_q} \frac{1}{\sqrt{q^n}} \left|(v_1+v_2),(v_1+v_3)\ldots,(v_n+v_{\frac{n}{2}})\right\rangle \otimes |v_1,v_2,v_3,\ldots,v_n\rangle \tag{5.4.6}$$

$$= \sum_{e_1,v_1,v_3\ldots,v_n\in\mathbb{F}_q} \frac{1}{\sqrt{q^n}} \left|e_1,(v_1+v_3)\ldots,(v_n+v_{\frac{n}{2}})\right\rangle \otimes |v_1,e_1-v_1,v_3,\ldots,v_n\rangle \tag{5.4.7}$$

$$= \sum_{e_1,e_2,v_1,v_4\ldots,v_n\in\mathbb{F}_q} \frac{1}{\sqrt{q^n}} \left|e_1,e_2\ldots,(v_n+v_{\frac{n}{2}})\right\rangle \otimes |v_1,e_1-v_1,e_2-v_1,v_4,\ldots,v_n\rangle \tag{5.4.8}$$

... (5.4.9)

$$\sum_{e_1,\ldots,e_{n-1},v_1\in\mathbb{F}_q} \frac{1}{\sqrt{q^n}} |e_1,\ldots,e_{n-1}\rangle \otimes \left|v_1,e_1-v_1,e_2-v_1,\ldots,e_{n-1}-v_{\frac{n}{2}}\right\rangle \tag{5.4.10}$$

$$= |\mathsf{BPM}_{q,n}\rangle\,, \tag{5.4.11}$$

where all operations are $\pmod q$ and all the equalities hold by rewriting the elements in $\mathbb{F}_q$.
Finally, given that the F_q gates are all applied in parallel to the vertex qudits, and the chromatic number of a balanced binary tree is equal to 3, all the $\mathsf{SUM}_{i,j}$ gates can be parallelized in order to be implemented in constant depth. □

We now show that these states can be easily converted to GHZ states.

Lemma 60. *There is a* QAC^0_q *circuit* C_n *that maps* $|\mathsf{BPM}_{q,n}\rangle$ *to the n-qudit state* $\left|\mathsf{GHZ}^0_{q,n}\right\rangle$.

Proof. In the circuit C_n, we first measure the edge qudits leading to the output string $e_1\ldots e_{n-1}\in\mathbb{F}_q^{n-1}$. The collapsed quantum state is

$$|e_1,\ldots,e_{n-1}\rangle \otimes \frac{1}{\sqrt{q}} \sum_{v_1\in\mathbb{F}_q} |v_1,v_2,\ldots,v_n\rangle\,. \tag{5.4.12}$$

where $v_i = e_i - v_{\lfloor\frac{i}{2}\rfloor}$. Notice that by opening this recurrence relation, we have that

$$v_i = \sum_{j=0}^{depth(i)-1} (-1)^{depth(i)-depth(q^j(i))} e_{q^j(i)} + (-1)^{depth(i)+1} v_1, \tag{5.4.13}$$

where $depth(i)$ denotes the depth of node i in the tree and $q^j(i)$ denotes the j-th predecessor of i in the tree.

[5] The $\mathsf{SUM}_{i,j}$ gate serves as the qudit equivalent of the CNOT gate and operates according to $\mathsf{SUM}_{\mathrm{i,j}}\,|n\rangle_i\,|m\rangle_j = |n\rangle\,|n+m\rangle$.

We notice that for each i, the value $c_i := \sum_{j=0}^{depth(i)-1}(-1)^{depth(i)-depth(q^j(i))}e_{q^j(i)}$ can be computed by $poly(n)$-size AC^0 circuits on input, since it depends on at most $\log(n)$ values e_j. In this case, a $poly(n)$-size AC^0 exists that computes $c_1...c_n$. Given these corrections, we can apply the gate X^{c_i} to the i-th vertex qubit (which is a classically-controlled X gate), and the resulting state is:

$$|e_1, ..., e_{n-1}\rangle \otimes \frac{1}{\sqrt{q}} \sum_{v_1 \in \mathbb{F}_q} \bigotimes_{i=1}^{n} \left|(-1)^{depth(i)} v_1\right\rangle . \tag{5.4.14}$$

Finally, we can apply the following operation to the qudits corresponding to the vertex of odd depth: $V|m\rangle = |p-m\rangle$ and the result state is

$$|e_1, ..., e_{n-1}\rangle \otimes \frac{1}{\sqrt{q}} \sum_{v_1 \in \mathbb{F}_q} |v_1\rangle^{\otimes n} \qquad \square.$$

The reduction we previously discussed cannot be achieved using QNC^0 or i-QNC^0 circuits since they cannot implement the essential corrections needed to transition from the poor man's cat states to the GHZ states. In contrast, we explore the additional computational power of QAC^0 circuits and show that they can produce the qudit GHZ states through a process involving mid-circuit measurements. As a consequence, we can show the following.

Theorem 5.4.2. *For any distinct primes p and q, a QAC^0_q circuit can solve $\mathcal{R}^{q\cdot n}_{q,p}$ with one-sided error at most $\frac{1-\cos\left(\frac{2\pi}{p}\right)}{q}$.*

Proof. Combining Lemma 59 and Lemma 60, we obtain that QAC^0_q circuits can create qudit GHZ states of the type $\left|\mathsf{GHZ}^0_{q,n}\right\rangle$. In addition, all the QNC^0_q circuits of Theorem 5.4.1 are contained in QAC^0_q. Therefore, all modular relation problems of the form $\mathcal{R}^{q\cdot n}_{q,p}$ fixed prime numbers p and q can be solved by QAC^0 circuits. □

5.4.2 Classical lower bounds

In this subsection, we prove the classical lower bounds for $\mathcal{R}^{q\cdot n}_{p,q}$. For that, we reduce it to computing Mod_p function and use the Razborov-Smolensky separations (Theorem 5.3.1) for $\mathsf{AC}^0[q]$. Then, we consider the k-$\mathcal{R}^{q\cdot n}_{q,p}$ problem, and we show that $\mathsf{AC}^0[q]$ circuits can solve them with only a negligible probability at best by using the Vazirani-XOR lemma.

Lemma 61. *Let p and q be two distinct primes. Any depth d circuit in $\mathsf{AC}^0[q]$ that solves the $\mathcal{R}^{q\cdot n}_{q,p}$ problem with one-sided error probability at most $1/n^{\Omega(1)}$ has size $\exp\left(n^{1/(2d-\Theta(1))}\right)$.*

Proof. Let us suppose, for the purpose of a contradiction, that there exists a circuit C in $\mathsf{AC}^0[q]$ of depth $d-3$ and size equal to $\exp\left(n^{1/(2d-\Theta(1))}\right)$ that computes $\mathcal{R}^{q\cdot n}_{q,p}$, with one-sided error probability at most $\frac{1}{n^{o(1)}}$. We will assume here that C computes the yes-instance with probability 1, and the arguments follow analogously if C computed the no-instances with probability 1.
We will now prove that if such a C exists, then there exists a $\mathsf{AC}^0[q]$ circuit C' that computes the Mod_p function correctly with probability $1 - negl(n)$. C' consists of running $O(n^2)$ times the circuit C in parallel and then applying a Mod_q gate to each output string, and then applying a NOR gate to the output of this layer of Mod_q gates. Figure 31 provides a representation of C'.
If $|x| \mod p = 0$, after running C on x, the resulting strings y_i are such that $|y_i| \mod q = 0$. Thus, the Mod_q gates to each of these strings lead to output 0. Therefore, after the NOR gate, the result for this final bit will always be 1.

On the other hand, if $|x| \mod p \neq 0$, with probability $1/n^{o(1)}$, the strings y_i after running C have the desired property $|y_i| \mod q \neq 0$. Therefore, with the same probability, the Mod_q gate applied to this string will result in the Boolean value 1, and the probability of at least one of them being equal to 1 is at least $1 - negl(n)$. Therefore, the output of the NOR gate is 0 with probability $1 - negl(n)$.
This leads to a depth-d circuit and size $\exp\left(n^{1/(2d-\Theta(1))}\right)$, that computes Mod_p in $\mathsf{AC}[q]$, contradicting Theorem 5.3.1. Thus, circuit C with such properties cannot exist. □

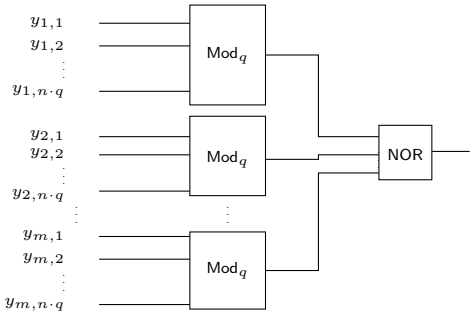


Figure 31: Representation of a classical $\mathsf{AC}^0[q]$ circuit, parameterized by q and n, that reduces a solution to the $\mathcal{R}_{q,2}^{q\cdot n}$ problem to a solution of $\mathsf{Mod}_2(x)$, where x is the initial input string of size n.

Subsequently, we show that polynomial-sized $\mathsf{AC}^0[q]$ circuits solve the parallel repetition of the modular relation problems with probability close to zero. To obtain this bound, we focus on the success of $\mathsf{AC}^0[q]$ circuits in computing the Mod_2 function with one-sided error.

Corollary 11. *The* Mod_2 *function can not be solved by any one-sided error* $\mathsf{AC}^0[q]$ *circuit with* q *prime* $q \neq 2$ *with a probability higher than* $1/2 + 1/n^{\Omega(1)}$.

Proof. Let's assume there exists an $\mathsf{AC}^0[p]$ circuit that can solve the Yes instances of the Mod_2 perfectly, then we know by Lemma 61 that the No instances can almost be solved correctly with probability $1/n^{o(1)}$. This implies any circuit of this type solves the Mod_2 function over a random input w.p. at most $1/2 + 1/n^{o(1)}$. □

Using all these tools, we can show that the class of $\mathsf{AC}^0[q]$ fails to compute completely parallel modular relation problems with a one-sided error for some parameter regimes.

Lemma 62. *Let* $q \neq 2$ *be a prime and* $k \in \Theta(\log(n))$. *No* $\mathsf{AC}^0[q]$ *circuit can solve* k-$\mathcal{R}_{q,2}^{q\cdot n}$ *on at least* $\frac{1}{2} + \frac{1-\cos\left(\frac{2\pi}{p}\right)}{2q}$ *fraction of the parallel instances with one-sided errors and a probability higher then* $n^{-\Omega(1)}$.

Proof. Let us suppose that there exists a circuit $C \in \mathsf{AC}^0[q]$, where $q \neq 2$, can solve the k-$\mathcal{R}_{q,2}^{q\cdot n}$ with the properties of the statement. Then, we can construct a circuit C' that solves the following search problem

$$k\text{-}\mathcal{D}_{2,2}(x_1, ..., x_k) = \left\{ (y'_1, ..., y'_k) \;\middle|\; \forall x_i \in \mathbb{F}_2^n,\ y'_i \in \mathbb{F}_2,\ y'_i + |x_i| \pmod 2 = 0 \right\}, \tag{5.4.15}$$

with the same success, where C' is equal to C with an extra layer of Mod_q gates, thereby transforming the outcome strings from k-$\mathcal{R}_{q,2}^{q\cdot n}$ into bits using the mapping $\mathsf{Mod}_q(y_i) = y'_i$.
The circuit C' induces a probability distribution $\mathcal{X}$ on $\{0,1\}^k$, corresponding on picking random $x_1, ..., x_k \in \{0,1\}^n$ and outputting k output bits as follows,

$$(|x_1| + C'_1(x_1, \ldots, x_n) \pmod 2, ..., |x_k| + C'_k(x_1, \ldots, x_n) \pmod 2). \tag{5.4.16}$$

Our goal now is to prove that $\mathcal{X}$ is close to a uniform distribution $\{0,1\}^k$ on half of its bits, either for all x_i with even or odd Hamming weight, if we assume one-sided error. For that, we notice that we can describe for all $S \subseteq [k]$ the following expectation value as

$$\left|\mathbb{E}_{z\in\mathcal{X}}[(-1)^{\chi_S(z)}]\right| = \frac{1}{2^{|S|\cdot n}} \sum_{x_1,\ldots,x_k\in\mathbb{F}_2^n} (-1)^{\oplus_{i\in S}(|x_i|+C_i'(x_1,\ldots,x_n) \pmod 2)} \tag{5.4.17}$$

$$= \frac{1}{2^{|S|\cdot n}} \sum_{x_1,\ldots,x_k\in\mathbb{F}_2^n} (-1)^{\left(\sum_{i\in S}|x_i|+\sum_{i\in S}C_i'(x_1,\ldots,x_n)\right) \pmod 2}, \tag{5.4.18}$$

where $\chi_S(z) = \oplus_{i\in S} z_i$.
This allows us to bind the previous value for all the subsets S based on the number of times the following equation is fulfilled for uniformly selected strings x_i,

$$\sum_{i\in S}|x_i| \equiv \sum_{i\in S} C_i'(x_1,\ldots,x_n) \mod 2. \tag{5.4.19}$$

We observe that determining the values for the outcome bits $i \in S$ such that the condition of Equation (5.4.19) is satisfied by an $\mathsf{AC}^0[q]$ circuit C' is as hard as computing the parity function across the concatenated input strings $||_{i\in S}x_i$. This assertion is supported by the fact that a circuit, which, if capable of determining valid outcome bits under the given condition, can be reduced to one that does compute the parity of the concatenated input strings $||_{i\in S}x_i$ with a polynomial-size $\mathsf{AC}^0[q]$ circuit. The latter follows from the fact that polynomial size $\mathsf{AC}^0[q]$ circuits can effectively compute the parity of a string of length $\log(n)$. Consequently, the probability that Equation (5.4.19) is satisfied is directly related to the probability that an $\mathsf{AC}^0[q]$ circuit computes the parity function across the concatenated input strings $||_{i\in S}x_i$, albeit with a one-sided error. This condition is met with a maximum probability of $1/2 + n^{-\Omega(1)}/2$, as inferred from Corollary 11. Therefore, we deduce that for any characteristic function χ_S, the value of $|\mathbb{E}_{x\in D}[(-1)^{\chi_S(x)}]|$ is limited by $\frac{1}{n^{\Omega(1)}}$, for all input strings of either even or odd Hamming weight.
Now, we can apply Vazirani's XOR Lemma (Lemma 56) to demonstrate that the distribution $\mathcal{X}$ is at most $1/n^{\Omega(1)} \cdot 2^{k/2}$ deviated in total variational distance from the uniform distribution, on those inputs. Then, we can show by using the Chernoff bound that the event X described as sampling a string with at least $1/2 + \frac{1-\cos\left(\frac{2\pi}{p}\right)}{2q}$ of 0's from this distribution, has a probability that decreases with,

$$\Pr\Big(X\Big) < e^{-2\cdot(4/\frac{1-\cos\left(\frac{2\pi}{p}\right)}{2q})^2(1/2)^2/(k^2)} < e^{-\Omega(k^2)}\ . \tag{5.4.20}$$

Thus, we obtain that the distribution $\mathcal{X}$ contains the same string with probability at most,

$$n^{-\Omega(1)} \cdot 2^{k/2} + e^{-\Omega(k^2)} = n^{-\Omega(1)+k/(2\log(n))}. \tag{5.4.21}$$

Finally, we conclude that the previous value is bounded by $n^{\Omega(1)}$, given that $k = \log(n)$. □

5.4.3 Separations

In this section, we combine the results from Section 5.4.1 and Section 5.4.2 to achieve unconditional separation between the classical and quantum circuit classes.

Theorem 5.4.3. *For each fixed prime q, there exists a relation problem that cannot be solved by polynomial-size circuits $\mathsf{AC}^0[q]$ with success probability at least $n^{-\Omega(1)}$, whereas there is a QNC^0_q/qpoly circuit and a QAC^0_p circuit that can solve it with probability of least $1 - o(1)$.*

Proof. For the case where $q = 2$, the result is derived from Theorem 6 in [27]. For all $q \neq 2$, we apply the lower bounds and the respective instances of the parallel modular relation problems $k\text{-}\mathcal{R}^{q\cdot n}_{q,2}$ discussed in Lemma 62. These instances establish that $\mathsf{AC}^0[q]$ circuits solve this set of problems with a probability of at most $n^{-\Omega(1)}$.
In contrast, the QNC^0_q/qpoly circuit, described in Theorem 5.4.1, can solve each instance of the parallel repetition game with a one-sided error of $\frac{1-\cos\left(\frac{2\pi}{p}\right)}{q}$. Similarly, QAC^0_p circuits achieve the same success probability, as detailed in Theorem 5.4.2. By combining these individual solutions, we deduce that the resulting global strategies effectively solve the $k\text{-}\mathcal{R}^{q\cdot n}_{q,2}$ problems, as outlined in Lemma 62, with a probability of at least $1 - o(1)$. This follows by considering their respective individual success probabilities and applying the Chernoff bounds. □

Remark. *We implicitly used the fact that the* Mod_2 *operation is not contained in any other* $\mathsf{AC}^0[q]$ *class with* $q \neq 2$*, which simplifies the proof since it allowed us to uniquely use Vazirani's XOR Lemma. However, we notice that the proof could follow with any other prime values and modular functions.*

One drawback of the previous results is that we need to pick a different qudit dimension for each modular relation problem. However, if we consider infinite-size gatesets, we can show that all the previous quantum circuits can be implemented with qubits and prove the separation for $\mathsf{AC}^0[p]$ for all primes p.

Theorem 5.4.4. *For all* p *prime,* $\text{i-}\mathsf{QAC}^0_2 \not\subseteq \mathsf{AC}^0[p]$.

Proof. From [107] and [108], we have that QNC^0_p circuits can be implemented using single and two-qubit operations contained in $\text{i-}\mathsf{QNC}^0_2$. Furthermore, all quantum operations in QAC^0_p used for the creation of the qudit-GHZ state in Lemmas 59 and 60 are controlled on a constant number of qudits and therefore these operations can be implemented $\text{i-}\mathsf{QNC}^0_2$.
Finally, we need to show that corrections needed to transform the poor man's cat states into qudit-GHZ states can be executed in $\text{i-}\mathsf{QAC}^0_2$. As described in the proof of Lemma 60, these corrections can be realized using a classical AC^0 circuit, which is encompassed by $\text{i-}\mathsf{QAC}^0_2$. Consequently, all quantum circuits from Theorems 5.4.2 and 5.4.3, for any prime p, can be constructed using $\text{i-}\mathsf{QAC}^0_2$ circuits. □

5.4.4 Classical upper bounds for modular relation problems

In this subsection, we consider the classical upper bound on the modular relation problems, which indicates the limitations on the quantum/classical separations using our techniques.

Lemma 63. *For any fixed primes* q,p *and fixed* $k_1, k_2, m \in \mathbb{N}$, $\mathcal{R}^m_{q^{k_1},p^{k_2}}$ *can be solved by an* $\mathsf{NC}^0[p]$ *circuit.*

Proof. We start by showing that there exists an $\mathsf{NC}^0[p]$ circuit that computes $\mathsf{Mod}_{p^{k_1}}$ for any $k_1 \in \mathbb{N}$. We use the $\mathsf{AC}^0[p]$ circuits described in [337] and argue that they can be implemented in $\mathsf{NC}^0[p]$.
We prove this statement by induction on the exponent k_1, showing that any Mod_{p^j} gate can be computed if we have access to $\mathsf{Mod}_{p^{j-1}}$ gates. The circuit construction from [337] is as follows:

1. For each $i \in [n]$, compute $y_i = \mathsf{Mod}_{p_{j-1}}(x)$.
2. For all $i \in 2, \ldots, n$, compute $z_i = \mathsf{AND}(y_{i-1}, \neg y_i)$.
3. Compute $b_j = \mathsf{AND}(\mathsf{Mod}_{p^{j-1}}(\sum_{i\in 2,\ldots,n} z_i), \mathsf{Mod}_p(|x|)) = \mathsf{Mod}_{p^j}(x)$.

Given that the final value $b_j = \mathsf{Mod}_{p^j}(x)$ [337], in order to show that Mod_{p^j} is in $\mathsf{NC}^0[p]$, we only need to show that each of these three steps can be implemented in $\mathsf{NC}^0[p]$. Let's evaluate these steps:

- The first step is inherently within $\mathsf{NC}^0[p]$ due to inductive reasoning; given the prior case produced an $\mathsf{Mod}_{p^{j-1}}$ gate, only the next steps need validation.
- The second step can be achieved with an NC^0 circuit since it solely relies on bounded fan-in AND gates.
- The third step follows since both modular gates are within $\mathsf{NC}^0[p]$ from our induction, combined with the AND gate having bounded fan-in.

Thus, any $\mathsf{Mod}_{p^{k_1}}$ gate can be implemented within $\mathsf{NC}^0[p]$ provided that k_1 is a constant, which determines the depth of the previously proposed circuit.

Finally, to prove our statement, i.e., to output a valid string for the modular relation problem $\mathcal{R}^m_{q^{k_1},p^{k_2}}$, one needs to use the $\mathsf{NC}^0[p]$ circuit to compute the value of $\mathsf{Mod}_{p^{k_1}}(x)$ and then append it to a string like 0^{m-1}. We have that $|b_{k_1}0^{m-1}| \pmod{q^{k_2}} = 0$ iff $|x| \pmod{p^{k_1}} = 0$. □

Given the containment of the $\mathsf{NC}^0[p]$ circuit classes within the $\mathsf{AC}^0[p]$ classes, we have that all modular relation problems are solvable by TC^0 circuits.

Corollary 12. *For any primes q,p and $k_1, k_2, m \in \mathbb{N}$, $\mathcal{R}^m_{q^{k_1},p^{k_2}}$ can be solved in* TC^0.

5.5 Shallow depth quantum circuits with infinite-size gate sets

This section considers constant-depth quantum circuits utilizing qubits and qudits, employing infinite gate sets and modular gates with unbounded fan-in while exploring their equivalences.

5.5.1 Constant-depth qudit circuits with quantum modular operators

In this subsection, we extend [77] and [338], and prove that the hierarchy of constant-depth quantum circuits with an infinite gateset over qudits collapses when we have access to a multi-qubit quantum/classical modular gate with unbounded fan-in gates.
We start by focusing on the quantum OR function, defined over Hilbert spaces of prime dimension p as follows,

$$\mathsf{qOR}\,|x_1, x_2, \ldots, x_n\rangle := |x_1 + \mathsf{OR}(x_2, x_3, \ldots, x_n) \pmod{p}, x_2, \ldots, x_n\rangle\,. \tag{5.5.1}$$

with the classical OR operating over strings in $\mathbb{F}_p$ defined as follows[6],

$$\mathsf{OR}(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i > 0, \\ 0 & \text{otherwise.} \end{cases}. \tag{5.5.2}$$

We first claim that a qudit analog of Høyer's and Špaleks OR reduction [53], which allows us to reduce the problem of computing qOR on an arbitrary state to the evaluation of qOR on another state with exponentially fewer qudits.

Lemma 64. *Let p be a prime. There exists a* i-$\mathsf{QNC}^0_p[p]$ *circuit C, such that for every n qudit state $|\psi\rangle = \sum_{x\in\mathbb{F}_p^n} \alpha_x |x\rangle$, we have*

[6] The intuition behind the AND and OR functions over strings in $\mathbb{F}_p$ is that, similar to the binary case, the AND function discriminates against only one string, while the OR function includes all strings except one.

$$C\left|\psi\right\rangle\left|0\right\rangle^{\otimes t}=\sum_{x\in\mathbb{F}_p^n,\ y\in\mathbb{F}_p^{\log_p(n)}}\alpha_x\beta_{y,x}\left|y\right\rangle\left|x\right\rangle\left|0\right\rangle^{\otimes t'}=\left|\psi^*\right\rangle\left|0\right\rangle^{\otimes t'}, \tag{5.5.3}$$

where and $\mathsf{qOR}\left|0\right\rangle\left|\psi\right\rangle=\mathsf{qOR}\left|0\right\rangle\left|\psi^*\right\rangle$ *with respect to the first qudit, while the second* qOR *does operate uniquely over the first* $\log(n)$ *qudits of* $\left|\psi^*\right\rangle$.

Proof. We describe now our circuit C. In the first step of this proposed circuit, we have a layer of n fanout gates, which are defined as follows

$$\mathsf{fanout}_p\left|x_1,x_2,\ldots,x_n\right\rangle:=\left|x_1,(x_1+x_2)\mod p,\ldots,(x_1+x_n)\mod p\right\rangle. \tag{5.5.4}$$

The k-th fanout_p gate is applied to the k-th qudit of $\left|\psi\right\rangle$ and a fresh $\log_p(n)$-qudit state initialized on $\left|0\right\rangle^{\otimes\log(n)}$[7].
In parallel to these fanout_p gates, we create $\log_p(n)$ copies of n-qudit GHZ states. These can be created by an $\mathsf{QNC}_p^0[p]$ circuit using the method proposed in [339]. We have the following state

$$\left|\psi^{(1)}\right\rangle=C_{step\ 1}\left|\psi\right\rangle\left|0\right\rangle^{\otimes t}=\sum_{x\in\mathbb{F}_p^n}\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\left|\mathsf{GHZ}_{p,n}^0\right\rangle^{\otimes\log_p(n)}\left|0\right\rangle^{\otimes t'}. \tag{5.5.5}$$

For each $1\leq k\leq\log_p(n)$, we apply n controlled rotations, where the j-th rotation of the k-th block will have as control the j-th qudit of the k-th copy of $\left|x\right\rangle$ and its target is the j-th qudit of the k-th GHZ state. The amplitude for all rotations in the k-th block is set to a fixed value of $\frac{2\pi}{p^k}$. Importantly, these operations can be implemented in constant depth due to the parallel copies of $\left|x\right\rangle$.
For a fixed x, this process modifies the k-th GHZ state to

$$\left(\bigotimes_{i=1}^{n}\mathsf{R}_Z^p\left(\frac{2\pi x_i}{p^k}\right)\right)\left|\mathsf{GHZ}_{p,n}^0\right\rangle=\sum_{x\in\mathbb{F}_p^n}\alpha_x\left(\frac{1}{\sqrt{p}}\sum_{j=0}^{p-1}e^{i\cdot\left(j\cdot\frac{2\pi}{p}\cdot\frac{|x|}{p^k}\right)}\left|j\right\rangle^{\otimes n}\right) \tag{5.5.6}$$

$$=\sum_{x\in\mathbb{F}_p^n}\alpha_x\sum_{j=0}^{p-1}c_{j,k,x}\left|\mathsf{GHZ}_{\mathsf{p,n}}^{\mathsf{j}}\right\rangle, \tag{5.5.7}$$

with $(c_{j,k,x})^2=e^{i\left(j\left(\frac{2\pi}{p}(1-\frac{|x|}{p^k})\right)\right)}$.
Note that for $|x|=0$, all these GHZ states remain unchanged and are equal to $\left|\mathsf{GHZ}_{p,n}^0\right\rangle$. However, in contrast, we will now prove that for all basis states $\left|x\right\rangle$ with $|x|\geq 1$, one of the $\log_p(n)$ GHZ states we started with ends up in the form $\left|\mathsf{GHZ}_{p,n}^m\right\rangle$, where $m\neq 0$. In other words, for each nonzero x, there exists one value of k for which $c_{j,k,x}$ equals 1 for a single value of j and 0 for all others. Meanwhile, for the same value of x and all other values of k, $c_{j,k,x}$ takes non-unitary values for all values of j. To establish this, we utilize the fact that any integer number $|x|$ can always be decomposed as

$$|x|=p^{a_x}(p\cdot b_x+m_x), \tag{5.5.8}$$

with $a_x,b_x\in\mathbb{N}$ and $m_x\in[1,2,\ldots,p-1]$. Hence, if $|x|>0$, then we have that

$$\left(\bigotimes_{i=1}^{n}\mathsf{R}_Z^p\left(\frac{2\pi x_i}{p^{(a_x+1)}}\right)\right)\left|\mathsf{GHZ}_{p,n}^0\right\rangle=\left|\mathsf{GHZ}_{p,n}^{m_x}\right\rangle, \tag{5.5.9}$$

[7] Note that this fanout gate can be equally created from the qMOD_p gate, $\mathsf{fanout}_p=(\mathsf{F}_p)^{\otimes n}\mathsf{qMOD}_p(\mathsf{F}_p)^{\otimes n}$.

with $m_x \neq 0$ by definition. Therefore, $c_{j=m_x,k=a_x,x} = 1$ and $c_{j\neq m_x,k=a_x,x} = 0$, while for all the other GHZ states, these coefficients $c_{j,k\neq a_x,x}$ are non-unital as $\frac{2\pi}{p}(1 - \frac{|x|}{p^k})$ is not a multiple of 2π. Thus, the state resulting from this set of controlled rotations is the following form,

$$\left|\psi^{(2)}\right\rangle = C_{step\ 2}\left(\sum_{x\in\mathbb{F}_p^n} \alpha_x \left|x\right\rangle^{\otimes \log_p(n)} \left|\mathsf{GHZ}^0_{p,n}\right\rangle^{\otimes \log_p(n)}\right) \otimes \left|0\right\rangle^{\otimes t'} \tag{5.5.10}$$

$$= \left(\alpha_{0^n} \left|0\right\rangle^{\otimes n\log_p(n)} \left|\mathsf{GHZ}^0_{p,n}\right\rangle_0 \left|\mathsf{GHZ}^0_{p,n}\right\rangle_1 \ldots \left|\mathsf{GHZ}^0_{p,n}\right\rangle_{\log_p(n)} + \sum_{x\in\mathbb{F}_p^n\setminus 0^n} \alpha_x \left|x\right\rangle^{\otimes \log_p(n)}\right. \tag{5.5.11}$$

$$\left. \ldots \left(\sum_{j=0}^{p-1} c_{j,a_x,x} \left|\mathsf{GHZ}^j_{p,n}\right\rangle\right) \left|\mathsf{GHZ}^{m_x}_{p,n}\right\rangle_{(a_x+1)} \left(\sum_{j=0}^{p-1} c_{j,(a_x+2),x} \left|\mathsf{GHZ}^j_{p,n}\right\rangle\right) \ldots \right) \otimes \left|0\right\rangle^{\otimes t'}. \tag{5.5.12}$$

We then apply Fourier gates $\mathsf{F}_p^{\otimes n}$ to the qudits initially entangled in GHZ states. Employing Lemma 58, we deduce that when $|x| = 0$, the resultant states form a superposition of strings with a Hamming weight congruent to zero modulo p. Conversely, for $|x| \geq 0$, at least one state transitions into a superposition of strings congruent modulo p to the inverse of m_x. It is important to note that since $m_x \mod p$ is invariably non-zero, its additive inverse $-m_x \mod p$ is likewise non-zero. Subsequently, we apply a qMOD_p gate with each of these states as control and as a target a new single qudit in a freshly initialized computational basis state $|0\rangle$ in our third register. The existence of the state described earlier ensures the qOR's intended effect, as at least one qudit in the third register will deterministically shift to a basis different from $|0\rangle$. For $|x| \geq 0$, since the GHZ states prior to the Fourier gates are superpositions in the Qudit-GHZ orthogonal X-basis, these strings after the Fourier gates are superpositions that align modulo p with potentially any value in $\mathbb{F}_p$. Consequently, in the third register, all other qudits assume superpositions dependent on the input, yet they do not influence the qOR operation.
Afterward, we apply a layer of the inverses of the Fourier gates $\mathsf{F}_p^{\otimes n}$ to exactly the same qudits in the second register as before, and consequently, they revert to the same states as described in Equations 5.5.11 and 5.5.12, being for all $k \in \{1, \ldots, \log(n)\}$ of the form

$$\sum_{x\in\mathbb{F}_p^n} \alpha_x \left(\frac{1}{\sqrt{p}} \sum_{j=0}^{p-1} e^{i\cdot\left(j\cdot\frac{2\pi}{p}\cdot\frac{|x|}{p^k}\right)} \left|j\right\rangle^{\otimes n}\right) = \sum_{x\in\mathbb{F}_p^n} \alpha_x \sum_{j=0}^{p-1} c_{j,k,x} \left|\mathsf{GHZ}^j_{p,n}\right\rangle. \tag{5.5.13}$$

This is possible because previously, we simply applied a qMOD_p with these states as control and did not change the target qudits. Thus, these states return to the same bases and states after the use of the inverse Fourier gates. Therefore, after these operations, we obtain the subsequent state, for which we do not represent the qudits of the third register in its position but next to each GHZ state that generated each one of them for a simpler representation.

$$C_{step\ 3}\left|\psi^{(2)}\right\rangle \left|0\right\rangle^{\otimes t'} = \left(\alpha_{0^n} \left|0\right\rangle^{\otimes (n\cdot\log_p(n))} \left|\mathsf{GHZ}^0_{p,n}\right\rangle_0 \left|\mathsf{GHZ}^0_{p,n}\right\rangle_1 \ldots \left|\mathsf{GHZ}^0_{p,n}\right\rangle_{\log_p(n)} \left|0\right\rangle^{\otimes \log_p(n)}\right. \tag{5.5.14}$$

$$+\sum_{x\in\mathbb{F}_p^n\setminus 0^n}\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left(\sum_{j=0}^{p-1}c_{j,a_x,x}\left|\mathsf{GHZ}_{p,n}^{j}\right\rangle\left|-j\right\rangle\right)\left|\mathsf{GHZ}_{p,n}^{m_x}\right\rangle_{(a_x+1)}\left|-m_x\right\rangle \tag{5.5.15}$$

$$\left(\sum_{j=0}^{p-1}c_{j,(a_x+2),x}\left|\mathsf{GHZ}_{p,n}^{j}\right\rangle\left|-j\right\rangle\right)\ldots\Bigg)\otimes\left|0\right\rangle^{\otimes t''}. \tag{5.5.16}$$

At this stage, all the qudits in the three registers are entangled. However, we can disentangle the first and third registers from the second, which contains the GHZ states. This disentanglement can be achieved using the third register that retains the phase information of the equivalent of the GHZ states, creating the entanglement between these two registers. For that, we use generalized controlled rotations, with the control based on the qudits of the third register, applied to one qudit of the respective GHZ state of amplitude $\frac{2\pi}{p^k}$. For simplicity, we will consider first the effect on a single basis of $\left|x\right\rangle$,

$$\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left(\sum_{j=0}^{p-1}c_{j,a_x,x}\left(\mathsf{CGR}\left(\frac{2\pi}{p^k}\right)\left|\mathsf{GHZ}_{p,n}^{j}\right\rangle\left|-j\right\rangle\right)\right)\ldots \tag{5.5.17}$$

$$=\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left(\frac{1}{\sqrt{p}}\sum_{j=0}^{p-1}e^{i\cdot\left(j\cdot\frac{2\pi}{p}\cdot\frac{|x|}{p^k}\right)}\left(\mathsf{CGR}\left(\frac{2\pi}{p^k}\right)\left|j\right\rangle^{\otimes n}\left|-j\right\rangle\right)\right)\ldots \tag{5.5.18}$$

$$=\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left(\frac{1}{\sqrt{p}}\sum_{j=0}^{p-1}e^{i\cdot\left(j\cdot\frac{2\pi}{p}\cdot\frac{|x|}{p^k}\right)}e^{i\cdot\left(-j\cdot\frac{2\pi}{p}\cdot\frac{|x|}{p^k}\right)}\left|j\right\rangle^{\otimes n}\left|-j\right\rangle\right)\ldots \tag{5.5.19}$$

$$=\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left(\sum_{j=0}^{p-1}\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle\left|-j\right\rangle\right)\ldots. \tag{5.5.20}$$

Now as the previous transformation is independent of the respective basis $\left|x\right\rangle$, we obtain the following state,

$$C_{step\ 4}\left|\psi^{(3)}\right\rangle\left|0\right\rangle^{\otimes t''}=\Bigg(\alpha_{0^n}\left|0\right\rangle^{\otimes(n\cdot\log_p(n))}\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle_0\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle_1\ldots\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle_{\log_p(n)}\left|0\right\rangle^{\otimes\log_p(n)} \tag{5.5.21}$$

$$+\sum_{x\in\mathbb{F}_p^n\setminus 0^n}\alpha_x\left|x\right\rangle^{\otimes\log_p(n)}\ldots\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle\left(\sum_{j=0}^{p-1}c_{j,a_x,x}\left|-j\right\rangle\right)\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle_{(a_x+1)}\left|-m_x\right\rangle \tag{5.5.22}$$

$$\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle\left(\sum_{j=0}^{p-1}c_{j,(a_x+2),x}\left|-j\right\rangle\right)\ldots\Bigg)\otimes\left|0\right\rangle^{\otimes t''}. \tag{5.5.23}$$

Consequently, we end up with $\log_p(n)$ GHZ states of the form $\left|\mathsf{GHZ}_{p,n}^{0}\right\rangle$, which are no longer entangled with the other registers, as intended. These states and the repetitions of the initial state can be easily removed using the fanout gates and their inverses[8] obtaining the following state,

[8] The inverse of the fanout_p gate is simply $(\mathsf{fanout}_p)^{p-1}$.

$$\left|\psi^{(5)}\right\rangle = C_{step\ 5}\left|\psi^{(4)}\right\rangle = \left(\alpha_{0^n}\left|0\right\rangle^{\otimes(n+\log_p(n))} + \sum_{x\in\mathbb{F}_p^n\setminus 0^n}\alpha_x\left|x\right\rangle\ldots\left(\sum_{j=0}^{p-1}c_{j,a_x,x}\left|-j\right\rangle\right)\left|-m_x\right\rangle\right.$$

(5.5.24)

$$\left.\left(\sum_{j=0}^{p-1}c_{j,(a_x+2),x}\left|-j\right\rangle\right)\ldots\right)\otimes\left|0\right\rangle^{\otimes t'''} = \sum_{x\in\mathbb{F}_p^n,\ y\in\mathbb{F}_p^{\log_p(n)}}\alpha_x\beta_{y,x}\left|x\right\rangle\left|y\right\rangle\otimes\left|0\right\rangle^{\otimes t'''}.$$

(5.5.25)

Finally, if we take $|\psi^*\rangle$ as $|\psi^5\rangle$ with the first and second register swapped, we obtain that the former state satisfies the property $\mathsf{qOR}\left|0\right\rangle\left|\psi\right\rangle = \mathsf{qOR}\left|0\right\rangle\left|\psi^*\right\rangle$ on the first qudit, with the second qOR applied only to the first $\log(n)$ qudits. This is justified by that fact that the basis $\left|0\right\rangle^{\otimes\log_p(n)}$ in $|\psi^*\rangle$ has the same amplitude as the basis $\left|0\right\rangle^{\otimes n}$ in $|\psi\rangle$. The same principle forcefully applies to set all the bases orthogonal to the previously described, as demonstrated in Equations 5.5.24 and 5.5.25. Therefore, the effect of the qOR function on the first qudit state is equal for both cases considered completing the proof. □

The second step shows an exponential-size i-QNC_p^0 circuit for qOR. For that, we use the following theorem on the representation of functions as a multi-linear polynomial over $\mathbb{F}_p$, expanding previous uses of these objects over $\mathbb{F}_2$ and $\mathbb{F}_3$ [338, 95, 340].

Theorem 5.5.1. *[341] Every function $f:\mathbb{F}_p^n\mapsto\mathbb{R}$ can be expressed as a polynomial*

$$f(x) = \sum_{k_1,k_2,\ldots,k_n\in\mathbb{F}_p}\widehat{f}(\chi_{k_1,k_2,\ldots,k_n})\cdot\chi_{k_1,k_2,\ldots,k_n}(x), \tag{5.5.26}$$

with $\widehat{f}(\chi_{k_1,k_2,\ldots,k_n})$ being a real coefficient, and $\chi_{k_1,k_2,\ldots,k_n}(x)$ multi-linear function in $\mathbb{F}_p$.

We now extend the techniques of [77] to show that our exponential-size circuit for qOR.

Lemma 65. *For any prime p,* qOR *can be implemented on an n-qudit state using $\mathcal{O}(n\cdot p^n)$ operations in* i-$\mathsf{QNC}_p^0[p]$.

Proof. Let us describe the i-QNC_p^0 circuit that computes qOR. The first layer of the circuit consists of $n-1$ fanout gates, where the i-th gate is controlled by the $(i+1)$-th qudit of the initial state to a new register initialized to $\left|0\right\rangle^{\otimes p^{n-1}}$. This maps the initial state $\left|0\right\rangle\left|\psi\right\rangle\left|0\right\rangle^{\otimes t}$ to

$$\left|0\right\rangle\sum_{x\in\mathbb{F}_p^{n-1}}\alpha_x\left|x\right\rangle\left(\left|x_2,\ldots,x_n\right\rangle\right)^{\otimes p^{n-1}}\left|0\right\rangle^{\otimes t'}. \tag{5.5.27}$$

We then apply p^{n-1} parallel qMOD_p gates, where the i-th gate is applied on the i-th block of the previous state, and the qMOD_p computes a different $\chi_{k_1,k_2,\ldots,k_n}$ leading to

$$\left|0\right\rangle\sum_{x\in\mathbb{F}_p^n}\alpha_x\left|x\right\rangle\left(\bigotimes_{k_1,\ldots,k_{n-1}\in\mathbb{F}_p}\left|x_2,\ldots,x_n\right\rangle\left|\chi_{k_1,k_2,\ldots,k_{n-1}}(x_2,\ldots,x_n)\right\rangle\right)\left|0\right\rangle^{\otimes t'}. \tag{5.5.28}$$

We then create a $\left|\mathsf{GHZ}_{p,p^{n-1}}^0\right\rangle$ state and we apply generalized controlled rotations as follows. The i-th rotation has as control the register where $\chi_{k_1,k_2,\ldots,k_n}(x_2,\ldots,x_n)$ is computed and the target is the i-th

qudit of the GHZ state. The angle of the rotation is $\widehat{f}_{\overline{\mathsf{OR}}}(\chi_{k_1,k_2,\ldots,k_n})$. These angles are defined by the $\overline{\mathsf{OR}}$ function which is defined as follows,

$$\overline{\mathsf{OR}}(x) = \begin{cases} p-1 & \text{if } \sum_{i=1}^{n} x_i > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.5.29}$$

and that we decompose, according to Theorem 5.5.1 as

$$\overline{\mathsf{OR}}(x) = \sum_{k_1,k_2,\ldots,k_n \in \mathbb{F}_p} \widehat{f}_{\overline{\mathsf{OR}}}(\chi_{k_1,k_2,\ldots,k_n}) \cdot \chi_{k_1,k_2,\ldots,k_n}(x) \tag{5.5.30}$$

to obtain the following state on this last register,

$$\begin{aligned} \left|\overline{\mathsf{OR}}\right\rangle &= \sum_{x\in\mathbb{F}_p^n} \alpha_x \left(\frac{1}{\sqrt{p}} \sum_{j=0}^{p-1} e^{j\cdot\frac{\pi}{p}\left(\sum_{k_1,k_2,\ldots,k_{n-1}\in\mathbb{F}_p} \widehat{f}_{\overline{\mathsf{OR}}}(\chi_{k_1,k_2,\ldots,k_{n-1}})\cdot\chi_{k_1,k_2,\ldots,k_{n-1}}(x_2\ldots x_n)\right)} |j\rangle^{\otimes p^{n-1}} \right) \\ &= \sum_{x\in\mathbb{F}_p^n} \alpha_x \left(\frac{1}{\sqrt{p}} \sum_{j=0}^{p-1} e^{\frac{j\cdot\pi\cdot\overline{\mathsf{OR}}(x_2\ldots x_n)}{p}} |j\rangle^{\otimes p^{n-1}} \right) = \sum_{x\in\mathbb{F}_p^n} \alpha_x \left| \mathsf{GHZ}_{p,p^{n-1}}^{\overline{\mathsf{OR}}(x_2,\ldots,x_n)} \right\rangle . \end{aligned}$$

In the last step, we apply the Fourier gate to all the qudits in the newly created state in the third register and a single use of the quantum modular gate qMOD_p from all the qudits to the first qudit in the state $|0\rangle$. Using Lemmas 57 and 58, we obtain the state

$$\begin{aligned} &\sum_{x\in\mathbb{F}_p^n} \alpha_x \left|\overline{\mathsf{OR}}(x_2\ldots x_n)^{-1}\right\rangle |x_1,x_2,\ldots,x_n\rangle \\ &\left(\bigotimes_{k_1,\ldots,k_{n-1}} |x_2,\ldots,x_n\rangle \left|\chi_{k_1,k_2,\ldots,k_{n-1}}(x_2,\ldots,x_n)\right\rangle \right) \left(\frac{1}{\sqrt{p^{n-2}}} \sum_{\substack{z\in\mathbb{F}_p^{n-1} \\ |z| \mod p = -\overline{\mathsf{OR}}(x_2,\ldots,x_n)}} |z\rangle \right). \end{aligned}$$

Next, we apply a SUM gate to the first qudit, with the first qudit from the initial state $|\psi\rangle$. This ensures we obtain the expected output state from applying qOR function on the input state $|\psi\rangle$ given that $-\overline{\mathsf{OR}}(x_2,\ldots,x_n) = \mathsf{OR}(x_2,\ldots,x_n)$. Finally, we execute the inverse operations on the auxiliary qudits to decouple these from the input state. We can do this because all operations acting on the input state and auxiliary qudits are unitary. In addition, the inverse operations of fanout_p, qMod_p, and of the controlled rotations—are all within $\mathsf{QNC}_p^0[p]$. See Appendix 5.3.2 for details on the first two operations. □

We now prove our first collapse of the different classes on constant-depth quantum circuits.

Theorem 5.5.2. i-$\mathsf{QNC}_p[p] = $ i-$\mathsf{QAC}_p[p]$*, for any p prime.*

Proof. To prove this collapse, we need only to show that qOR and the qAND operation can be implemented by an i-$\mathsf{QNC}_p[p]$ circuit. The latter gate is already the Toffoli X gate, which we can use to obtain any other multi-qudit Toffoli gate.

To implement qOR, we integrate circuits C_1 and C_2 from Lemmas 64 and 65, respectively. Circuit C_1 effectively reduces the number of qudits in the state $|\psi\rangle$, generating a new state $|\psi^*\rangle$ that preserves the qOR output alongside a fresh $|0\rangle$, and C_2 performs the actual qOR computation. By applying C_1 to all qudits of $|\psi\rangle$ except the first, and subsequently feeding $|0\rangle|\psi^*\rangle$ to C_2. The global output is generated by

applying the SUM gate to the first qudit $\sum_{x\in[p^{n-1}]}\alpha_{x_2,\ldots,x_n}|\mathsf{OR}(x_2,\ldots,x_n)\rangle$ resulting from C_2 and the first qudit of the initial $|\psi\rangle$ state.
To conclude the proof, we return all auxiliary qudits to their initial state, $|0\rangle^{\otimes t}$. Initially, we observe that $\log(n)$ auxiliary qudits arise from the OR reduction. We then apply the exponentially large qOR operator, described in Lemma 65, to these $\log_p(n)$ qudits. Importantly, this step does not introduce additional auxiliary qudits. Thus, our goal now is to reverse the state of these $\log(n)$ auxiliary qudits to $|0\rangle^{\otimes \log_p(n)}$, without impacting the result obtained from the qOR operation. We notice that undoing the operation on the $\log(n)$ qubits is possible since the operations of Lemma 64 can be reverted even if the first and second registers are entangled with the qubit where the qOR was applied. This implies that the $\log(n)$ qudits were unnecessary to achieve the intended final state. Hence, we can reversibly return these qudits to their initial state without losing any information.
The reversal process is straightforward. We follow the steps used to generate these $\log_p(n)$ qudits but omit the application of the qMOD operator, as delineated in Lemma 64. Instead, we employ the qMOD^{p-1} operator. This ensures that all qudits are restored to the $|0\rangle$ state, completing the process unitarily. □

For the next collapse, we need to demonstrate that the quantum threshold gate is in $\text{i-QNC}^0_p[p]$. For that, we first demonstrate that the quantum version of the Exact_k function is contained in the same circuit class.

Definition 67. Exact_k. *The* Exact_k *function is defined as follows,*

$$\mathsf{Exact}_k(x)=\begin{cases}1 & \text{if } \sum_{i=1}^n x_i=k\\ 0 & \text{otherwise}\end{cases}, \tag{5.5.31}$$

and its quantum analog as the following operation over Hilbert spaces of prime dimension p

$$\mathsf{qExact}_k|x_1,x_2,\ldots,x_n\rangle:=|x_1+\mathsf{Exact}_k(x_2,x_3,\ldots,x_n)\pmod p,x_2,\ldots,x_n\rangle. \tag{5.5.32}$$

Lemma 66. *For any prime* p, qExact *can be implemented in* $\text{i-QNC}^0_p[p]$.

Proof. We approach this problem once again using a constructive method, retaining most elements from the circuit used for the qOR function. The qOR function fundamentally discerns whether a computational basis state has a Hamming weight of zero. However, the initial step involving the qOR reduction will be modified. Previously, this reduction applied controlled rotations to GHZ states, effectively implementing a rotation with an amplitude proportional to the Hamming weight of the computational basis states, denoted as $\frac{2\pi|x|}{p^i}$. The new circuit introduces additional rotations to each i-th GHZ state with an amplitude of $-\frac{2\pi k}{p^i}$ respectively. This supplemental rotation emulates the behavior of the original circuit, equating the effects of a string with a Hamming weight of zero to those of a string with a Hamming weight of k. As a result, when these modified states are subsequently processed by the remaining qOR circuitry, the outcome will precisely compute the qExact_k function. □

Using Lemma 66, we push the collapse of Theorem 5.5.2 further to constant-depth quantum circuits with threshold gates.

Theorem 5.5.3. *For any prime* p, $\text{i-QNC}_p[p]=\text{i-QTC}_p$.

Proof. To prove this statement, it is sufficient to show that qTH_k can be implemented in $\text{i-QNC}_p[p]$. The circuit for qTH_k is composed of $n-k$ qExact_t operations (described in Lemma 66) for which t will range from k to n. In this setup, the fanout gate is employed to parallelize the application of the qExact_t operations. As a result, we generate states of the form $\sum_{x\in\mathbb{F}_p^{n-1}}\alpha_x|\mathsf{Exact}_k(x_2\ldots x_n)\rangle|x\rangle$ based on the input state $|\psi\rangle=\sum_{x\in\mathbb{F}_p^n}\alpha_x|0\rangle^{\otimes(n-k)}|x\rangle$.

Subsequently, the qOR gate will be applied to all these states, and a computational basis state $|0\rangle$, obtained with the following state,

$$\sum_{x\in\mathbb{F}_p^n} \alpha_x \left|\mathsf{OR}(\mathsf{Exact}_k(x_2\dots x_n), \mathsf{Exact}_{k+1}(x_2\dots x_n), \dots, \mathsf{Exact}_n(x_2\dots x_n))\right\rangle \tag{5.5.33}$$

$$\left|\mathsf{Exact}_k(x_2\dots x_n)\right\rangle_k \dots \left|\mathsf{Exact}_k(x_2\dots x_n)\right\rangle_n |x\rangle^{\otimes(n-k)} \tag{5.5.34}$$

$$= \sum_{x\in\mathbb{F}_p^n} \alpha_x \left|\mathsf{TH}_k(x_1\dots x_n)\right\rangle \dots |x\rangle^{\otimes(n-k)} \ . \tag{5.5.35}$$

It requires only applying the SUM gate between the first qudit and the first qudit of the state $|\psi\rangle$, and we obtain the state resulting from applying the qTH_k on the input state $|\psi\rangle$. The last step resumes applying all the inverse unitaries on the auxiliary qudits to detangle those from the resulting state. Lastly, we apply the inverse unitaries to all auxiliary qudits, disentangling them from the final state while preserving the desired output. □

5.5.2 Constant-depth qudit circuits with classical modular operators

In this section, we consider the collapse of constant-depth circuits with infinite-size quantum gate sets and classical modular gates.
We first show that quantum modular gates can be implemented with constant depth measurement patterns and classical modular gates.

Lemma 67. *The* qMOD_p *gate can be implemented in* $\mathsf{i\text{-}QNC}^0_p[p]_\mathsf{c}$.

Proof. Our first step in this proof is to show that a circuit in $\mathsf{i\text{-}QNC}^0_p[p]_\mathsf{c}$ can implement the fanout_p gate over an arbitrary quantum state, $|\psi\rangle = \sum_{x\in\{0,1,\dots,p-1\}^n} \alpha_x |x\rangle$.
For that, we first create an $\left|\mathsf{GHZ}^0_{p,n}\right\rangle$ state and apply a SUM where the control-qudit is the first qudit of $|\psi\rangle$ and the target is the first qudit of the GHZ state. This leads to the state,

$$\left|\psi^{(1)}\right\rangle = \sum_{x\in\mathbb{F}_p^n} \alpha_x (\mathsf{I}^{\otimes n} \otimes (\mathsf{X}_p)^{x_1} \otimes \mathsf{I}^{\otimes(n-1)}) |x\rangle \left|\mathsf{GHZ}^0_p\right\rangle \ . \tag{5.5.36}$$

Subsequently, we measure the first qudit of the second register with outcome $m_1 \in \{0, 1, \dots, p-1\}$. The post-measured state will then be

$$\left|\psi^{(2)}\right\rangle = \sum_{x\in\mathbb{F}_p^n} \alpha_x |x_1, x_2\dots x_n, m_1, (x_1+m_1) \mod p, \dots, (x_1+m_1) \mod p\rangle \ . \tag{5.5.37}$$

We then apply $(\mathsf{X}_p)^{-m_1}$ to each of the $n-1$ qudits of the second register, leading to the state

$$\left|\psi^{(3)}\right\rangle = \sum_{x\in\mathbb{F}_p^n} \alpha_x |x_1, x_2, \dots, x_n, m_1, x_1, \dots, x_1\rangle \ . \tag{5.5.38}$$

Using this state, we can apply a layer of $n-1$ parallel SUM gates where the control gate is one of the $n-1$ last qudits, and the target is one of the qudits from the 2nd to the n-th qudit in the first register. Resulting in the following state

$$\left|\psi^{(4)}\right\rangle = \sum_{x\in\mathbb{F}_p^n} \alpha_x |x_1, (x_2+x_1) \mod p, \dots, (x_n+x_1) \mod p, m_1, x_1, \dots, x_1\rangle \ , \tag{5.5.39}$$

which is equivalent to applying the fanout_p gate on state $\left|\psi^{(3)}\right\rangle$.
The last step is to remove the entanglement between the first n qudits and the last $n-1$ ones. For that, we apply the Fourier gate to the last $n-1$ qudits, and we have

$$\left|\psi^{(5)}\right\rangle = \frac{1}{\sqrt{p^n}} \sum_{m_2 \ldots m_n \in \mathbb{F}_p^{n-1}} \sum_{x \in \mathbb{F}_p^n} \alpha_x \cdot e^{i\frac{-2\pi x_1 (\sum_{i=2}^n m_i)}{p}} \tag{5.5.40}$$

$$\left|x_1, (x_2 + x_1) \bmod p, \ldots, (x_{n-1} + x_1) \bmod p, m_1, \ldots, m_n\right\rangle \ . \tag{5.5.41}$$

We can then measure the last $n-1$ qudits with outcome $m_2, ..., m_n \in \mathbb{F}_p$ to determine the value of $\phi = -\sum_{i=2}^n m_i \pmod p$ and apply a controlled generalized rotation CGR_Z^p with control on x_1 and parameterized by ϕ to remove the relative phase introduced by the Fourier gates[9]. Tracing out the measured states, we have the state

$$\left|\psi^{(6)}\right\rangle = \sum_{x \in \mathbb{F}_p^n} \alpha_x \left|x_1, (x_2 + x_1) \bmod p, \ldots, (x_n + x_1) \bmod p\right\rangle \ . \tag{5.5.42}$$

which corresponds to state $(\mathsf{fanout}_\mathsf{p} \left|\psi\right\rangle)$, as intended.
Finally, we only need to use the circuit that translates the fanout_p gate to the qMOD_p gate, using simply the standard translation presented by qubits [53] in higher dimensions with the Fourier gates. This and the fact that all the sub-processes described previously are contained in $\text{i-QNC}_p^0[p]_\mathsf{c}$ completes the proof. □

Using this lemma and the results in the previous subsection we can prove the following.

Lemma 68. *For any prime p,* $\text{i-QNC}_p[p]_\mathsf{c} = \text{i-QTC}_p$.

Proof. It follows from Lemma 67 that $\text{i-QNC}_p[p]_\mathsf{c} = \text{i-QNC}_p^0[p]$. The result follows since $\text{i-QTC}_p = \text{i-QNC}_p^0[p]$ from Theorem 5.5.3. □

Lemma 68 shows that the implementation of constant-depth MBQC over prime dimensional qudits is capable of the same type of circuit collapses as all the ones that have been demonstrated for qubits. Now we show how these circuit classes compare to each other.

Theorem 5.5.4. *For any p and q prime.* $\text{i-QTC}_p^0 \subseteq \text{i-QNC}_2^0[q]_\mathsf{c}$.

Proof. The proof proceeds in two steps. Firstly, we show that i-QNC_q^0 and i-QNC_p^0 are contained in i-QNC_2^0 for any primes p and q. To establish this, we need a mapping from the qudit basis to the qubit basis for both prime dimensions, p and q. For example, qutrits can be mapped to qubits using the following correspondence: $|0\rangle \mapsto |00\rangle$, $|1\rangle \mapsto |01\rangle$ and $|2\rangle \mapsto |11\rangle$, utilizing only a subspace of the Hilbert space of dimension 4. Consequently, any qudit can be translated into $2\lceil p/2 \rceil$ qubits. Additionally, any unitary operator over a Hilbert space of dimensions p and $2p$ can be mapped to a unitary operator over a Hilbert space of dimensions $2\lceil p/2 \rceil$ and $4\lceil p/2 \rceil$. This new operator applies the same transformation over the encoded qubit basis as the original did over the qudits, while the basis ignored by the encoding is operated by the identity locally.
This argument can be augmented with the demonstration that all such fixed sizes can be constructed in constant depth using single and two-qubit gates, as described in [107] and [108]. These methods prove that any unitary operator on k qubits can be synthesized using at most $\mathcal{O}(k^3 4^k)$ two-qubit gates. This further suggests that all operations in i-QNC_q^0 and i-QNC_p^0 can be executed by an i-QNC_2^0 circuit with a specific encoding for each prime dimension.

[9] The value of ϕ is classical and can be determined using the classical Mod_p gates with a binary outcome in the form of p bits, with only one of them being equal to 1, determining the value ϕ.

Given that all operations in i-QNC$_q^0$ can be achieved within i-QNC$_2^0$, it follows that all operations in i-QNC$_q^0[q]_c$ also belong to i-QNC$_2^0[q]_c$. This implies that i-QNC$_2^0[q]_c$ contains the Mod$_p$ gate. Coupled with the fact that i-QNC$_2^0$ can implement all the operations in i-QNC$_p^0$, we deduce that i-QNC$_2^0[q]_c$ can execute all operations in i-QNC$_p^0[p]_c$ for any given prime p. The final piece to consider is Lemma 68, which allows us to arrive at the desired result. □

Building on this result, we can also extend the quantum-classical separations illustrated by [76] and [77], which separated i-QNC$_2^0[2]$ from the classical AC$^0[p]$ classes.

Corollary 13. *For all p and q prime,* i-QNC$_2^0[q]_c \not\subseteq$ AC$^0[p]$.

Proof. The i-QNC$_2^0[q]_c$ class contains all the operations of i-QNC$_q^0$; therefore, it contains all the operations in i-QTC$_q^0$, which include all the Mod$_k$ operations, with k prime. This is combined with Razborov-Smolensky separations (Theorem 5.3.1), which states that all AC$^0[p]$ classes fail to compute any Mod$_k$ gate other than the Mod$_p$ gate with p and k being two distinct prime numbers finishes the proof. □

Chapter 6

Rick: "That is something I should know, or at least something I should have predicted, which I did."

Heistotron: "But what does it even matter if I'm random or planned? ... My heists will still rule. [I have defeated you]"

Rick: "Well, if the brain inside of you is actually Heistotron's, it matters, because it explodes when it hits six levels of contrivance. [You have a fixed amount of times you can fool me]"

Heistotron: "Clap, clap, clap, clap, clap, ... [self-destruction]"

— *Rick and Morty*, Season 4 Episode 3

Verifcation-inspired quantum circuit benchmarking

Overview

In this section, we introduce a new approach to quantum benchmarking inspired by quantum verification, motivating new paradigms in quantum benchmarking. Our proposed benchmark not only serves as a robust indicator of computational capability within BQP but also offers scalability, customizability, and universality. By providing formal statements regarding the quality of quantum devices while assuming device consistency, we eliminate the reliance on heuristics. We establish a deep connection between quantum verification and quantum benchmarking. For practical applications, we present a concrete benchmarking protocol derived from a quantum verification protocol and prove that it aligns with our redefined standards for quantum benchmarking. In particular, this protocol enables the certification of quantum devices for the correct execution of various classes of quantum circuits, and consequently a robust device characterization. The paper providing the content for this section can be found in [144].

6.2 Introduction

Advances in quantum computation, mainly in hardware development, have taken significant steps toward the goal of building fault-tolerant quantum computers. With this constant improvement, it is of great interest to assess and compare the performance of various quantum devices. This benchmarking process plays a vital role in gauging the quality of ever-improving quantum hardware while simultaneously aiding end-users in determining the suitability of a specific quantum device for their intended applications. Currently, no universally accepted standard benchmarking technique exists. Moreover, so far, most considered quantum benchmarks have been largely heuristic and do not generate strong or reliable predictions about the computational performance of quantum devices. Recognizing the need for a more formal underpinning of predictive statements, recent efforts have emerged to strengthen the statements derived from heuristic quantum benchmarks, as discussed in [89]. From this standpoint, there is a clear desirability to transition benchmarking from its current heuristic nature to a more formal and predictable framework. Our aim is to make a meaningful contribution to this ongoing quest by integrating benchmarking with a previously distinct approach: quantum verification.

As already pointed out in [140], quantum benchmarking is naturally related to the field of quantum verification as both involve the certification of quantum computations. As our main contribution, we establish a strong link between these two fields by presenting a generic way of compiling blind verification schemes to benchmarking protocols. This compiler relies on a reduction from worst-case to average-case computations that blind verification schemes with strong security inherently exhibit. We further prove that quantum benchmarking schemes obtained in this verification-inspired way achieve the new stronger paradigms. We then apply this compiler to the robust quantum verification scheme of [91] to obtain a concrete benchmarking protocol that inherits its favorable properties of scalability, efficiency and noise robustness. The employed verification techniques do not entail a quantum hardware overhead beyond the size of the target computation. The new scheme does not rely on any heuristics or on any assumptions regarding the structure of noise and is inherently platform-agnostic. It offers statements that can be seen as certificates of computational power rather than a heuristic performance indicator.

6.2.1 Related work

Within quantum benchmarking, several methodologies have been assessed. A significant portion of them does not fully meet the exhaustive criteria desired for a quantum benchmark. Drawing from our own work and the guidelines outlined in [94], we conclude that an effective quantum benchmark should be unambiguous to clarify its derived metrics. It would ideally remain platform-agnostic, ensuring that its characterization of quantum capabilities is universal and not constricted to specific architectures or instructions. An essential facet is the ability of the benchmark to deduce optimal global properties, even when the data at hand is limited. That is, a benchmark should not vastly underestimate the capabilities of a device. Furthermore, it is crucial for a benchmark to offer predictive insights into a device's computational performance which are rigorously certified.

Building on these criteria, it is vital to analyse existing benchmarks with respect to their limitations. Some benchmarks are considerably specialized, focusing on certifying specific quantum processes or states rather than providing a global device evaluation [84]. Prominent techniques like quantum volume and mirrored quantum volume shed light on a quantum device's deviations when executing a generic type of random unitary defined by specific width and depth in a circuit, as illustrated in [85, 86]. More recently, even more, fragmented approaches have been introduced to assess larger quantum devices with limited connectivity, resulting in certifications that are less robust [87]. While they offer insights into how circuits of specific dimensions may be affected by noise, they do not provide the predictive power we deem desirable, especially regarding the fidelity of executing specific unitaries.

Application-based benchmarks gauge quantum device performance across various quantum algorithms and processes. They pinpoint the performance thresholds and highlight the deficiencies of metrics like quantum volume when applied to practical quantum solutions. As [88] elucidates, errors influence circuit compilation of certain quantum routines differently than they do generic random unitary operations of similar dimensions. This distinction sometimes undermines predictions from quantum volume metrics. However, these benchmarks have their inadequacies, as they are not algorithm-oblivious, potentially leading to biased results for devices reacting differently to distinct algorithmic solutions.
The CLOPS benchmark, as cited in [93], emphasizes the speed of quantum operations, shedding light on the time resources required for specific quantum tasks. This approach is motivated by classical benchmarking where the main focus of benchmarking lies on the speed of operations due to high confidence in the correct execution of operations. However, on quantum hardware fast execution alone cannot be used as an indicator for computational performance as all available devices struggle with noise.
The work of [89] underscores the potential of predictive benchmarks, utilizing both error models and neural networks. Notably, using quantum device data the error models excelled in determining realistic gate fidelities, sometimes diverging from manufacturer claims. Meanwhile, the neural network models predicted the success probabilities of specific unitaries, even ones not in the training data. However, both models have scalability questions unanswered and do not provide concrete guarantees on their predictions.
Lastly, verification protocols, as discussed in [90, 91, 92], offer a means to certify particular quantum algorithms or protocols with minimal overhead, even in the presence of noise. While yet to be adapted into general benchmarking with predictive attributes, they excel at offering a rigorous certification of a quantum algorithm's validity. These protocols harness the potential of blind delegation of computations, with embedded traps ensuring the algorithm's integrity. Recognizing deviations at these trap points offers insights into the overall fidelity of the process. We aim to integrate these techniques and protocols, crafting benchmarks that resonate with these attributes.

6.2.2 New paradigms for quantum benchmarking

Noise model agnostic. To assess the performance of quantum devices, certain assumptions regarding the impact of noise on computations have been posited. These assumptions fix a distinct noise model that may not accurately reflect the actual noise effects present in the device. One illustration is the employment of gate fidelities, operating as an unital noise model. While they can offer predictions on the fidelity of a unitary or a circuit, they overlook certain quantum device issues like crosstalk, which typically amplifies the noise interference in computations. In this context, we aim to introduce a benchmark devoid of any noise model assumptions, accommodating all potential noise disturbances.

Hardware agnostic. The benchmarking of quantum devices should be independent of the hardware platform. Therefore, the benchmark should remain unaffected by any provider's specific operations/instructions and must be framed to ensure this. This independence is crucial, especially given the ongoing exploration of multiple hardware platforms to establish a sense of comparability across hardware platforms. It is worth noting that while these benchmarks and ours are crafted for the end user, benchmarks that are hardware-specific could prove valuable for hardware providers, assisting in the evolution of their quantum devices.

Algorithm agnostic. Benchmarks should not determine the capabilities of a machine solely based on a collection of quantum algorithms or routines. While some classical benchmarks are constructed this way [88], our grasp of the quantum algorithmic landscape is still evolving, and recent discoveries of algorithms exhibiting exponential quantum advantages [114, 342] indicate the potential for a spectrum of quantum routines that offer computational benefits. Hence, it is premature to design machines optimized solely for the current known algorithms, potentially overlooking emergent solutions. A benchmark should aim at

providing strong statements about the performance of the device that should be useful as indicator across a multitude of possible applications. This underscores an advantage of our benchmark: it can determine the potential of new algorithmic approaches based solely on their circuit size/scaling, independent of specific instructions or architecture.

Predictive & **Certifiable.** Quantum benchmarks should not only rank quantum devices based on their efficacy in executing quantum computations but also offer insights into the performance of other quantum computations and routines one might wish to deploy on those devices. Essentially, they should globally profile the machine, equipping the end user with insights into the device's performance across diverse and user-specified applications. Beyond the given predictive capacity, there should be a level of assurance, or to put it more directly, a guarantee of accuracy. This ensures that the predictions align with a recognized benchmark. Our method fulfills both these criteria, differentiating it from previous solutions.

Explicit, Practical & **Scalable.** Lastly, a benchmark should be detailed, providing a straightforward set of instructions to execute without being overly complex. Also, it should not incur significant computational overheads that lead to an underestimation in the assessment of quantum devices. Finally, the overall computational resources demanded by the benchmark should grow at most polynomially, guaranteeing its feasibility and scalability beyond small-sized quantum devices.

6.2.3 Chapter Organization

In section 6.3 we introduce the technical definitions that we will work with throughout the paper. Notably, we formalize the behavior of quantum devices, which is needed to prove heuristic-free statements, and define quantum verification protocols in an abstract way. In section 6.4 we introduce a formal definition of a certifiably implementable class of computations (CICC) benchmark. We elaborate on which properties each of them should have to investigate the computability of a class of computations on a given quantum device. In addition, we prove that from any verification scheme one can efficiently compile a CICC benchmark. In section 6.4.4 we instantiate a concrete efficient CICC benchmark utilizing the robust VBQC protocol from [90].

6.3 Preliminaries

In this section, we aim to present a rigorous examination of the primary subject of our analysis. Initially, we will describe quantum devices, outlining only the essential requirements to ensure the hardware-agnostic proprieties. Following this, we will define quantum verification protocols, which will serve as the tools to assess the performance of quantum devices. Finally, we present the measurement-based quantum computation model which will be the fundamental model used to describe the respective protocols/benchmarks.

6.3.1 Quantum devices

In Quantum Benchmarking we want to assess the performance of a quantum device carrying out tasks that involve quantum information processing. For the end-user it is favorable not to make any assumptions on the quantum device and treat it as a black box to get unbiased performance indicators.
The benchmark that we propose justifies its statements about quantum devices formally which requires a formal description of the behavior of physical quantum devices. We introduce an interactive quantum algorithm (Definition 68) that describes the action of an a general quantum device partaking in an interactive protocol. The formalism we choose in this work is not unique but it covers the any behavior a quantum device can exhibit in a protocol type setting. We only consider the interfaces and treat the devices themselves as a black box because we do not want to make assumptions about the the inner workings

of a device to gauge performance in carrying out a desired task. An interactive quantum algorithm is designed to be part of a most general type of quantum protocols: In between the rounds of communication the parties/devices can act on the received quantum states with quantum channels and use an arbitrary quantum and classical memory to store information and entangle states from round to round. Classical information is also encoded in quantum states in this picture.

Definition 68 (Interactive Quantum Algorithm)**.** *Let $\mathcal{I}_i$, $\mathcal{O}_i$, and $\mathcal{D}i$ denote the input quantum states, output quantum states, and internal memory states of the quantum device in each round, respectively. Consider a sequence of n quantum channels, $\{\mathcal{E}^i : \mathcal{S}(\mathcal{I}_i \otimes \mathcal{D}_{i-1}) \longrightarrow \mathcal{S}(\mathcal{O}_i \otimes \mathcal{D}_i)\}_{i=1}^n$ represents the space of all density matrices over the Hilbert space $\mathcal{H}$. We define an algorithm that iteratively executes the following steps for each i from 1 to n:*

1. ***Receive Input***: ρ_I^i

2. ***Return Output***: $\rho_O^i = \mathrm{Tr}_{\mathcal{D}_i}[\mathcal{E}^i(\rho_I^i \otimes \sigma^{i-1})]$

3. ***Update Internal Memory***: $\sigma^i = \mathrm{Tr}_{\mathcal{O}_i}[\mathcal{E}^i(\rho_I^i \otimes \sigma^{i-1})]$

an interactive quantum algorithm **Q**.

The motivation for this very technical definition and somewhat trivial summary of interactive quantum algorithms is that we will need formal objects that can model the entire behavior of a physical quantum device to make formal assumptions and statements about their performance. One fundamental assumption that underlies our benchmark is that the interactive quantum algorithm **Q** that a quantum device implements stays the same over all interactions with the device. The formalism still allows for randomness in the behavior (quantum channels constituting **Q**) but the underlying quantum process must be consistent. This might seem like a very strong assumption but in fact it is necessary for any kind of benchmark: if the benchmarked device is changing after the benchmark, we have no information about the new behavior.
Note that we will also refer to the quantum device itself as **Q**. The difference of the quantum device to the interactive quantum algorithm is merely that a device repeats the same interactive quantum algorithm **Q** multiple times, just like a physical device would. We expect a device **Q** to act in the same way everytime we interact with it. It is also worth noting that not only the behavior of the quantum device that is to be benchmarked can be formalized like in definition 68 but actually the entire benchmarking protocol we cover in this work will follow this formalism.
However, we will not use this tedious notation to define the interactive algorithms that come up in this work but we will stick to a more intuitive, graphical depiction instead, which can be seen in figure 32.
When interacting two quantum algorithms we use the notation $\langle \mathbf{A}, \mathbf{B} \rangle$. Whenever we use this notation it is implicitly assumed that the interfaces of **A** and **B** match. The resulting structure will again be an interactive quantum algorithm if it has an open input/output interface on one or both sides. In favor of readability this will sometimes be neglected while at other points we will use the notation of an inward facing arrow for inputs and an outward facing arrow for outputs like so:

$$\text{output} \leftarrow \langle \mathbf{A}, \mathbf{B} \rangle \leftarrow \text{input} \tag{6.3.1}$$

It is also possible, if the communication rounds of **A** fit m times the one of **B**, to interact multiple versions of the same algorithm which we denote as $\langle \mathbf{A}, \mathbf{B}^m \rangle$.

6.3.2 Quantum Verification Protocols

Quantum verification protocols are designed for delegating quantum computations in a secure and certifiable way to a server or a device. To make them accessible for our benchmark we introduce verification in an interactive quantum algorithm picture. We define the client side of a verification protocol as interactive quantum algorithm **V** which delegates a computation from a class of computations: $C \in \mathfrak{C}$ via interacting with a quantum device **Q**. With a computation C we mean a classical description of a quantum computation, in most general terms: a CPTP map.

Definition 69 (Verification Algorithm)**.** *We call an interactive quantum algorithm* **V** *that:*

1. *Takes input* C*, the classical information describing a quantum computation in class* $\mathfrak{C}$
2. *Interacts with a quantum device* **Q**
3. *Outputs the result of the computation and a flag indicating whether to accept or reject the result*

At the same time it fulfills:

a)

$$\exists \delta > 0 : \forall C \in \mathfrak{C}, \forall \mathbf{Q} : \Pr[\text{acc} \wedge \text{wrong} \mid (\text{res}, \text{flag}) \leftarrow \langle \mathbf{V}, \mathbf{Q} \rangle \leftarrow C] \leq \delta \tag{6.3.2}$$

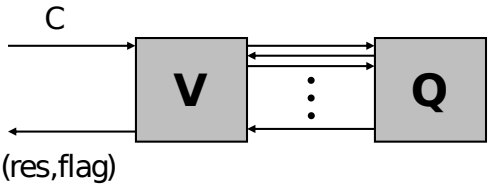
b)

$$\exists \mathbf{P} : \forall C \in \mathfrak{C} : \Pr[\text{acc} \mid (\text{res}, \text{flag}) \leftarrow \langle \mathbf{V}, \mathbf{P} \rangle \leftarrow C] = 1 \tag{6.3.3}$$

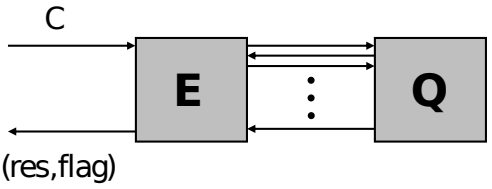
a verification algorithm **V**

Condition *a*. Guarantees that there is an upper bound $\delta > 0$ on the probability of erroneously accepting a wrong result. This upper bound should preferably approach zero in some parameter for example when scaling the size of the verification scheme.

Condition *b*. Makes sure that a verification scheme is accompanied by a prover algorithm **P**, representing an ideal quantum device flawlessly implementing the server side of the scheme. Also, it excludes schemes that reject always and thus trivially fulfill 6.3.2 while not being useful at all.
Provided the formal treatment of the verification algorithms that will be employed we do point to figure 32 for a graphical representation of the structure of these processes.



(a) Quantum verification as interactive quantum algorithm. This first input is a classical description of a quantum computation $C \in \mathfrak{C}$. Then it interacts with a quantum device **Q** to delegate the computation. As a last output it provides the result of the computation and a flag indicating accept or reject.



(b) Extractor Algorithm **E** delegating computation C to quantum device **Q** yielding result and flag: $(\text{res}, \text{flag}) \leftarrow \langle \mathbf{E}, \mathbf{Q} \rangle \leftarrow C$. The first input computation $C \in \mathfrak{C}$ is meant to come from the client and the result and flag are returned also to the client.

Figure 32: Quantum verification and extraction as interactive algorithms.

6.3.3 Measurement Based Quantum Computing and Universal Blind Quantum Computation

This introduction mainly follows [90]. Measurement based quantum computing (MBQC) utilizes gate teleportation to implement universal quantum computations based on graph-states as entanglement resource and single qubit measurements [343]. Graph states from graph $G = (V, E)$ where V is the set of vertices or nodes and E is the set of edges between them are defined as follows:

$$|G\rangle = \prod_{(a,b)\in E} CZ(a,b) \, |+\rangle^{\otimes V} \tag{6.3.4}$$

While it is perfectly possible to use quantum input and output computations with this model, in this work we restrict to the classical input/output case where we can carry out MBQC in a delegation scheme where the client only communicates the measurement angles to a large quantum server that can operate a graph state with large scale entanglement.
Computations are done in the MBQC picture as so-called measurement patterns. That means we choose a graph $G = (V, E)$ and a set of measurement angles $\{\phi(i)\}_{i\in V}$ for all the qubits in the graph state $|G\rangle$ and a flow function f which defines a partial order on the qubits of $|G\rangle$. Carrying out a computation then means measuring according to the flow f each qubit i of $|G\rangle$ in the X-Y plane according to the (corrected) measurement angles $\phi(i)$. Depending on the measurement outcome of each qubit, the computation will be carried out with a different set of measurement angles $\phi'(i)$ that are adaptively computed using the classical information of previous measurement outcomes we keep track of. That means that the initial set $\{\phi(i)\}_{i\in V}$ is only used if all the intermediate measurement results are zero.
MBQC can be realized in a delegation setting where we have a client with small quantum capacities i.e. the client is able to prepare and send single qubits states from a finite set. Then the server has universal quantum computational capacities and carries out the entangling operations needed for producing the graph states. Then the client communicates layer by layer (graph states can often be measured out layer-wise i.e. multiple qubits are measured simultaneously because their outcomes do not influence the measurement angles of each other) the measurement angles. Then the server carries out the measurement in the correct basis and communicates the outcome. Then the client updates the angles for the next layer until the entire graph state is measured out. The last layer will typically be the outcome of the computation.
Unlike one might assume from the graph state construction MBQC can be done with less operational qubits than constitute the graph state. Early qubits and later qubits do not necessarily be stored in quantum memory at the same time, but can be provided "on the fly" [344].
We provide the protocol of delegated measurement based quantum computing as the benchmarking protocol that we propose will be based on a slightly adapted version. We use the notation of [90] and particularly,

$$|\pm_\theta\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm e^{i\theta} \, |1\rangle). \tag{6.3.5}$$

The delegated version of MBQC that we discussed here allows for a client with limited quantum resources i.e. preparing and sending single qubit states to delegate classical input and output computations to a computationally more capable server.
However, in this way the server gets full access to the classical description of the data and algorithm involved in the computation. Especially in a cloud setting this is highly undesirable. Instead, what a client can do is to hide all of the information that is used in protocol 1 except for the graph G and flow f, which merely disclose the upper limit of resources employed by the server, by completely randomizing the communicated angles. We provide here the universal blind computation protocol from [345] but do not prove its desirable properties of composable security for delegating quantum computations in a blind way. We describe the protocol rather than explain it here. For a more rigorous understanding of the protocol

Protocol 1 Measurement Based Quantum Computing

1. Client sends Graph $G = (V, E)$, flow f and $S_X(j)$ and $S_Z(j)$, the subsets of qubits that influence the angle of qubit j, derived from G and f to the server
2. Server prepares $|G\rangle$
3. Client sends measurement angles $\{\phi(i)\}_{i\in V}$ to server which encode the target computation
4. Server measures qubits in the order induced by f in the basis $\left|\pm_{\phi'(i)}\right\rangle$. $\phi'(i)$ is computed by the server as

$$\phi'(i) = (-1)^{s_X(i)}\phi(i) + s_Z(i)\pi \tag{6.3.6}$$

$$s_X(i) = \bigoplus_{j\in S_X(i)} b(j), \quad s_Z(i) = \bigoplus_{j\in S_Z(i)} b(j) \tag{6.3.7}$$

   where $b(j)$ is the measurement outcome of qubit j.
5. Server communicates last measurement layer to the client as result of the target computation

---

the reader is referred to [345]. Again we use the notation of [90].

Protocol 2 Universal Blind Quantum Computing

1. Client sends Graph $G = (V, E)$ and flow f to the server
2. Client prepares and sends all the qubits i in V as either
   - $\left|+_{\theta(i)}\right\rangle$ where $\theta(i)$ is randomly sampled for every qubit $i \in V$ from $\theta = \{0, \frac{\pi}{4}, \ldots, \frac{7\pi}{4}\}$
   - or $X^{r(i)}|0\rangle$ where $r(i)$ is randomly sampled from $\{0, 1\}$. For X-Y basis qubits $r(i) = 0$ per default.
3. Server entangles all the qubits just as in protocol 1 $|G\rangle$
4. Client sends measurement angles $\{\delta(i)\}_{i\in V}$ to server for every measurement, right before the measurement. $\delta(i)$ is computed by the client as:

$$\delta(i) = (-1)^{s_X(i)}\phi(i) + s_Z(i)\pi + \theta(i) + r(i)\pi \tag{6.3.8}$$

$$s_X(i) = \bigoplus_{j\in S_X(i)} b(j), \quad s_Z(i) = \bigoplus_{j\in S_Z(i)} b(j) \tag{6.3.9}$$

   where $b(j)$ is the measurement outcome of qubit j.
5. Server measures qubits in the basis $\left|\pm_{\delta(i)}\right\rangle$ and communicates the result directly to the client who uses the information for the next measurement layer.
6. Server communicates last measurement layer to the client as result of the target computation.

6.4 A formal treatment of benchmarking

On a more abstract level, benchmarking can be described as the characterization of quantum processes. The most comprehensive form of this characterization is quantum process tomography, which, unfortunately, is not feasible due to poor scaling behavior. However, benchmarking can be viewed as an attempt to ascertain the quantum information processing capabilities of a quantum device using fewer resources than tomography requires. To achieve this, a practical benchmark is divided into two stages. In the first stage, the client runs an efficient benchmarking algorithm, **B**, interacting with the quantum device, **Q**, that is undergoing benchmarking. The second stage involves delegating actual computations to the quantum device via an extractor algorithm, **E**. The extractor serves as a set of instructions, specifically the quantum channels the client uses to convey their quantum computations to the device. The insights gained by the client in the first stage provide a lower bound on the success probability of executing computations belonging to class $\mathfrak{C}$. Any non-zero lower bound on this success probability indicates that computations of class $\mathfrak{C}$ can be made computable through repetition.
Expanding upon this initial overview, we will refine the precise definitions, properties, and their relationships to verification, which will be segmented into the following subsections.

- In section 6.4.1 we make the formal foundation for our benchmark. We start by defining the form of extractor algorithm which makes concrete how the client will implement computations on a device. Secondly, we introduce the form of our benchmark in the picture of interactive quantum algorithms, without fixing any properties yet.
- We go on in section 6.4.2 to define properties of a before defined benchmark that are desirable. The properties are chosen such that it will be easy to prove them in the following section. The important part is what follows from fulfilling all of them which will also be explained.
- Finally, in section 6.4.3 we introduce a scheme of utilizing multiple verification protocols that qualifies as a CICC benchmark in the sense that we defined. We go on to show that it fulfills all the properties that we want from the benchmark to be useful.

6.4.1 Definitions

The goal of any benchmark is to make statements about the performance of quantum devices on tasks that require quantum information processing. The exact statements that we will make concern what we call the computability of computational classes. Additionally, we aim to remove any heuristics from our benchmark. We achieve this by introducing the extractor algorithm to anchor the computability statements: If a device **Q** is certified via our benchmark for a class of computations $\mathfrak{C}$, then there exists an interactive quantum algorithm (the extractor **E**) which the client can use to delegate any computation in $\mathfrak{C}$ to the device and is guaranteed formally to get the correct result in at most a small number of tries with no heuristic arguments involved. For a more intuitive understanding of what the extractor algorithm does, we include a graphical depiction of it in figure 32 b). Furthermore, we define the interactive quantum algorithm that the client ideally expects from the device in the delegation protocol as the prover algorithm **P**. Thus, we assume $\langle \mathbf{E}, \mathbf{P} \rangle$ perfectly implements any computation in $\mathfrak{C}$. We can now define our proposed benchmark.

Definition 70 (Certified Implementable Class of Computations Benchmark)**.** *Let $\mathfrak{C}$ be a class of computations. We call the tuple of interactive quantum algorithms $(\mathbf{B}, \mathbf{E}, \mathbf{P})$ where*

- **B***, referred to as the benchmarking algorithm, engages in interactions with a quantum device* **Q** *for a total of m times, ultimately producing an output $z \in [0,1] \leftarrow \langle \mathbf{B}, \mathbf{Q}^m \rangle$.*

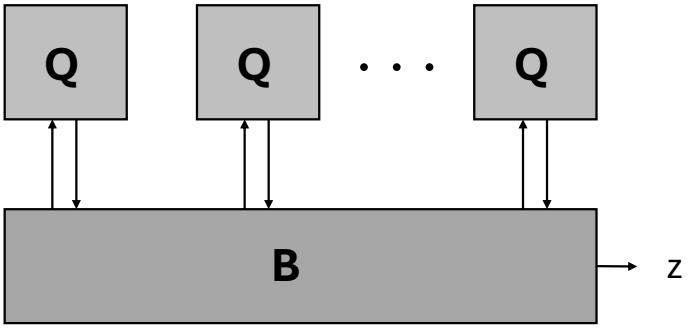


Figure 33: Benchmarking Algorithm **B** interacting with a quantum device **Q** m times outputting information $z \in [0, 1]$ that is based on **Q**.

- **P**, *known as the prover algorithm, outlines the expected behavior of a quantum device within the delegation scheme of* **E**.
- **E**, *functioning as the extractor algorithm, takes input* $C \in \mathfrak{C}$*, engages in interactions with a quantum device, and ultimately produces the result along with a flag indicating acceptance:* $(\text{res}, \text{flag}) \leftarrow \langle \mathbf{E}, \mathbf{P} \rangle$*. It consistently accepts and is accurate for all* $C \in \mathfrak{C}$*.*

a Certified Implementable Class of Computations Benchmark.

We implicitly assume that all three algorithms that constitute the benchmark match on their interfaces. The structure of the benchmarking algorithm is depicted in figure 33.
The final goal of our benchmark will be to output a $z \in [0, 1]$ that is a lower bound on the success probability of running any computation $C \in \mathfrak{C}$ on **Q** via the extractor **E**. However, at this point it is not clear yet how this property will be achieved. As of now, we merely created the abstract form of a CICC benchmark in a protocol picture. In the next section (6.4.2) we will define properties that a CICC benchmark should fulfill. It becomes useful for practical applications only if all the properties are really fulfilled. For now we want to give a few more insights on the concepts that are already introduced.

Classes of Computations The class of computations $\mathfrak{C}$ is always linked to the choice of the extractor algorithm. For example, any delegation scheme (extractor) of finite size necessarily restricts the size of computations. In this work we also restrict everything to classical input / classical output quantum computations. Note that we only handle the classical description C of these computations in the communication to the extractor and we assume this description to be efficient.

When comparing devices with a CICC benchmark, the idea is to check which device can successfully implement a larger set of computations $\mathfrak{C}$ or, as the sets might not be inclusive, which device can implement a set that includes a desired subset of computations $\mathfrak{c}$.

Benchmarking Stage. For the CICC benchmark it is important that the quantum device **Q** stays the same through out the interaction with **B** as well as extending to the computational stage when the client performs computations via $\langle \mathbf{E}, \mathbf{Q} \rangle$. Every interaction with the device must implement the same interactive quantum algorithm **Q** on the device side. In particular, a device has no historic data to adapt its strategy. This is vital for the integrity of the benchmark.

Extractor Algorithm. The choice of the extractor is crucial. The more resources the extractor uses, the less powerful the statement about the quantum device is. That is because the more quantum resources the extractor uses to delegate a computation the less quantum resources need to be employed on the device side to implement computations. Thus the overhead of the extractor is important. Also, the resources which a client has access to are determining which extractor to use in a benchmark.

6.4.2 Properties of a CICC Benchmark

In this section, we introduce the essential properties that a CICC benchmark must fulfill and explain why these properties lead to a useful benchmark. The first property we discuss is correctness, which ensures that an optimal device receives an optimal score with high probability. This is a fundamental expectation of any effective benchmark, as it stands to reason that a perfect device should score perfectly.

Definition 71 (Correctness)**.** *We say Benchmark* $(\mathbf{B}, \mathbf{E}, \mathbf{P})$ *is* α*-correct for a shift of* β *if*

$$\Pr[z \geq 1 - \beta \mid z \leftarrow \langle \mathbf{B}, \mathbf{P^m} \rangle] \geq 1 - \alpha \tag{6.4.1}$$

The reason why we introduce the β-shift is to add some kind of robustness. For a process that inherently involves randomness we do not want to include a hard condition such as $z = 1$. It also suffices to ensure there is a high probability $(\geq 1 - \alpha)$ to score close to optimal $(\geq 1 - \beta)$ with the optimal device **P**.
Recall from the definition of the CICC benchmark that the ideal device implementing the prover algorithm **P** is simultaneously the perfect counterpart to the extractor algorithm **E**, which, by definition, has a success probability of exactly 1 and should, therefore, score very high on the benchmark. Building on this, the next two properties introduce what we call the extractability of the benchmark. This property aims to encapsulate the client's ability to obtain correct results, which are also flagged as "accept", for the computations they implement ensuring that the benchmark not only measures performance but also the practical usability of the device in delivering reliable outcomes.

Definition 72 (Correctly-Flagged Extractability)**.** *We say a benchmark* $(\mathbf{B}, \mathbf{E}, \mathbf{P})$ *is* γ*-correctly-flagged extractable if for all* **Q** *of the form of* **P** *it holds that:*

$$\forall C \in \mathfrak{C} : \Pr[\Pr[\text{flag} = \text{acc} \mid (\text{res}, \text{flag}) \leftarrow \langle \mathbf{E}, \mathbf{Q} \rangle] \leq z \mid z \leftarrow \langle \mathbf{B}, \mathbf{Q}^m \rangle] \leq \gamma \tag{6.4.2}$$

This property of a CICC benchmark refers to the computation stage where the client implements a computation $C \in \mathfrak{C}$ via the extractor **E** on the quantum device **Q**. It makes sure that is unlikely (small γ) that the acceptance rate is lower than the value z that the benchmark is suggesting. Also, note that this will be the only property that actually links the benchmarking algorithm **B** to the extractor algorithm **E**.
We are now only missing a property that we call verifiable extractability:

Definition 73 (Verifiable Extractability)**.** *We call benchmark* $(\mathbf{B}, \mathbf{E}, \mathbf{P})$ δ*-verifiably extractable if for all* **Q** *of the form of* **P** *it holds that:*

$$\forall C \in \mathfrak{C} : \Pr[\text{res} = \text{wrong} \wedge \text{flag} = \text{acc} \mid (\text{res}, \text{flag}) \leftarrow \langle \mathbf{E}, \mathbf{Q} \rangle] \leq \delta \tag{6.4.3}$$

We use the notation "res $=$ wrong" to denote the stochastic event where the result returned by the interaction between the extractor and the device, given an input $C \in \mathfrak{C}$, is incorrect. Since we are dealing with classical input/classical output quantum computations, the distinction between "correct" and "wrong" results is clearly defined. This property pertains exclusively to the extraction algorithm, ensuring that the likelihood of obtaining an incorrect result that is mistakenly marked as correct is low (denoted by a small δ).

If a CICC benchmark fulfills properties 72 and 73 then the retrieved value of z can be used as a lower bound on the probability of getting a (flagged) correct result on a computation. To see this we follow the subsequent argument [1].
Correctly-flagged extractability 72 gives us directly:

$$\Pr[\Pr[\text{acc}] > z] > 1 - \gamma \tag{6.4.4}$$

[1] Note that we have introduced an abbreviated notation where "acc" refers to the probability of success for implementing a computation via the extractor.

At the same time verifiable extractability 73 gives us:

$$\Pr[\text{wrong} \wedge \text{acc}] \leq \delta \tag{6.4.5}$$

Now we use:

$$\Pr[\text{acc}] = \Pr[\text{correct} \wedge \text{acc}] + \Pr[\text{wrong} \wedge \text{acc}] \tag{6.4.6}$$

Combining equations 6.4.4 - 6.4.6 we find:

$$\Pr[\Pr[\text{correct} \wedge \text{acc}] > z - \delta] > 1 - \gamma \tag{6.4.7}$$

This means that z, retrieved from the benchmark, shifted by δ, serves as a lower bound for the fraction of accepted correct results with probability of $1 - \gamma$. If a CICC benchmark meets all these criteria and returns a non-zero success probability z, then by repeating the computation with **E** multiple times, the likelihood of not obtaining a single correct result decreases exponentially with the number of runs, provided $z - \delta > 0$. The convergence speed depends on z.
Even though the original output of the benchmark is the lower-bound z, it is more interesting to note that as long as the success probability is non-zero then, with a small overhead in time (and not in the size of the quantum device), the device can certainly carry out every computation from the class $\mathfrak{C}$. Thus, one can hand out a certificate of computability for a device that was benchmarked in this way.

Remark. We want to stress that a specific benchmark that comes with values of δ and γ is only useful if there is a systematic way how to make these values approach zero exponentially fast. The benchmark that we suggest will achieve this via a repetition scheme where both δ and γ will be negligible in the number of rounds.

6.4.3 CICC Benchmark from Verification

In this subsection, we delve into the process of constructing a CICC benchmark based on an arbitrary verification scheme, **V**, as introduced in section 6.3.2, which is pertinent to classical input classical output quantum computations. The core concept is fairly simple: we employ a verification scheme repeatedly to estimate its success rate. Following this, we delineate the CICC benchmark formally as $(\mathbf{B_V}, \mathbf{V}, \mathbf{P})$, ensuring it fulfills the critical properties 71 - 73, and explain the constituents subsequently.

Benchmarking algorithm $\mathbf{B_V}$. The benchmark repeats verification protocol **V** m times on quantum device **Q**. Then, it returns $z \in [0, 1] = \frac{\#acc}{m} - \beta$ the fraction of accepted results, shifted by $\beta > 0$ so that z serves as a lower bound on the success probability estimate of the verification protocol **V**. We provide a schematic of the verification based benchmarking algorithm in34.

Extractor algorithm V. The extractor algorithm which is needed to define the CICC benchmark is simply the verification protocol: $\mathbf{E} = \mathbf{V}$.

Prover algorithm P. The prover algorithm **P** is the same that comes with the verification protocol, which was explained in section 6.3.2.

We now go on to prove that the benchmark defined by $(\mathbf{B_V}, \mathbf{V}, \mathbf{P})$ fulfills properties 71 - 73.

Proof of Correctness. Per definition, a verification protocol has an ideal counterpart algorithm that it can interact with perfectly i.e. $\mathbf{P} : (res, flag) \leftarrow \langle \mathbf{V}, \mathbf{P} \rangle$ gives the correct result and accepts at the same time deterministically. Therefore we trivially get that $\#acc = m$ and thus $z = 1 - \beta$ will be returned deterministically. Here, we made the choice of β in the definition of the CICC benchmark deliberately such

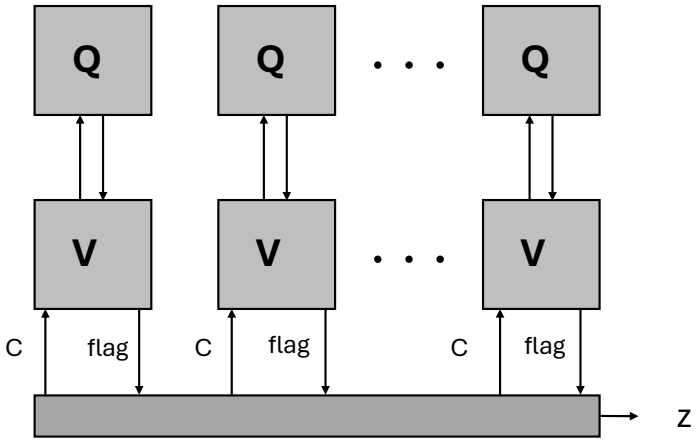


Figure 34: Verification-inspired CICC benchmarking algorithms $\mathbf{B_V}$ utilizing m rounds of the verification algorithm $\mathbf{V}$

that it matches the one used in the correctness property. We get:

$$\mathsf{Pr}[z \geq 1-\beta \mid z \leftarrow \langle \mathbf{B_V}, \mathbf{P}^m \rangle] \geq \mathsf{Pr}[z = 1-\beta \mid z \leftarrow \langle \mathbf{B_V}, \mathbf{P}^m \rangle] = 1 \geq 1-\alpha \tag{6.4.8}$$

This makes the benchmark α-correct with shift β for any $\alpha \geq 0$. □

Proof of Correctly-Flagged Extractability. During the Benchmark we execute $\langle \mathbf{V}, \mathbf{Q} \rangle$ m times. Both interactive quantum algorithms $\mathbf{V}$ and $\mathbf{Q}$ stay the same that means that the probability which underlies the sampling of $flag \leftarrow \langle \mathbf{V}, \mathbf{Q} \rangle$ stays the same everytime. Also it is determined fully by $\mathbf{V}$ and $\mathbf{Q}$ independent of the computation C which is chosen in the beginning of the benchmark due to the blindness of the verification protocol. This means that we effectively have m rounds of sampling the same, independent binary probability distribution of accept/reject. The Hoeffdings inequality for IID binary binomial distribution is explained in Appendix B.1. We identify accept with the outcome 1 and reject with the outcome 0. Then we are exactly in the case that is discussed in the appendix. We can just take the inequality that results for m times repeating a biased coinflip with probability p of obtaining 1 and threshold β: $0 < \beta < \frac{\#acc}{m}$ and $z = \frac{\#acc}{m} - \beta$:

$$\mathsf{Pr}[p \leq z] \leq \exp\left(-2m\beta^2\right) \tag{6.4.9}$$

Where random variable $z \in [0,1]$ is the averaged value after sampling m times the same coin flip shifted by small β with unknown probability p.
Note that in the case of our benchmark p the probability of getting result 1 is assigned to the probability of getting an accept. Thus we will replace p in equation 6.4.9 with $\mathsf{Pr}[\mathsf{flag} = \mathsf{acc} \mid \mathsf{flag} \leftarrow \langle \mathbf{V}, \mathbf{Q} \rangle]$. At the same time we identify the threshold parameter β as the one from correctness. This gives us:

$$\mathsf{Pr}[\mathsf{Pr}[\mathsf{flag} = \mathsf{acc} \mid (\mathsf{res}, \mathsf{flag}) \leftarrow \langle \mathbf{V}, \mathbf{Q} \rangle] \leq z \mid z \leftarrow \langle \mathbf{B_V}, \mathbf{Q}^m \rangle] \leq \exp\left(-2m\beta^2\right) \tag{6.4.10}$$

Which is nothing else than γ-correctly-flagged extracability for $\gamma = \exp(-2m\beta^2)$

□

The effect of β is the following: If we want to find out the highest lower-bound on the success probability we will decrease β but to achieve the same confidence we will then need a squared increase in rounds which comes from the exponent in 6.4.10. Usually β does not need to go to zero though because a finite β will suffice to obtain a reasonable lower-bound.

Proof of Verifiable Extractability. The δ-verifiable extracability property comes directly from the δ-verifiability of the verification protocol that is used as explained in 6.3.2.

□

6.4.4 Optimized Benchmarking Protocol from Robust VBQC

In this section, we adapt the generic CICC benchmark (section 6.4.3) using the robust VBQC protocol from [90]. Due to the round-based structure of the robust VBQC protocol, we can reduce the time overhead during the benchmarking stage to get a slightly more efficient version of the generic CICC benchmark.
Notice that to avoid confusion, we stick to the following notions. When talking about "rounds" we refer to the sub-graph rounds in the robust VBQC MBQC pattern. The number n will always refer to the number of rounds that constitute the verification scheme. Also, when referring to "repetitions" we mean the number of times a benchmarking algorithm is repeating a verification scheme or interacting with a quantum device m times.

Recap: robust VBQC protocol

The robust VBQC protocol was introduced in [90]. It is written in the Measurement Based Quantum Computing framework [343] where computations are mapped to feed forward measurement sequences on large graph states. It is a provably secure (in the Abstract Cryptography framework) protocol to delegate quantum computations from a client that can only prepare and send single qubit states to a server that can prepare large scale entangled resource states (graph states) and use the clients input qubits to calculate the desired computation via feed forward measurements. The robust VBQC protocol uses the universal blind computation protocol from [345] to delegate computations in a completely blind way, making it impossible for the server to distinguish between different interactions as long as the underlying graph state remains the same.

The Protocol The Robust VBQC protocol has a round based structure. For n rounds the client implements a blind MBQC pattern where it delegates its target computation to the server on the same graph state (subgraph G). A fraction of the rounds are randomly chosen by the client to feature a trap computation instead. Those are computations with a deterministic outcome known to the client. If the server's result deviates from the expected one the server's malicious or faulty behavior is revealed. As all the rounds are indistinguishable for the server, deviations must affect trap and computational rounds equally. The indistinguishability leads to an effective error model on subgraph G that is sampled n times that is, the error probability distribution is identical and independent in each round. After running n rounds (computations and traps) the client counts the number of total activated traps. If the trap activation rate is below threshold $c_t \cdot n$ the client accepts the result, performing a majority vote over all computational rounds. The trap computations are chosen in such a way that the probability of the joint event of the trap activation rate being below the threshold (i.e. client accepting result) while the result of the majority vote is actually not the correct result of the desired computation is negligible in n:

$$\Pr[\text{accepts} \wedge \text{wrong}] \propto \exp\left(-n^2\right) \tag{6.4.11}$$

Thus, a small overhead in time, not in the size of the quantum system (in the form of subgraph G), yields exponential security in the correctness of an accepted computation. Thus, the protocol can verifiably delegate quantum computations to the server without hardware overheads.

Optimizing the Generic CICC Benchmark for Robust VBQC

We propose here a concrete efficient benchmarking protocol that utilizes the round-based robust VBQC protocol from [90]. It slightly differs from the generic case that we investigated in section 6.4.3 in favor of only using trap computations.
We choose to instantiate our CICC benchmark from 6.4.3 with the robust VBQC protocol because it is efficient. It only uses a small overhead in time and not in the size of the quantum system. Also, the round-based structure allows for a slight optimization that improves the confidence level. We could instead simply use the generic CICC benchmark from section 6.4.3 to construct from robust VBQC protocol $\mathbf{V}_n$ a benchmark $(\mathbf{B}_{\mathbf{V}_n}, \mathbf{V}_n, \mathbf{P})$. However, this only achieves a γ-correctly-flagged extracrability with $\gamma = \exp(-2m\beta^2)$, where the n-round verification protocol $\mathbf{V}_n$ is repeated m times. We will show that there is an optimized version of the benchmark which achieves $\gamma = \exp(-2mn\beta_t^2)$ decreasing the number of times m we need to run the verification protocol to get the same confidence in the result of the benchmark.

Optimized CICC Benchmark from robust VBQC

We instantiate a CICC benchmark $(\mathbf{B}_{\mathbf{V}_n}, \mathbf{V}_n, \mathbf{P})$ and slightly alter the benchmarking algorithm $\mathbf{B}_{\mathbf{V}_n}$:

- For all m repetitions the computational rounds of $\mathbf{V}_n$ are used for trap computations, randomly sampled from the trappified scheme that comes with $\mathbf{V}_n$, leaving all the $m \cdot n$ sub-graphs to implement trap computations.
- Instead of counting accepts vs. rejects it counts the total number of activated traps
- Defines small shift β_t such that $0 < \beta_t \ll |\frac{\#activated}{m \cdot n} - c_t|$ where c_t is the threshold trap activation rate from the robust VBQC protocol
- Defines $z_t = \frac{\#activated}{m \cdot n} + \beta_t$ which serves as upper bound on trap activation rate
- If $z_t < c_t$ the benchmarking algorithm returns $z = 1 - \exp(-2n^3(c_t - z_t)^2)$ as the lower bound on the success probability
- If $z_t \geq c_t$ it returns $z = 0$

We call the updated benchmarking algorithm $\hat{\mathbf{B}}_{\mathbf{V}_n}$. The optimized benchmark differs merely in the classical interpretation of the data that comes back from the quantum device during the benchmark. Introducing trap computations in every computational round is very reasonable because trap computations are just one special type of possible computations and in the generic form 6.4.3 these rounds are discarded. We acknowledge that the trap activation rate of a device is the quantity which governs the success probability of running the robust VBQC protocol on it. Therefore, we use every $m \cdot n$ sub-graphs to gauge this quantity. We now prove that the optimized benchmark $(\hat{\mathbf{B}}_{\mathbf{V}_n}, \mathbf{V}_n, \mathbf{P})$ fulfills all the properties from section 6.4.2.

Proof of Correctness. This proof is analogous to the generic case in section 6.4.3. When interacting with the ideal device, not a single trap will ever activate such that the returned value from the benchmark will always be $z = 1 - \exp(-2n^3(c_t - \beta_t)^2)$. To see why this concludes the proof refer to the proof of correctness in 6.4.3. □

Proof of Correctly-Flagged Extractability. We first note that the the probability of a trap activating is the same on all of the $m \cdot n$ sub-graph trap computations. Why? By definition, the behavior of the device is unchanged for all m repetitions of the all-trappified robust VBQC protocol. In section 6.4.4 we also explained that for a robust VBQC protocol $\mathbf{V}_n$ the global error model, that defines probability of errors on G^n, can be reduced to a local error model on the sub-graph G that is sampled n times. Effectively, we

therefore have $m \cdot n$ independent trials of randomly sampling both trap and error for each sub-graph state G. Every of these $m \cdot n$ rounds will have the same probability of activating a trap. This probability p_t is unknown upfront. However, we are again in the situation of $m \cdot n$ Bernoulli trials.
We define $z_t = \frac{\#activated}{m \cdot n} + \beta_t$ for some $0 < \beta_t \ll |\frac{\#activated}{m \cdot n} - c_t|$. Hoeffding's inequality yields for trap activation probability p_t in each round:

$$\mathsf{Pr}[p_t < z_t] > 1 - \exp\left(-2mn\beta_t^2\right) \tag{6.4.12}$$

Now we want to use 6.4.12 to find out the acceptance probability of $\mathbf{V}_n$. Whether the robust VBQC accepts or not is determined by the number of activated traps. We recognize the execution of $\langle \mathbf{V}_n, \mathbf{Q} \rangle$ again as an n round Bernoulli trial with random variables X_i:

$$\forall i \in [n] : X_i \sim \begin{cases} 1 & \text{with probability } p_t \\ 0 & \text{with probability } 1 - p_t \end{cases} \tag{6.4.13}$$

where we identify outcome 1 with and activated trap and 0 for not activating a trap for each round. We denote the precise condition under which the verification scheme accepts:

$$S_n = \sum_{i=1}^{n} X_i < c_t \cdot n \tag{6.4.14}$$

Where the exact value of c_t is defined in the specific choice of the scheme. Refer to 6.4.4.
Just like in the definition of $\hat{\mathbf{B}}_{\mathbf{V}_n}$ we now make a case distinction:

Case 1: $z_t < c_t$ From 6.4.12 we know that with probability $p_t < z_t < c_t$ which is what assume for the following argument. We use Hoeffdings inequality:

$$\mathsf{Pr}[S_n \geq \mathbb{E}[S_n] + t] \leq \exp\left(-2nt^2\right), t > 0 \tag{6.4.15}$$

We note that $\mathbb{E}[S_n] = np_t$ and set $t = n(c_t - p_t) > 0$ to obtain:

$$\mathsf{Pr}[S_n \geq nc_t] \leq \exp\left(-2n^3(c_t - p_t)^2\right) \tag{6.4.16}$$

However, $S_n \geq nc_t$ is exactly the condition under which $\mathbf{V}_n$ **rejects** the result. We conclude:

$$\mathsf{Pr}[\mathbf{V}_n \text{ accepts}] > 1 - \exp\left(-2n^3(c_t - p_t)^2\right) \overset{z_t > p_t}{>} 1 - \exp\left(-2n^3(c_t - z_t)^2\right) \tag{6.4.17}$$

Now we remember that the assumption $p_t < z_t$ was actually only justified with high probability (6.4.12). We get:

$$\mathsf{Pr}\left[\mathsf{Pr}[\mathbf{V}_n \text{ accepts}] > 1 - \exp\left(-2n^3(c_t - z_t)^2\right)\right] > 1 - \exp\left(-2mn\beta_t^2\right) \tag{6.4.18}$$

which is equivalent to :

$$\mathsf{Pr}\left[\mathsf{Pr}[\mathbf{V}_n \text{ accepts}] \leq 1 - \exp\left(-2n^3(c_t - z_t)^2\right)\right] \leq \exp\left(-2mn\beta_t^2\right) \tag{6.4.19}$$

We remind ourselves of the output of $\hat{\mathbf{B}}_{\mathbf{V}_n}$: $z = 1 - \exp(-2n^3(c_t - z_t)^2)$. We can directly identify correctly-flagged extractability from this expression in the case of $z_t < c_t$.

Case 2: $z_t \geq c_t$ The benchmark simply returns $z = 0$. It trivially holds:

$$\Pr[\Pr[\mathbf{V}_n \text{ accepts}] \leq 0] = 0 < \exp\left(-2mn\beta_t^2\right) \tag{6.4.20}$$

We combine **Case 1** and **Case 2** to conclude that for any value z_t, and therefore in general, the CICC benchmark defined by $(\hat{\mathbf{B}}_{\mathbf{V}_n}, \mathbf{V}_n, \mathbf{P})$ is γ correctly-flagged extractable with $\gamma = \exp(-2mn\beta_t^2)$. □

Proof of Verifiable Extractability. Verifiable extractability only considers the extracting algorithm alone which is the same as in the generic case of section 6.4.3. □

Remark. One might wonder why we define the benchmarking algorithm to return $z = 0$ if $z_t > c_t$. The reason is that if the trap activation probability is above the threshold for the scheme, then increasing the number of rounds n which is needed to make the verification scheme work in the first place, diminishes the success proability as is to expected. The inherent border of computable vs. not computable for the robust VBQC protocol lies in the question whether the trap activation lies above or below the threshold of the scheme. It is the properterty that ultimately defines if $z \mapsto 1$ or $z \mapsto 0$ exponentially fast.

6.5 A Concrete Efficient CICC Benchmarking Protocol

In Section 6.4, we have established an abstract link between quantum verification and quantum benchmarking. In particular, we have shown that it is possible to compile a benchmarking protocol from any verification protocol with sufficiently strong security, by reinterpreting its results. This will allow any further improvements on the side of quantum verification to carry over directly to the realm of benchmarking, through our generic compiler.
While Section 6.4 treated benchmarking in a very abstract way that shows conceptual advantages of reinterpreting verification as benchmarking, in this section we aim to give an explicit and self-contained description of a concrete benchmarking protocol without the abstract language of the previous section. Security and correctness of the presented protocol directly follow from the proofs given in Section 6.4. The description of the concrete CICC benchmarking protocol based on robust VBQC protocol is given as Protocol 3.

Protocol 3 Efficient CICC Benchmark from Robust VBQC Protocol

Parameters:

- A graph G.
- A distribution of trap computations $\mathcal{D}$, and a threshold $\omega \in [0, 1]$.
- A number of repetitions n.

The protocol:

1. The client repeats n times:
   - Randomly sample a trap computation $T \sim \mathcal{D}$.
   - Delegate the trap computation T to the quantum device, using the UBQC Protocol 2.
   - Check the obtained measurement outcomes.
2. The client computes $z = \frac{\#\text{failed traps}}{n}$.
3. Iff $z < \omega$, the client returns Accept, otherwise Reject.

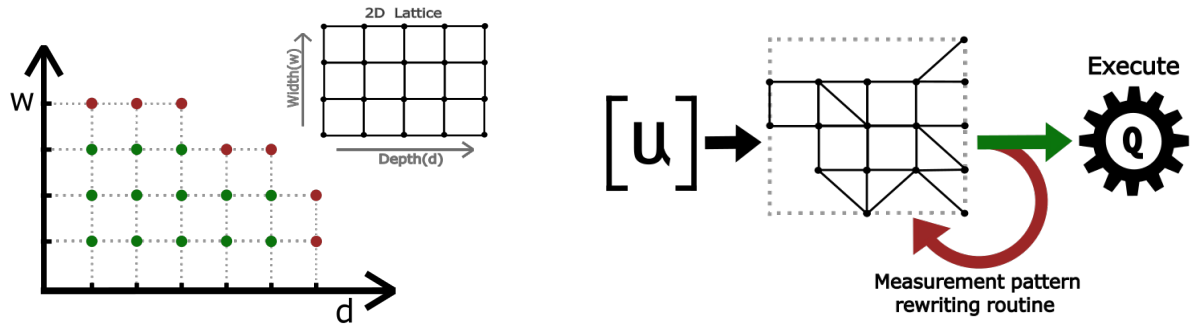
The choice of the trap computations and the threshold will dictate the confidence in the result of the protocol and the guarantees that the benchmark can provide. As a simple but efficient example, we refer to the generalized stabilizer traps from [90]. To sample from this distribution of traps, uniformly at random sample a stabilizer of the graph state $|G\rangle$ and delegate the corresponding stabilizer measurement to the quantum device, after randomization by virtue of the UBQC protocol. On a perfect and noiseless device, these measurements yield deterministic outcomes and can therefore be checked for errors. If the outcome of the stabilizer measurement is not as expected, the trap is considered to have failed. As generalized stabilizer traps have an error detection rate of $1/2$, any threshold $\omega < 1/4$ yields a secure protocol, with the gap between ω and $1/4$ and the number of repetitions n dictating the security error. For a full analysis of the security error, we refer to [90].
In this sense, one can also understand this benchmarking protocol as a kind of resource state certification process with an added layer of randomization. The resource states that we certify are graph states, and are checked for their usefulness to implement MBQC computations correctly.

6.5.1 Graph State based Quantum Device Characterization

The specific proposal for a CICC benchmark, as introduced in section 6.4.4, was derived from formal verification protocols. It certifies computational classes $\mathcal{C}$ based on the graphs that result from translating our initial computation or unitary operation into the MBQC model. Furthermore, the deployment of our benchmark can be viewed as a process designed to characterize quantum devices in terms of the computational classes they can successfully implement. This process also involves listing these classes to provide valuable information for any future users of the devices.
In particular, as the classes under our proposal take the form of graphs, the test could be derived by benchmarking graphs of sequentially larger sizes. However, the search space does grow rapidly to an exponential size. Nevertheless, we propose, as a very informative and general characterization, the use of 2D cluster states, parameterized by two values: its width and depth. This choice results from the fact that the previous resource state is a universal resource state for the MBQC model, and the structure of the graphs resembles the characterization provided by previous benchmarks such as quantum volume, while this time providing some formal guarantees. This benchmark for quantum devices would provide a figure of merit in the form represented in Figure 35a.
Furthermore, we would like to highlight that with such a characterization, a device user would be able to tell if its computation will be executed correctly on the device. To determine this, it is only necessary to translate the unitary into a planar MBQC measurement pattern and check if the resulting graph is equal to or a subgraph of a certified graph. Additionally, if the initial graph derived does not match any certified graph, the unitary could be rewritten according to specific rules as a deeper or more parallel computation, which might then correspond to one of the verified graphs (Figure 35b). This further integrates the benchmark into a practical pipeline of processes that allow for optimal resource usage and computational applications of quantum devices.
Remark. It's important to note that the chosen graph structure is not the only possibility; the benchmark's versatility accommodates various exhaustive characterizations. Device manufacturers might tailor their device benchmarks to reflect the qubit connectivity, offering a precise characterization of quantum devices. This specificity fosters the development of circuit synthesis routines optimized for individual quantum devices, enhancing their computational performance.

6.6 Discussion

In previous verification protocols, the focus has been on establishing the correctness of specific computational instances. Our work extends this concept by eliminating the need for executing computational



(a) Depiction of a device characterization using tests based on graph states on 2D lattices.

(b) Illustration of a methodology for compiling any unitary operation onto a graph state within an MBQC pattern, accompanied by a thorough adequacy assessment to ensure accuracy.

Figure 35: After a quantum device is certified for different graph states, one can ask which computations, i.e., unitaries, can be certifiably run on the device. Unitaries can be compiled onto a graph state for the equivalent MBQC pattern. However, one unitary can be mapped onto different graph states. Some of these might be certified on the device, others might not. Therefore, one can attempt to rewrite the computation to find a graph state that is certified on the device so it can implement the computation.

stages. Instead, we utilize uniquely trap computations as a generalized mechanism for characterizing quantum devices. This approach culminates in a concrete benchmarking protocol possessing all the attributes outlined in Section 6.2.2, filling some gaps in the existing benchmarking methods while removing the dependency on heuristics.

Practically speaking, our protocol is enunciated within the Measurement-Based Quantum Computing (MBQC) model, although translations to more commonly used circuit models can be achieved with minimal overhead [346, 347, 348]. This translation implies that our benchmark is hardware agnostic. Furthermore, the univerality of MBQC on 2D cluster states allows us to incorporate and enhance current benchmarking techniques by benchmarking exactly those graph states capable of implementing the computations which constitute the old benchmark.

Importantly, our benchmark focuses solely on certifying the dimensions of the employed graph states. This allows for an exhaustive benchmark of graph state dimensions, effectively scanning the entire quantum device. Armed with such comprehensive information, one can then synthesize circuits with the highest likelihood of successful execution on a given quantum device and obtain well-informed guidance on how the quantum solution at hand executes on a device in consideration [349].

Regarding device requirements, our protocol mandates the minimal condition of supporting the interactive processes described herein. Specifically, it can be implemented in the circuit model but necessitates mid-circuit measurements. Currently, these are not universally available from all current hardware providers. However, numerous organizations are making concerted efforts to integrate this feature, attracted by its significant benefits in improving fault tolerance and aiding in circuit synthesis [350, 351, 352, 353]. Apart from this, our proposal does not necessitate a precise match between the quantum device's structure and the benchmark's specifications. Operating on a larger quantum device while consistently implementing the same processes suffices. The critical factor is ensuring appropriate communication bandwidth. We regard the remainder of the device as a "black box," thereby alleviating the need for a predefined structural framework on the quantum device, even though a specific implementation might suggest otherwise.

One interesting future question is how to reduce the vulnerability of the proposed protocol to secret-dependent noise in the communication with the quantum device, which compromises the protocol's integrity and effectiveness. Addressing verification in the presence of secret-dependent noise remains an

open question and constitutes a future research avenue. Another perceived drawback is that our benchmark appears to predominantly reject current noisy hardware. This criticism, however, stems from the non-stringent requirements of existing benchmarks rather than a flaw in our approach. Achieving more rigorous and formal guarantees inevitably comes at the cost of higher stringency, a trade-off we consider justified. Furthermore, it is worth noting that the current state of hardware has advanced to the point where the first instances of verification protocols are now realizable [141, 354].

Chapter 7
Conclusions and outlook

The rapid advancement of quantum computing has brought both excitement and challenges, particularly in understanding the boundaries between classical and quantum computational power. Ideally, we would like to pinpoint the scale at which a quantum computer outperforms the best classical devices in specific computational processes where quantum systems have an asymptotic advantage. However, our understanding remains fundamentally limited: we have yet to fully identify which computational tasks demonstrate a meaningful quantum advantage, and at what point quantum devices become large and reliable enough to consistently realize this potential. Beyond identifying tasks where quantum superiority arises, there is the critical challenge of certifying the correctness of quantum device operations. Additionally, evaluating the broader computational capabilities of quantum devices is equally vital. Addressing these challenges—determining the nature of problems and quantum solutions that yield an advantage, while establishing reliable and formal certification processes to assess quantum performance—is essential to unlocking the full potential of quantum computing.

Compiling constant-depth quantum circuits with local, and multi-qubit gates. By focusing on the potential quantum advantages for near-term quantum devices operating at shallow depths, we have explored the potential of parallel quantum computation in comparison to parallel classical computation within this regime. In Chapter 3, we identified formal methods to synthesize parallel quantum circuits capable of solving a family of relation problems parameterized by all possible Boolean functions. Notably, with additional access to unbounded fan-in gates, such as parity gates, these circuits can evaluate all Boolean functions in constant depth [355]. While these constant-depth quantum circuits may not efficiently scale for most Boolean functions, they serve as essential components in more complex quantum circuits [77, 338]. Notably, integrating these circuits with additional quantum layers enhances their resource efficiency to evaluate arbitrary symmetric Boolean functions, where threshold functions serve as an example of practical interest. Thus, we expect that the circuits and Fourier analytic techniques discussed here may enable efficient Boolean logic operations in constant-depth quantum circuits, especially with the addition of multi-qubit gates like fanout gates. Recent results have further advanced this research [105], showing reductions in resource estimates for constant-depth quantum circuits that utilize multi-qubit controlled gates. Finally, beyond their computational advantages over parallel classical computation, this research about parallel quantum circuits also contributes to reducing resource costs in quantum algorithms with more significant advantages, such as integer factorization, which involves numerous arithmetic operations as subprocesses [356].
Additionally, in relation to the previous topic of using multi-qubit gates for parallel quantum computation, it has been exciting to see that, on the hardware front, significant progress has been made in realizing multi-qubit gates in constant-depth circuits as native operations in finite time steps, independent of the number of qubits affected, as demonstrated by recent experiments [79, 80, 81, 82, 326]. In this thesis, we contribute even more directly to this topic, in Chapter 5, by showing that the constant-depth circuit collapses as previously demonstrated for qubits also hold for qudits of prime dimensions [77]. Moreover,

we have shown that the parity gate is not unique to qubits in achieving such circuits, as equivalent circuits can be realized using unbounded fan-in gates with different arithmetic operations. This broadens the range of methods for implementing quantum circuits and achieves computational advantages previously identified in different hardware platforms. Along similar lines, recent research has explored the inclusion of other multi-qubit gates, such as the quantum threshold gate, which has been shown to generate the quantum fanout gate in a completely unitary manner [48]. This expansion of the gate set enhances potential quantum computational power in highly parallel settings, mainly to be equal to the one of the $\mathsf{QNC}^0[2]$ class. Thus, further investigating the computational capabilities of these circuit classes in solving algebraic problems within small classical complexity classes, such as NC^1 and NC^2, with polylogarithmic runtimes, presents a valuable area for future research. Such exploration could deepen our understanding of potential quantum advantages and facilitate the integration of known quantum solutions into larger computational problems.

Looking ahead, quantum circuit synthesis compilation remains a vibrant and promising field of research, with techniques focused on multi-qudit gates and parallel computation offering inspiring opportunities to pursue practical and early computational advantages, presenting various exciting research questions to explore.

Expanding the set of problems with quantum advantage. A different research direction involves identifying computational problems that exhibit quantum advantage using shallow-depth quantum circuits composed solely of bounded fan-in gates without additional multi-qubit gates. While multi-qubit gates hold promise, they require further research and technical development. Currently, we have access to universal quantum computation with bounded fan-in gates that achieve higher fidelities. Therefore, relying on a finite gate set is well-justified for practical, noise-resilient, and near-term quantum advantages.

In particular, from the relational problems of the $\mathcal{R}_{\langle f,p\rangle}$ type considered in Chapter 3, defined by a corresponding Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ and $p = 2$, a natural question arises: for which Boolean functions f, other than when f is the Least Significant Bit function, does the respective relation entail an unconditional quantum advantage, as the one shown in the seminal work by Bravyi et al.? This could lead to insights into other sets of relation problems and potentially reveal the underlying structure needed to achieve similar quantum advantages.

In this thesis, we did not fully explore this question. In Chapter 3, our primary focus was on optimizing quantum circuits and strategies for XOR non-local games, while in Chapter 4, we considered functions analogous to those used by Bravyi et al.[357] in a higher-dimensional setting, moving beyond the Boolean context. In Chapter 5, we addressed a different set of problems for which we demonstrated quantum advantage. However, this was not achieved using the simplest constant-depth quantum circuits—these circuits belong to $\mathsf{QAC}^0_{\mathsf{cf}}$, rather than QNC^0, as used in [357]. The missing element in potentially extending known separations and comparing our quantum solutions to the best classical ones lies in deriving upper bounds on classical strategies or, equivalently, lower bounds on the size of classical circuits capable of solving these problems. Initially, this was a possible direction we chose not to pursue, given the choice for the equally intriguing question analyzed in Chapter 4: to investigate the extent to which the existing quantum advantage can be scaled.

Nevertheless, expanding on the former question is certainly of great interest, as it could significantly broaden the set of relation problems for which we know unconditional quantum advantages beyond the handful of examples currently known [357, 35, 73, 143]. The main technical challenge lies in developing methods to upper-bound the optimal classical strategies for these non-local games. Current techniques are limited to linear Boolean functions, and extending them to non-linear Boolean functions in multiparty non-local games, such as XOR non-local games, will require new approaches to analyzing higher-degree functions. Moreover, overcoming this challenge will not only address the proposed question but also provide valuable insights of broader significance.

Finally, beyond the specific family of relation problems central to this thesis from Chapter 3 to Chapter 5, exploring other types of relation problems that may offer similar quantum advantages—and, importantly, hold greater practical applications—remains a key avenue for advancing our understanding of parallel quantum computations. Some progress has already been made in this area, particularly in settings that integrate multi-qubit quantum gates [78]. Overall, there is strong motivation to pursue quantum advantages that, while perhaps more modest in scale, still enable quantum circuits to outperform classical counterparts under comparable conditions.

Gauging/certifying the scale of quantum advantage. As previously mentioned in Chapter 4, we have explored the extent to which quantum advantage surpasses classical computation in more immediate settings. Building on the findings of [357] and the extension of this quantum advantage concerning AC^0 circuits [27], one might wonder whether proving a computational advantage beyond AC^0 is possible. It is conjectured that the parity halving problem may be sufficient for establishing a separation against $\mathsf{AC}^0[p]$ circuits for all prime p except 2. In scenarios where this holds, the problems proposed in Chapter 4 could confirm such a separation, including the case where $p = 2$. Unfortunately, existing proof techniques fall short in this context, requiring new tools to resolve this potential separation. This poses an intriguing question about the computational power of shallow-depth quantum computations, while the necessary techniques to lower bound $\mathsf{AC}^0[p]$ circuits regarding relation/search problems to settle this question would be of independent interest.

Further exploring this topic, recent research in the interactive setting [74] has shown that QNC^0 circuits are not contained within TC^0. Proving such a separation in the plain model remains challenging with our current proof techniques, particularly due to the absence of superpolynomial lower bounds for this circuit class. Nonetheless, we have demonstrated that the computational power of QNC^0 circuits surpasses that of AC^0. This was achieved by employing a class of biased polynomial threshold circuits of constant depth, which, put simply, use biased threshold gates. Interestingly, even with bias parameters that enable substantial computational power, this class cannot solve the parity halving problem. This limitation underscores the notable computational power of shallow-depth quantum circuits, which, for certain problems, significantly exceeds that of AC^0 circuits.

As these circuits serve as restricted models for neural networks while retaining significant computational power and potential practical efficacy, it is anticipated that these separations will facilitate exploration of potential quantum advantages in machine learning problems and the superiority of quantum models over classical ones. In particular, some studies have extended these separations to learning models and have demonstrated differences in standard machine learning metrics, such as the Kullback-Leibler divergence, between classical and quantum models [321]. This opens venues for further research in this domain. Concurrently, our findings suggest that even simple and potentially easy-to-learn quantum circuits can outperform their classical counterparts. This supports other research indicating that efficiently learnable quantum circuits could offer a quantum advantage in practical information processing tasks [317, 322, 323, 318, 321, 319, 324, 325, 358].

A common and significant objection regarding quantum systems is that they are inherently affected by noise, which complicates error correction significantly, casting doubt on the *robustness* of any computational advantage under such conditions. This is a crucial issue in both theoretical and practical contexts, especially in the NISQ era. Noise can severely diminish the computational advantages of even the most powerful quantum models over classical counterparts. For example, recent research demonstrates that even minor constant error rates can drastically reduce the capabilities of multi-prover interactive proofs with entangled provers (MIP^*) to those without entanglement (MIP) [264]. However, our research in Chapter 4 show that the separations we identified remain robust against noise: even when all stages of quantum computation are subjected to local stochastic noise, a specific family of modified relation problems can still be solved by these noisy quantum circuits provided with logical magic states, yet these

remain challenging for noiseless $\mathsf{bPTF}^0(k)$ circuits.
Our findings prompt further intriguing questions within this active research field. Having explored error-corrected qudit circuits for establishing quantum separations, we are now questioning whether advancements in high-dimensional error-correcting codes could enable earlier experimental demonstrations [296]. Moreover, we have reiterated the utility of magic states in constant-depth fault-tolerant regimes [65, 271]. While we conjecture that magic state factories may not be parallelizable to the extent of falling within the QNC^0 circuit class, determining the simplest circuit class capable of facilitating these processes and allowing for adaptive corrections during error-corrected circuit operations remains a key area of interest. This exploration could potentially extend known quantum advantages to simpler fault-tolerant executions [272]. Concurrently, examining more complex and alternative noise models could help to align research more closely with the challenges in practical quantum computing architectures [273].
Finally, we emphasize the importance of formally investigating the computational power of parallel quantum circuit classes that incorporate multi-qubit controlled gates. Our main result in the first part of Chapter 5, demonstrates that the addition of not only unbounded fan-in quantum fanout gates but also multi-qubit Toffoli gates and classical fanout gates suffices to show an unconditional quantum advantage over equivalent classical circuit classes and even larger circuit classes with additional unbounded fan-in modular gates, expressed as $\mathsf{QAC}^0_{\mathsf{cf}} \not\subseteq \mathsf{AC}^0[p]$ for any arbitrary prime p. This result further highlights the value of additional operations that quantum circuits provide for parallel computations, enabling more efficient calculations beyond what classical circuits can achieve.

Benchmarking and certifying quantum processes. In our exploration of shallow-depth quantum circuits, we have achieved significant quantum advantages, successfully overcoming numerous challenges. However, as we aim to scale these systems, the theoretical tools of error correction alone may not suffice to guarantee the correctness of computations; ensuring that error correction schemes are correctly implemented and functioning requires their own verification. This leads us to a critical issue: quantum verification. While quantum verification is theoretically well understood [359, 91] and initial experiments have been conducted [141], it primarily allows us to check individual circuit runs rather than offering a streamlined solution that instills high confidence in the capabilities of a specific quantum device across a broad range of quantum circuits.
This gap highlights the transition from quantum verification protocols to quantum benchmarking. Benchmarking should not only assesses and compare the performance of various quantum devices but also aid end-users to determine the suitability of a specific quantum device for their intended applications. To date, quantum benchmarking has been largely heuristic, with no universally accepted standard technique, though it primarily aims to predict the computational performance of quantum devices. Recent efforts [89] have been made to provide a more formal underpinning to these heuristic assessments, indicating a shift towards a more structured and predictable benchmarking framework.
In Chapter 6, we proposed new benchmarking paradigms that extend previously sought properties of benchmarking schemes while introducing the necessary formality for making predictive statements. Specifically, we demonstrated that quantum benchmarking schemes derived from this verification-inspired approach adhere to such more rigorous paradigms. Furthermore, our approach to quantum benchmarking diverges from earlier methods that primarily relied on heuristics or the verification of specific algorithmic instances. Instead, we leverage cryptographic tools for blindness and verification to perform comprehensive quantum device characterization, establishing a robust benchmarking framework that addresses existing methodological gaps. In addition, the resulting device characterizations facilitate a straightforward design of circuits tailored to individual quantum devices, offering detailed insights into their operational capabilities.
While this research has yielded positive outcomes, several concurrent and pertinent questions remain unaddressed, particularly concerning verifying and formally benchmarking other forms of quantum com-

putations. Notably, our current methodologies focus exclusively on BQP decision problems—a significant and large class. However, we lack formal methods to verify and, consequently, benchmark devices for sampling problems. These problems offer exponential quantum advantages under robust computational assumptions and serve as prime candidates for demonstrating quantum advantage on near-term quantum devices. Therefore, developing verification and benchmarking techniques for this class would be highly beneficial.

Interplay between quantum foundations and quantum computation. Finally, as a concluding consideration, all the results in this thesis were built upon the foundations of quantum mechanics and information theory. The first three chapters, from Chapter 3 to Chapter 5, focus on utilizing non-local games to design quantum circuits and demonstrate quantum advantage. In the later chapter, although indirectly, the discussion on quantum benchmarking is based on quantum verification protocols, with the latter incorporating stabilizer tests. These tests presuppose the existence of correct quantum states in specific superpositions, as seen in quantum communication protocols like BB84. Furthermore, recent developments have taken an inverted approach, where computational assumptions and cryptographic tools have facilitated and simplified the design of quantum experiments to test fundamental phenomena such as non-locality and contextuality [360, 361, 362, 363]. This approach helps to address challenges like maintaining space-like separations between experimenters or mitigating memory effects. Although it may not be surprising that a fundamental understanding of phenomena explored for information processing has significant implications for the respective computational machines, these examples—among many others—highlight the strong and productive interplay between quantum foundations and quantum computation, with valuable progress for theoretical computer science.

Bibliography

[1] Markus Frembs, Sam Roberts, Earl T Campbell, and Stephen D Bartlett. Hierarchies of resources for measurement-based quantum computation. *New Journal of Physics*, 25(1):013002, jan 2023.

[2] Lena Strobl, William Merrill, Gail Weiss, David Chiang, and Dana Angluin. What formal languages can transformers express? a survey. *Transactions of the Association for Computational Linguistics*, 12:543–561, 2024.

[3] William Merrill, Ashish Sabharwal, and Noah A. Smith. Saturated transformers are constant-depth threshold circuits. *Transactions of the Association for Computational Linguistics*, 10:843–856, 2022.

[4] Kai-Yeung Siu and Jehoshua Bruck. On the power of threshold circuits with small weights. *SIAM Journal on Discrete Mathematics*, 4(3):423–435, 1991.

[5] Alberto Bertoni and Beatrice Palano. Structural complexity and neural networks. In *Neural Nets: 13th Italian Workshop on Neural Nets, WIRN VIETRI 2002 Vietri sul Mare, Italy, May 30–June 1, 2002 Revised Papers 13*, pages 190–216. Springer, 2002.

[6] William Merrill and Ashish Sabharwal. The parallelism tradeoff: Limitations of log-precision transformers. *Transactions of the Association for Computational Linguistics*, 11:531–545, 2023.

[7] Jiří Šíma and Pekka Orponen. General-purpose computation with neural networks: A survey of complexity theoretic results. *Neural Computation*, 15(12):2727–2778, 2003.

[8] Ojas Parekh, Cynthia A Phillips, Conrad D James, and James B Aimone. Constant-depth and subcubic-size threshold circuits for matrix multiplication. In *Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures*, pages 67–76, 2018.

[9] Michael E Beverland, Prakash Murali, Matthias Troyer, Krysta M Svore, Torsten Hoefler, Vadym Kliuchnikov, Guang Hao Low, Mathias Soeken, Aarthi Sundaram, and Alexander Vaschillo. Assessing requirements to scale to practical quantum advantage. *arXiv:2211.07629*, 2022.

[10] Craig Gidney and Martin Ekerå. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, 2021.

[11] Artur Scherer, Benoıt Valiron, Siun-Chuon Mau, Scott Alexander, Eric Van den Berg, and Thomas E Chapuran. Concrete resource analysis of the quantum linear-system algorithm used to compute the electromagnetic scattering cross section of a 2d target. *Quantum Information Processing*, 16:1–65, 2017.

[12] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[13] Michael J Bremner, Richard Jozsa, and Dan J Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, 2011.

[14] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 333–342, 2011.

[15] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, 2019.

[16] Joonsuk Huh, Gian Giacomo Guerreschi, Borja Peropadre, Jarrod R McClean, and Alán Aspuru-Guzik. Boson sampling for molecular vibronic spectra. *Nature Photonics*, 9(9):615–620, 2015.

[17] Juan Miguel Arrazola and Thomas R Bromley. Using gaussian boson sampling to find dense subgraphs. *Physical review letters*, 121(3):030503, 2018.

[18] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[19] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 77–82, New York, NY, USA, 1987. Association for Computing Machinery.

[20] Andrew Chi-Chih Yao. Separating the polynomial-time hierarchy by oracles. In *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, pages 1–10. IEEE, 1985.

[21] David A Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in nc. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, pages 1–5, 1986.

[22] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical notes of the Academy of Sciences of the USSR*, 41:333–338, 1987.

[23] Pooya Hatami, William M. Hoza, Avishay Tal, and Roei Tell. Depth-□ threshold circuits vs. depth-(□+1) and-or trees. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 895–904, New York, NY, USA, 2023. Association for Computing Machinery.

[24] Yuval Filmus, Or Meir, and Avishay Tal. Shrinkage Under Random Projections, and Cubic Formula Lower Bounds for AC0. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 89:1–89:7, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[25] Ryan Williams. Nonuniform acc circuit lower bounds. *J. ACM*, 61(1), 2014.

[26] Emanuele Viola. Extractors for circuit sources. *SIAM Journal on Computing*, 43(2):655–672, 2014.

[27] Adam Bene Watts, Robin Kothari, Luke Schaeffer, and Avishay Tal. Exponential Separation between Shallow Quantum Circuits and Unbounded Fan-in Shallow Classical Circuits. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 515–526, New York, NY, USA, 2019. Association for Computing Machinery.

[28] Brett Giles and Peter Selinger. Exact synthesis of multiqubit clifford+t circuits. *Phys. Rev. A*, 87:032332, Mar 2013.

[29] Yan Ge, Wu Wenjie, Chen Yuheng, Pan Kaisen, Lu Xudong, Zhou Zixiang, Wang Yuhan, Wang Ruocheng, and Yan Junchi. Quantum circuit synthesis and compilation optimization: Overview and prospects. *arXiv preprint arXiv:2407.00736*, 2024.

[30] Florian Fürrutter, Gorka Muñoz-Gil, and Hans J Briegel. Quantum circuit synthesis with diffusion models. *Nature Machine Intelligence*, pages 1–10, 2024.

[31] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, page 212–219, New York, NY, USA, 1996. Association for Computing Machinery.

[32] Marc G Davis, Ethan Smith, Ana Tudor, Koushik Sen, Irfan Siddiqi, and Costin Iancu. Towards optimal topology aware quantum circuit synthesis. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 223–234. IEEE, 2020.

[33] François Le Gall. Average-case quantum advantage with shallow circuits. In *34th Computational Complexity Conference (CCC 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[34] Matthew Coudron, Jalex Stark, and Thomas Vidick. Trading locality for time: certifiable randomness from low-depth circuits. *Communications in mathematical physics*, 382:49–86, 2021.

[35] Sergey Bravyi, David Gosset, Robert Koenig, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020.

[36] Daniel Grier, Nathan Ju, and Luke Schaeffer. Interactive quantum advantage with noisy, shallow clifford circuits. *arXiv:2102.06833*, 2021.

[37] Libor Caha, Xavier Coiteux-Roy, and Robert Koenig. A colossal advantage: 3d-local noisy shallow quantum circuits defeat unbounded fan-in classical circuits. *arXiv:2312.09209*, 2023.

[38] Adam Bene Watts and Natalie Parham. Unconditional quantum advantage for sampling with shallow circuits. *arXiv:2301.00995*, 2023.

[39] Anne Broadbent, Joseph Fitzsimons, and Elham Kashefi. Universal blind quantum computation. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 517–526, 2009.

[40] Scott Aaronson, Alexandru Cojocaru, Alexandru Gheorghiu, and Elham Kashefi. Complexity-Theoretic Limitations on Blind Delegated Quantum Computation. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1—6:13, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[41] Maxime Garnier, Dominik Leichtle, Luka Music, and Harold Ollivier. Composably secure delegated quantum computation with weak coherent pulses. In *2024 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, pages 221–225, 2024.

[42] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[43] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, USA, 1st edition, 2009.

[44] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, 2014.

[45] Heribert Vollmer. *Introduction to Circuit Complexity: A Uniform Approach*. Springer Publishing Company, Incorporated, 1st edition, 2010.

[46] Cristopher Moore and Martin Nilsson. Parallel quantum computation and quantum codes. *SIAM journal on computing*, 31(3):799–815, 2001.

[47] Cristopher Moore. Quantum circuits: Fanout, parity, and counting, 1999.

[48] Daniel Grier and Jackson Morris. Quantum Threshold is Powerful. *arXiv e-prints*, page arXiv:2411.04953, November 2024.

[49] Amir Shpilka, Amir Yehudayoff, et al. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010.

[50] Scott Aaronson. BQP and the Polynomial Hierarchy. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, pages 141–150, New York, NY, USA, 2010. Association for Computing Machinery.

[51] Atul Singh Arora, Andrea Coladangelo, Matthew Coudron, Alexandru Gheorghiu, Uttam Singh, and Hendrik Waldner. Quantum depth in the random oracle model. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1111–1124, New York, NY, USA, 2023. Association for Computing Machinery.

[52] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. On the Pauli Spectrum of QAC0. *arXiv e-prints*, page arXiv:2311.09631, November 2023.

[53] Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005.

[54] Richard Jozsa. An introduction to measurement based quantum computation. In *Quantum Information Processing*, pages 137–158. IOS Press, 2006.

[55] Kun Fang, Munan Zhang, Ruqi Shi, and Yinan Li. Dynamic quantum circuit compilation. *arXiv preprint arXiv:2310.11021*, 2023.

[56] Matthew Amy, Dmitri Maslov, and Michele Mosca. Polynomial-time t-depth optimization of clifford+ t circuits via matroid partitioning. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 33(10):1476–1489, 2014.

[57] Bob Coecke and Ross Duncan. Interacting quantum observables. In *International Colloquium on Automata, Languages, and Programming*, pages 298–310. Springer, 2008.

[58] Jacob Biamonte and Ville Bergholm. Tensor networks in a nutshell. *arXiv preprint arXiv:1708.00006*, 2017.

[59] Jiaqing Jiang, Xiaoming Sun, Shang-Hua Teng, Bujiao Wu, Kewen Wu, and Jialin Zhang. Optimal space-depth trade-off of cnot circuits in quantum logic synthesis. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 213–229. SIAM, 2020.

[60] Tom Peham, Nina Brandl, Richard Kueng, Robert Wille, and Lukas Burgholzer. Depth-optimal synthesis of clifford circuits with sat solvers. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 1, pages 802–813. IEEE, 2023.

[61] Lindsay Bassman Oftelie, Roel Van Beeumen, Ed Younis, Ethan Smith, Costin Iancu, and Wibe A de Jong. Constant-depth circuits for dynamic simulations of materials on quantum computers. *Materials Theory*, 6(1):13, 2022.

[62] Christopher S Wang, Jacob C Curtis, Brian J Lester, Yaxing Zhang, Yvonne Y Gao, Jessica Freeze, Victor S Batista, Patrick H Vaccaro, Isaac L Chuang, Luigi Frunzio, et al. Efficient multiphoton sampling of molecular vibronic spectra on a superconducting bosonic processor. *Physical Review X*, 10(2):021060, 2020.

[63] Dominik Hangleiter, Marcin Kalinowski, Dolev Bluvstein, Madelyn Cain, Nishad Maskara, Xun Gao, Aleksander Kubica, Mikhail D Lukin, and Michael J Gullans. Fault-tolerant compiling of classically hard iqp circuits on hypercubes. *arXiv preprint arXiv:2404.19005*, 2024.

[64] Carsten Robens, Iñigo Arrazola, Wolfgang Alt, Dieter Meschede, Lucas Lamata, Enrique Solano, and Andrea Alberti. Boson sampling with ultracold atoms in a programmable optical lattice. *Physical Review A*, 110(1):012615, 2024.

[65] Rawad Mezher, Joe Ghalbouni, Joseph Dgheim, and Damian Markham. Fault-tolerant quantum speedup from constant depth quantum circuits. *Phys. Rev. Res.*, 2:033444, 2020.

[66] Tom Douce, Damian Markham, Elham Kashefi, Eleni Diamanti, Thomas Coudreau, Perola Milman, Peter Van Loock, and Giulia Ferrini. Continuous-variable instantaneous quantum computing is hard to sample. *Physical review letters*, 118(7):070503, 2017.

[67] Matthew DeCross, Reza Haghshenas, Minzhao Liu, Enrico Rinaldi, Johnnie Gray, Yuri Alexeev, Charles H Baldwin, John P Bartolotta, Matthew Bohn, Eli Chertkov, et al. The computational power of random quantum circuits in arbitrary geometries. *arXiv preprint arXiv:2406.02501*, 2024.

[68] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G S L Brandao, David A Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P Harrigan, Michael J Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S Humble, Sergei V Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C Platt, Chris Quintana, Eleanor G Rieffel, Pedram Roushan, Nicholas C Rubin, Daniel Sank, Kevin J Satzinger, Vadim Smelyanskiy, Kevin J Sung, Matthew D Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[69] Jonathan Barrett, Carlton M. Caves, Bryan Eastin, Matthew B. Elliott, and Stefano Pironio. Modeling pauli measurements on graph states with nearest-neighbor classical communication. *Physical Review A*, 75(1), January 2007.

[70] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*, pages 11–20, 1993.

[71] Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3):107–108, 1990.

[72] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.*, 65:3373–3376, 1990.

[73] Libor Caha, Xavier Coiteux-Roy, and Robert Koenig. Single-qubit gate teleportation provides a quantum advantage. *arXiv e-prints*, page arXiv:2209.14158, sep 2022.

[74] Daniel Grier and Luke Schaeffer. Interactive Shallow Clifford Circuits: Quantum Advantage against NC^1 and Beyond. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 875–888, New York, NY, USA, 2020. Association for Computing Machinery.

[75] Kishor Bharti and Rahul Jain. On the power of geometrically-local classical and quantum circuits. *arXiv:2310.01540*, 2023.

[76] Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. Counting, Fanout and the Complexity of Quantum ACC. *Quantum Info. Comput.*, 2(1):35–65, dec 2002.

[77] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. *computational complexity*, 25:849–881, 2016.

[78] Jop Briët, Harry Buhrman, Davi Castro-Silva, and Niels M. P. Neumann. Noisy Decoding by Shallow Circuits with Parities: Classical and Quantum (Extended Abstract). In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:11, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[79] Dmitri Maslov and Yunseong Nam. Use of global interactions in efficient quantum circuit constructions. *New Journal of Physics*, 20(3):033018, 2018.

[80] Koen Groenland, Freek Witteveen, Kareljan Schoutens, and Rene Gerritsma. Signal processing techniques for efficient compilation of controlled rotations in trapped ions. *New Journal of Physics*, 22(6):063006, 2020.

[81] Nikodem Grzesiak, Andrii Maksymov, Pradeep Niroula, and Yunseong Nam. Efficient quantum programming using ease gates on a trapped-ion quantum computer. *Quantum*, 6:634, 2022.

[82] Sergey Bravyi, Dmitri Maslov, and Yunseong Nam. Constant-cost implementations of clifford operations and multiply-controlled gates using global interactions. *Physical Review Letters*, 129(23):230501, 2022.

[83] Yongxin Song, Liberto Beltrán, Ilya Besedin, Michael Kerschbaum, Marek Pechal, François Swiadek, Christoph Hellings, Dante Colao Zanuz, Alexander Flasby, Jean-Claude Besse, and Andreas Wallraff. Realization of Constant-Depth Fan-Out with Real-Time Feedforward on a Superconducting Quantum Processor. *arXiv e-prints*, page arXiv:2409.06989, September 2024.

[84] Gary J Mooney, Gregory AL White, Charles D Hill, and Lloyd CL Hollenberg. Generation and verification of 27-qubit greenberger-horne-zeilinger states in a superconducting quantum computer. *Journal of Physics Communications*, 5(9):095004, 2021.

[85] Andrew W Cross, Lev S Bishop, Sarah Sheldon, Paul D Nation, and Jay M Gambetta. Validating quantum computers using randomized model circuits. *Physical Review A*, 100(3):032328, 2019.

[86] Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. Measuring the capabilities of quantum computers. *Nature Physics*, 18(1):75–79, 2022.

[87] David C. McKay, Ian Hincks, Emily J. Pritchett, Malcolm Carroll, Luke C. G. Govia, and Seth T. Merkel. Benchmarking Quantum Processor Performance at Scale. *arXiv e-prints*, page arXiv:2311.05933, November 2023.

[88] Thomas Lubinski, Sonika Johri, Paul Varosy, Jeremiah Coleman, Luning Zhao, Jason Necaise, Charles H. Baldwin, Karl Mayer, and Timothy Proctor. Application-oriented performance benchmarks for quantum computing, 2023.

[89] Daniel Hothem, Jordan Hines, Karthik Nataraj, Robin Blume-Kohout, and Timothy Proctor. Predictive models from quantum computer benchmarks. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 01, pages 709–714, 2023.

[90] Theodoros Kapourniotis, Elham Kashefi, Dominik Leichtle, Luka Music, and Harold Ollivier. Unifying quantum verification and error-detection: theory and tools for optimisations. *Quantum Science and Technology*, 9(3):035036, may 2024.

[91] Dominik Leichtle, Luka Music, Elham Kashefi, and Harold Ollivier. Verifying bqp computations on noisy devices with minimal overhead. *PRX Quantum*, 2:040302, Oct 2021.

[92] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, 2020.

[93] Andrew Wack, Hanhee Paik, Ali Javadi-Abhari, Petar Jurcevic, Ismael Faro, Jay M Gambetta, and Blake R Johnson. Quality, speed, and scale: three key attributes to measure the performance of near-term quantum computers. *arXiv preprint arXiv:2110.14108*, 2021.

[94] Mirko Amico, Helena Zhang, Petar Jurcevic, Lev S. Bishop, Paul Nation, Andrew Wack, and David C. McKay. Defining Standard Strategies for Quantum Benchmarks. *arXiv e-prints*, page arXiv:2303.02108, March 2023.

[95] Michael de Oliveira, Luís S. Barbosa, and Ernesto F. Galvão. Quantum advantage in temporally flat measurement-based quantum computation. *Quantum*, 8:1312, April 2024.

[96] Alex Bredariol Grilo, Elham Kashefi, Damian Markham, and Michael de Oliveira. The power of shallow-depth Toffoli and qudit quantum circuits. *arXiv e-prints*, page arXiv:2404.18104, April 2024.

[97] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the Hierarchy of Constant-Depth Exact Quantum Circuits. In *2013 IEEE Conference on Computational Complexity*, pages 168–178, 2013.

[98] Anne Broadbent and Elham Kashefi. Parallelizing quantum circuits. *Theoretical Computer Science*, 410(26):2489–2510, 2009.

[99] Dan Browne, Elham Kashefi, and Simon Perdrix. Computational depth complexity of measurement-based quantum computation. In Wim van Dam, Vivien M. Kendon, and Simone Severini, editors, *Theory of Quantum Computation, Communication, and Cryptography*, pages 35–46, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[100] Oded Regev and Thomas Vidick. Quantum xor games. *ACM Trans. Comput. Theory*, 7(4), August 2015.

[101] Lorenzo Catani, Ricardo Faleiro, Pierre-Emmanuel Emeriau, Shane Mansfield, and Anna Pappa. Connecting xor and xor* games. *Phys. Rev. A*, 109:012427, Jan 2024.

[102] Matty J Hoban, Earl T Campbell, Klearchos Loukopoulos, and Dan E Browne. Non-adaptive measurement-based quantum computation and multi-party Bell inequalities. *New Journal of Physics*, 13(2):23014, feb 2011.

[103] Austin K Daniel and Akimasa Miyake. Quantum algorithms for classical boolean functions via adaptive measurements: Exponential reductions in space-time resources. *arXiv e-prints*, page arXiv:2211.01252, 2022.

[104] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. *computational complexity*, 25:849–881, 2016.

[105] Jonathan Allcock, Jinge Bao, João F. Doriguello, Alessandro Luongo, and Miklos Santha. Constant-depth circuits for Uniformly Controlled Gates and Boolean functions with application to quantum memory circuits. *arXiv e-prints*, page arXiv:2308.08539, August 2023.

[106] Anurag Anshu, Yangjing Dong, Fengning Ou, and Penghui Yao. On the Computational Power of QAC0 with Barely Superlinear Ancillae. *arXiv e-prints*, page arXiv:2410.06499, October 2024.

[107] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, Jul 1994.

[108] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52:3457–3467, Nov 1995.

[109] Yulin Chi, Jieshan Huang, Zhanchuan Zhang, Jun Mao, Zinan Zhou, Xiaojiong Chen, Chonghao Zhai, Jueming Bao, Tianxiang Dai, Huihong Yuan, et al. A programmable qudit-based quantum processor. *Nature communications*, 13(1):1166, 2022.

[110] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz. A universal qudit quantum processor with trapped ions. *Nature Physics*, 18(9):1053–1057, 2022.

[111] Christian Reimer, Stefania Sciara, Piotr Roztocki, Mehedi Islam, Luis Romero Cortés, Yanbing Zhang, Bennet Fischer, Sébastien Loranger, Raman Kashyap, Alfonso Cino, et al. High-dimensional one-way quantum processing implemented on d-level cluster states. *Nature Physics*, 15(2):148–153, 2019.

[112] Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. *Theory of Computing*, 10(1):133–166, 2014.

[113] Scott Aaronson. How Much Structure Is Needed for Huge Quantum Speedups? *arXiv preprint arxiv:2209.06930*, sep 2022.

[114] Takashi Yamakawa and Mark Zhandry. Verifiable quantum advantage without structure. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 69–74. IEEE, 2022.

[115] Vinayak M. Kumar. Tight correlation bounds for circuits between ac0 and tc0. In *Proceedings of the Conference on Proceedings of the 38th Computational Complexity Conference*, CCC '23, Dagstuhl, DEU, 2023. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[116] Daniel González-Cuadra, Torsten V. Zache, Jose Carrasco, Barbara Kraus, and Peter Zoller. Hardware efficient quantum simulation of non-abelian gauge theories with qudits on rydberg platforms. *Phys. Rev. Lett.*, 129:160501, 2022.

[117] Sivert Aasnæss. *Comparing two cohomological obstructions for contextuality, and a generalised construction of quantum advantage with shallow circuits*. PhD thesis, University of Oxford, Oxford, UK, 2021.

[118] Jay Lawrence. Mermin inequalities for perfect correlations in many-qutrit systems. *Physical Review A*, 95(4):042123, 2017.

[119] Yanglin Hu, Darya Melnyk, Yuyi Wang, and Roger Wattenhofer. Space complexity of streaming algorithms on universal quantum computers. In *Theory and Applications of Models of Computation: 16th International Conference, TAMC 2020, Changsha, China, October 18–20, 2020, Proceedings 16*, pages 275–286. Springer, 2020.

[120] Héctor Bombín. Resilience to time-correlated noise in quantum computation. *Physical Review X*, 6(4):041034, 2016.

[121] David Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of mathematical physics*, 47(12), 2006.

[122] Mark Howard, Eoin Brennan, and Jiri Vala. Quantum contextuality with stabilizer states. *Entropy*, 15(6):2340–2362, 2013.

[123] Uta Isabella Meyer, Ivan Šupić, Damian Markham, and Frédéric Grosshans. Bell Nonlocality from Wigner Negativity in Qudit Systems. *arXiv e-prints*, page arXiv:2405.14367, May 2024.

[124] Robert Raussendorf, Sergey Bravyi, and Jim Harrington. Long-range quantum entanglement in noisy cluster states. *Phys. Rev. A*, 71:062313, Jun 2005.

[125] Clémence Chevignard, Pierre-Alain Fouque, and André Schrottenloher. Reducing the number of qubits in quantum factoring. Cryptology ePrint Archive, Paper 2024/222, 2024.

[126] Thomas Lubinski, Sonika Johri, Paul Varosy, Jeremiah Coleman, Luning Zhao, Jason Necaise, Charles H Baldwin, Karl Mayer, and Timothy Proctor. Application-oriented performance benchmarks for quantum computing. *IEEE Transactions on Quantum Engineering*, 2023.

[127] Dolev Bluvstein, Simon J Evered, Alexandra A Geim, Sophie H Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, et al. Logical quantum processor based on reconfigurable atom arrays. *Nature*, pages 1–3, 2023.

[128] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla Lambrocco, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, 2015.

[129] Dominik Rauch, Johannes Handsteiner, Armin Hochrainer, Jason Gallicchio, Andrew S. Friedman, Calvin Leung, Bo Liu, Lukas Bulla, Sebastian Ecker, Fabian Steinlechner, Rupert Ursin, Beili Hu, David Leon, Chris Benn, Adriano Ghedina, Massimo Cecconi, Alan H. Guth, David I. Kaiser, Thomas Scheidl, and Anton Zeilinger. Cosmic bell test using random measurement settings from high-redshift quasars. *Phys. Rev. Lett.*, 121:080403, 2018.

[130] Anna Gál, Kristoffer Arnsfelt Hansen, Michal Koucký, Pavel Pudlák, and Emanuele Viola. Tight bounds on computing error-correcting codes by bounded-depth circuits with arbitrary gates. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 479–494, 2012.

[131] Yang Wang, Selwyn Simsek, Thomas M Gatterman, Justin A Gerber, Kevin Gilmore, Dan Gresh, Nathan Hewitt, Chandler V Horst, Mitchell Matheny, Tanner Mengle, et al. Fault-tolerant one-bit addition with the smallest interesting color code. *Science Advances*, 10(29):eado9024, 2024.

[132] Rajeev Acharya, Laleh Aghababaie-Beni, Igor Aleiner, Trond I Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Nikita Astrakhantsev, Juan Atalaya, et al. Quantum error correction below the surface code threshold. *arXiv preprint arXiv:2408.13687*, 2024.

[133] Bianca Schroeder, Eduardo Pinheiro, and Wolf-Dietrich Weber. Dram errors in the wild: a large-scale field study. In *Proceedings of the Eleventh International Joint Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '09, page 193–204, New York, NY, USA, 2009. Association for Computing Machinery.

[134] Oded Goldreich. Three xor-lemmas—an exposition. *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation: In Collaboration with Lidor Avigad, Mihir Bellare, Zvika Brakerski, Shafi Goldwasser, Shai Halevi, Tali Kaufman, Leonid Levin, Noam Nisan, Dana Ron, Madhu Sudan, Luca Trevisan, Salil Vadhan, Avi Wigderson, David Zuckerman*, pages 248–272, 2011.

[135] Gregory Rosenthal. Bounds on the $qac0$ complexity of approximating parity. *arXiv:2008.07470*, 2020.

[136] Frederic Green, Steven Homer, and Chris Pollett. On the complexity of quantum acc. In *Proceedings 15th Annual IEEE Conference on Computational Complexity*, pages 250–262. IEEE, 2000.

[137] Daniel Padé, Stephen Fenner, Daniel Grier, and Thomas Thierauf. Depth-2 qac circuits cannot simulate quantum parity. *arXiv:2005.12169*, 2020.

[138] M. Fang, S. Fenner, F. Green, S. Homer, and Y. Zhang. Quantum lower bounds for fanout. *Quantum Info. Comput.*, 6(1):46–57, jan 2006.

[139] Debajyoti Bera. A lower bound method for quantum circuits. *Information processing letters*, 111(15):723–726, 2011.

[140] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, jun 2020.

[141] P Drmota, DP Nadlinger, D Main, BC Nichol, EM Ainley, Dominik Leichtle, A Mantri, Elham Kashefi, R Srinivas, G Araneda, et al. Verifiable blind quantum computing with trapped ions and single photons. *arXiv preprint arXiv:2305.02936*, 2023.

[142] Cica Gustiani, Dominik Leichtle, Daniel Mills, Jonathan Miller, Ross Grassie, and Elham Kashefi. On-chip verified quantum computation with an ion-trap quantum processing unit. *arXiv preprint arXiv:2410.24133*, 2024.

[143] Min-Hsiu Hsieh, Leandro Mendes, Michael de Oliveira, and Sathyawageeswar Subramanian. Unconditionally separating noisy $qnc0$ from bounded polynomial threshold circuits of constant depth. *arXiv preprint arXiv:2408.16378*, 2024.

[144] Johannes Frank, Elham Kashefi, Dominik Leichtle, and Michael de Oliveira. Heuristic-free Verification-inspired Quantum Benchmarking. *arXiv e-prints*, page arXiv:2404.10739, April 2024.

[145] Scott Aaronson, DeVon Ingram, and William Kretschmer. The Acrobatics of BQP. In Shachar Lovett, editor, *37th Computational Complexity Conference (CCC 2022)*, volume 234 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20:1–20:17, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[146] Richard Jozsa and Noah Linden. On the role of entanglement in quantum-computational speed-up. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 459(2036):2011–2032, 2003.

[147] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, 2014.

[148] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. Contextuality as a resource for models of quantum computation with qubits. *Phys. Rev. Lett.*, 119:120505, Sep 2017.

[149] Ernesto F. Galvão. Discrete wigner functions and quantum computational speedup. *Phys. Rev. A*, 71:042302, Apr 2005.

[150] A. Mari and J. Eisert. Positive wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.*, 109:230503, Dec 2012.

[151] Lov K. Grover. The advantages of superposition. *Science*, 280(5361):228–228, 1998.

[152] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[153] Maarten Van den Nest, Akimasa Miyake, Wolfgang Dür, and Hans J. Briegel. Universal resources for measurement-based quantum computation. *Phys. Rev. Lett.*, 97:150504, Oct 2006.

[154] Janet Anders and Dan E. Browne. Computational power of correlations. *Phys. Rev. Lett.*, 102:050502, Feb 2009.

[155] Vincent Danos and Elham Kashefi. Determinism in the one-way model. *Phys. Rev. A*, 74(5):052310, nov 2006.

[156] Daniel E Browne, Elham Kashefi, Mehdi Mhalla, and Simon Perdrix. Generalized flow and determinism in measurement-based quantum computation. *New Journal of Physics*, 9(8):250, aug 2007.

[157] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations. *Phys. Rev. Lett.*, 117(8):080501, aug 2016.

[158] Matty J. Hoban, Joel J. Wallman, Hussain Anwar, Naïri Usher, Robert Raussendorf, and Dan E. Browne. Measurement-based classical computation. *Phys. Rev. Lett.*, 112:140505, Apr 2014.

[159] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, April 2017.

[160] Leonardo Novo, Juani Bermejo-Vega, and Raúl García-Patrón. Quantum advantage from energy measurements of many-body quantum systems. *Quantum*, 5:465, jun 2021.

[161] Masahito Hayashi and Yuki Takeuchi. Verifying commuting quantum computations via fidelity estimation of weighted graph states. *New Journal of Physics*, 21(9):93060, sep 2019.

[162] Juan Bermejo-Vega, Dominik Hangleiter, Martin Schwarz, Robert Raussendorf, and Jens Eisert. Architectures for Quantum Simulation Showing a Quantum Speedup. *Phys. Rev. X*, 8(2):021010, apr 2018.

[163] Jacob Miller, Stephen Sanders, and Akimasa Miyake. Quantum supremacy in constant-time measurement-based computation: A unified architecture for sampling and verification. *Phys. Rev. A*, 96:062320, Dec 2017.

[164] Ryuhei Mori. Periodic Fourier representation of Boolean functions. *Quantum Info. Comput.*, 19(5–6):392–412, may 2019.

[165] Jelena Mackeprang, Daniel Bhatti, Matty J Hoban, and Stefanie Barz. The power of qutrits for non-adaptive measurement-based quantum computing. *New Journal of Physics*, 25(7):073007, jul 2023.

[166] Daniel Collins, Nicolas Gisin, Sandu Popescu, David Roberts, and Valerio Scarani. Bell-Type Inequalities to Detect True *n*-Body Nonseparability. *Phys. Rev. Lett.*, 88(17):170405, apr 2002.

[167] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86(2):419–478, apr 2014.

[168] Dmitrijs Kravčenko. *Quantum Games, Quantum States, Their Properties and Applications*. PhD thesis, Latvijas Universitāte, 2013.

[169] William Slofstra. Lower bounds on the entanglement needed to play XOR non-local games. *Journal of Mathematical Physics*, 52(10):102202, 2011.

[170] Andris Ambainis, Jānis Iraids, Dmitry Kravchenko, and Madars Virza. Advantage of quantum strategies in random symmetric xor games. In Antonín Kučera, Thomas A. Henzinger, Jaroslav Nešetřil, Tomáš Vojnar, and David Antoš, editors, *Mathematical and Engineering Methods in Computer Science*, pages 57–68, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[171] Andris Ambainis and Janis Iraids. Provable Advantage for Quantum Strategies in Random Symmetric XOR Games. In Simone Severini and Fernando Brandao, editors, *8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)*, volume 22 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 146–156, Dagstuhl, Germany, 2013. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[172] Samuel Marcovitch and Benni Reznik. Implications of communication complexity in multipartite systems. *Phys. Rev. A*, 77:032120, Mar 2008.

[173] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. Information causality as a physical principle. *Nature*, 461(7267):1101–1104, 2009.

[174] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[175] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71(2):022101, feb 2005.

[176] A A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, feb 2003.

[177] Zhiqiang Zhang and Yaoyun Shi. Communication complexities of symmetric XOR functions. *Quantum Information and Computation*, 9(3):255–263, 2009.

[178] Pierre Botteron. NonLocal Boxes and Communication Complexity. Master's thesis, Université Paul Sabatier Toulouse III, 2022.

[179] Kwangil Bae and Wonmin Son. Generalized nonlocality criteria under the correlation symmetry. *Phys. Rev. A*, 98:022116, Aug 2018.

[180] Markus Frembs, Sam Roberts, and Stephen D Bartlett. Contextuality as a resource for measurement-based quantum computation beyond qubits. *New Journal of Physics*, 20(10):103011, oct 2018.

[181] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[182] François Le Gall. Average-Case Quantum Advantage with Shallow Circuits. In Amir Shpilka, editor, *34th Computational Complexity Conference (CCC 2019)*, volume 137 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1—21:20, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[183] Matthew Coudron, Jalex Stark, and Thomas Vidick. Trading locality for time: certifiable randomness from low-depth circuits. *Communications in mathematical physics*, 382(1):49–86, 2021.

[184] Sergey Bravyi, David Gosset, Robert König, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020.

[185] Atsuya Hasegawa and François Le Gall. Quantum Advantage with Shallow Circuits Under Arbitrary Corruption. In Hee-Kap Ahn and Kunihiko Sadakane, editors, *32nd International Symposium on Algorithms and Computation (ISAAC 2021)*, volume 212 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 74:1–74:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[186] Natalie Parham. On the Power and Limitations of Shallow Quantum Circuits. Master's thesis, University of Waterloo, 2022.

[187] Dmitri Maslov, Jin-Sung Kim, Sergey Bravyi, Theodore J Yoder, and Sarah Sheldon. Quantum advantage for computations with limited space. *Nature Physics*, 17(8):894–897, 2021.

[188] Farid Ablayev, Aida Gainutdinova, Marek Karpinski, Cristopher Moore, and Christopher Pollett. On the computational power of probabilistic and quantum branching program. *Information and Computation*, 203(2):145–162, 2005.

[189] D Shepherd and M. J. Bremner. Temporally unstructured quantum computation. *Proceedings of the Royal Society of London Series A*, 465(2105):1413–1439, jan 2009.

[190] Daniel M Greenberger, Michael A Horne, and Anton Zeilinger. Going Beyond Bell's Theorem. In Menas Kafatos, editor, *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72, Dordrecht, 1989. Springer Netherlands.

[191] Diogo Cruz, Romain Fournier, Fabien Gremion, Alix Jeannerot, Kenichi Komagata, Tara Tosic, Jarla Thiesbrummel, Chun Lam Chan, Nicolas Macris, Marc-André Dupertuis, and Clément Javerzac-Galy. Efficient Quantum Algorithms for GHZ and W States, and Implementation on the IBM Quantum Computer. *Advanced Quantum Technologies*, 2(5-6):1900015, 2019.

[192] R. F. Werner and M. M. Wolf. All-multipartite bell-correlation inequalities for two dichotomic observables per site. *Phys. Rev. A*, 64:032112, Aug 2001.

[193] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[194] Anastasiya Chistopolskaya and Vladimir V. Podolskii. Parity Decision Tree Complexity is Greater Than Granularity. *arXiv e-prints*, page arXiv:1810.08668, October 2018.

[195] A Canteaut and M Videau. Symmetric Boolean functions. *IEEE Transactions on Information Theory*, 51(8):2791–2811, 2005.

[196] Larry J Stockmeyer. On the combinational complexity of certain symmetric Boolean functions. *Mathematical systems theory*, 10(1):323–336, 1976.

[197] R F Arnold and M A Harrison. Algebraic Properties of Symmetric and Partially Symmetric Boolean Functions. *IEEE Transactions on Electronic Computers*, EC-12(3):244–251, 1963.

[198] An Braeken and Bart Preneel. On the algebraic immunity of symmetric boolean functions. In Subhamoy Maitra, C. E. Veni Madhavan, and Ramarathnam Venkatesan, editors, *Progress in Cryptology - INDOCRYPT 2005*, volume 3797 of *Lecture Notes in Computer Science*, pages 35–48, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[199] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002.

[200] Matthew Amy, Dmitri Maslov, Michele Mosca, and Martin Roetteler. A Meet-in-the-Middle Algorithm for Fast Synthesis of Depth-Optimal Quantum Circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 32(6):818–830, 2013.

[201] V V Shende, S S Bullock, and I L Markov. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(6):1000–1010, 2006.

[202] Juha J Vartiainen, Mikko Möttönen, and Martti M Salomaa. Efficient Decomposition of Quantum Gates. *Phys. Rev. Lett.*, 92(17):177902, apr 2004.

[203] Bei Zeng, Xie Chen, and Isaac L. Chuang. Semi-clifford operations, structure of $\rfloor_k$ hierarchy, and gate complexity for fault-tolerant quantum computation. *Phys. Rev. A*, 77:042313, April 2008.

[204] Gary J Mooney, Charles D Hill, and Lloyd C L Hollenberg. Cost-optimal single-qubit gate synthesis in the Clifford hierarchy. *Quantum*, 5:396, feb 2021.

[205] Nadish de Silva. Efficient quantum gate teleportation in higher dimensions. *Proceedings of the Royal Society A*, 477(2251):20200865, 2021.

[206] Daniel Gottesman and Isaac L Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, 1999.

[207] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *arXiv e-prints*, pages quant–ph/9807006, July 1998.

[208] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Fast and efficient exact synthesis of single-qubit unitaries generated by clifford and t gates. *Quantum Info. Comput.*, 13(7–8):607–630, jul 2013.

[209] Nicolas Brunner, James Sharam, and Tamás Vértesi. Testing the structure of multipartite entanglement with bell inequalities. *Phys. Rev. Lett.*, 108:110501, Mar 2012.

[210] Julia Kempe, Hirotada Kobayashi, Keiji Matsumoto, Ben Toner, and Thomas Vidick. Entangled Games Are Hard to Approximate. *SIAM Journal on Computing*, 40(3):848–877, 2011.

[211] Ravindran Kannan and Achim Bachem. Polynomial Algorithms for Computing the Smith and Hermite Normal Forms of an Integer Matrix. *SIAM Journal on Computing*, 8(4):499–507, 1979.

[212] Josh Alman and Virginia Vassilevska Williams. A refined laser method and faster matrix multiplication. In *Proceedings of the Thirty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '21, page 522–539, USA, 2021. Society for Industrial and Applied Mathematics.

[213] Yihui Quek, Eneet Kaur, and Mark M. Wilde. Multivariate trace estimation in constant quantum depth. *Quantum*, 8:1220, January 2024.

[214] Peter Selinger. Efficient Clifford+T Approximation of Single-Qubit Operators. *Quantum Info. Comput.*, 15(1–2):159–180, jan 2015.

[215] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Practical Approximation of Single-Qubit Unitaries by Single-Qubit Quantum Clifford and T Circuits. *IEEE Transactions on Computers*, 65(1):161–172, 2016.

[216] Neil J Ross. Optimal Ancilla-Free CLIFFORD+V Approximation of Z-Rotations. *Quantum Info. Comput.*, 15(11–12):932–950, sep 2015.

[217] Ethan Bernstein and Umesh Vazirani. Quantum Complexity Theory. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '93, pages 11–20, New York, NY, USA, 1993. Association for Computing Machinery.

[218] Alex Bocharov, Martin Roetteler, and Krysta M Svore. Efficient synthesis of probabilistic quantum circuits with fallback. *Phys. Rev. A*, 91(5):052317, may 2015.

[219] Alex Bocharov, Martin Roetteler, and Krysta M Svore. Efficient Synthesis of Universal Repeat-Until-Success Quantum Circuits. *Phys. Rev. Lett.*, 114(8):080502, feb 2015.

[220] Ingo Wegener. *The Complexity of Boolean Functions*. John Wiley & Sons, Inc., USA, 1987.

[221] Jaikumar Radhakrishnan. Better bounds for threshold formulas. In *[1991] Proceedings 32nd Annual Symposium of Foundations of Computer Science*, pages 314–323. IEEE Computer Society, 1991.

[222] Michael J Fischer, Albert R Meyer, and Michael S Paterson. $\Omega(N \log n)$ Lower Bounds on Length of Boolean Formulas. *SIAM J. Comput.*, 11(3):416–427, aug 1982.

[223] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, USA, 1st edition, 2009.

[224] David A Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in NC1. *Journal of Computer and System Sciences*, 38(1):150–164, 1989.

[225] Scott Aaronson and Alex Arkhipov. The Computational Complexity of Linear Optics. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 333–342, New York, NY, USA, 2011. Association for Computing Machinery.

[226] Peter W Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Review*, 41(2):303–332, 1999.

[227] Daniel R Simon. On the Power of Quantum Computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997.

[228] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Unger Falk. Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial. *Phys. Rev. Lett.*, 96(25):250401, jun 2006.

[229] Wim van Dam. Implausible consequences of superstrong nonlocality. *Natural Computing*, 12(1):9–12, 2013.

[230] Matthew Amy and Michele Mosca. T-Count Optimization and Reed–Muller Codes. *IEEE Transactions on Information Theory*, 65(8):4771–4784, 2019.

[231] Peter Bürgisser, Michael Clausen, and Mohammad A Shokrollahi. *Algebraic complexity theory*, volume 315. Springer Science & Business Media, 2013.

[232] Guang Hao Low and Isaac L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.*, 118:010501, Jan 2017.

[233] Jeongwan Haah. Product Decomposition of Periodic Functions in Quantum Signal Processing. *Quantum*, 3:190, oct 2019.

[234] Scott Aaronson, Shalev Ben-David, Robin Kothari, Shravas Rao, and Avishay Tal. Degree vs. Approximate Degree and Quantum Implications of Huang's Sensitivity Theorem. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, pages 1330–1342, New York, NY, USA, 2021. Association for Computing Machinery.

[235] Hao Huang. Induced subgraphs of hypercubes and a proof of the Sensitivity Conjecture. *Annals of Mathematics*, 190(3):949–955, 2019.

[236] Andris Ambainis, Kaspars Balodis, Aleksandrs Belovs, Troy Lee, Miklos Santha, and Juris Smotrovs. Separations in Query Complexity Based on Pointer Functions. *J. ACM*, 64(5), sep 2017.

[237] Peter Høyer and Robert Špalek. Quantum circuits with unbounded fan-out. In Helmut Alt and Michel Habib, editors, *STACS 2003*, pages 234–246, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[238] Paul Herringer and Robert Raussendorf. Classification of measurement-based quantum wire in stabilizer PEPS. *Quantum*, 7:1041, June 2023.

[239] Abhishek Anand. On the power of interleaved low-depth quantum and classical circuits. Master's thesis, University of Waterloo, 2022.

[240] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, aug 2018.

[241] Bülent Demirel, Weikai Weng, Christopher Thalacker, Matty Hoban, and Stefanie Barz. Correlations for computation and computation for correlations. *npj Quantum Information*, 7(1):1–8, 2021.

[242] Manoranjan Swain, Amit Rai, Bikash K Behera, and Prasanta K Panigrahi. Experimental demonstration of the violations of Mermin's and Svetlichny's inequalities for W and GHZ states. *Quantum Information Processing*, 18(7):218, 2019.

[243] Bo Yang, Rudy Raymond, Hiroshi Imai, Hyungseok Chang, and Hidefumi Hiraishi. Testing scalable bell inequalities for quantum graph states on ibm quantum devices. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 12(3):638–647, 2022.

[244] F. Baccari, R. Augusiak, I. Šupić, J. Tura, and A. Acín. Scalable bell inequalities for qubit graph states and robust self-testing. *Phys. Rev. Lett.*, 124:020402, Jan 2020.

[245] Ken X Wei, Isaac Lauer, Srikanth Srinivasan, Neereja Sundaresan, Douglas T McClure, David Toyli, David C McKay, Jay M Gambetta, and Sarah Sheldon. Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences. *Phys. Rev. A*, 101(3):032343, mar 2020.

[246] Wei-Jia Huang, Wei-Chen Chien, Chien-Hung Cho, Che-Chun Huang, Tsung-Wei Huang, and Ching-Ray Chang. Mermin's inequalities of multiple qubits with orthogonal measurements on IBM Q 53-qubit system. *Quantum Engineering*, 2(2):e45, 2020.

[247] Meron Sheffer, Daniel Azses, and Emanuele G Dalla Torre. Playing Quantum Nonlocal Games with Six Noisy Qubits on the Cloud. *Advanced Quantum Technologies*, 5(3):2100081, 2022.

[248] Vedran Dunjko, Theodoros Kapourniotis, and Elham Kashefi. Quantum-Enhanced Secure Delegated Classical Computing. *Quantum Info. Comput.*, 16(1–2):61–86, jan 2016.

[249] Stefanie Barz, Vedran Dunjko, Florian Schlederer, Merritt Moore, Elham Kashefi, and Ian A. Walmsley. Enhanced delegated computing using coherence. *Phys. Rev. A*, 93:032339, Mar 2016.

[250] Marco Clementi, Anna Pappa, Andreas Eckstein, Ian A Walmsley, Elham Kashefi, and Stefanie Barz. Classical multiparty computation using quantum resources. *Physical Review A*, 96(6):062317, 2017.

[251] Johan Hastad. On the correlation of parity and small-depth circuits. *SIAM Journal on Computing*, 43(5):1699–1708, 2014.

[252] Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. An average-case depth hierarchy theorem for boolean circuits. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1030–1048, 2015.

[253] Johan Hastad. An average-case depth hierarchy theorem for higher depth. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2016.

[254] Ruiwen Chen, Rahul Santhanam, and Srikanth Srinivasan. Average-Case Lower Bounds and Satisfiability Algorithms for Small Threshold Circuits. In Ran Raz, editor, *31st Conference on Computational Complexity (CCC 2016)*, volume 50 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1:1–1:35, Dagstuhl, Germany, 2016. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[255] Cody Murray and Ryan Williams. Circuit lower bounds for nondeterministic quasi-polytime: an easy witness lemma for np and nqp. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, page 890–901, New York, NY, USA, 2018. Association for Computing Machinery.

[256] Lijie Chen. Non-deterministic quasi-polynomial time is average-case hard for ACC circuits. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1281–1304, 2019.

[257] R. Ryan Williams. Complexity lower bounds from algorithm design. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, 2021.

[258] Prahladh Harsha, Tulasimohan Molli, and Ashutosh Shankar. Criticality of AC0-Formulae. In Amnon Ta-Shma, editor, *38th Computational Complexity Conference (CCC 2023)*, volume 264 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 19:1–19:24, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[259] Ian Parberry and Georg Schnitger. Parallel computation with threshold functions. *Journal of Computer and System Sciences*, 36(3):278–302, June 1988.

[260] Marvin Minsky and Seymour Papert. An introduction to computational geometry. *Cambridge tiass., HIT*, 479(480):104, 1969.

[261] Pierre Baldi and Roman Vershynin. Polynomial threshold functions, hyperplane arrangements, and random tensors, 2019.

[262] Anders Krogh. What are artificial neural networks? *Nature Biotechnology*, 26:195–197, 2008.

[263] Adam Bene Watts and Natalie Parham. Unconditional Quantum Advantage for Sampling with Shallow Circuits. *arXiv e-prints*, page arXiv:2301.00995, January 2023.

[264] Yangjing Dong, Honghao Fu, Anand Natarajan, Minglong Qin, Haochen Xu, and Penghui Yao. The Computational Advantage of MIP* Vanishes in the Presence of Noise. In Rahul Santhanam, editor, *39th Computational Complexity Conference (CCC 2024)*, volume 300 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:71, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[265] Ryan O' Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[266] András Hajnal, Wolfgang Maass, Pavel Pudlák, Mario Szegedy, and György Turán. Threshold circuits of bounded depth. *Journal of Computer and System Sciences*, 46(2):129–154, 1993.

[267] Kristoffer Arnsfelt Hansen and Vladimir V Podolskii. Polynomial threshold functions and boolean threshold circuits. *Information and Computation*, 240:56–73, 2015.

[268] Russell Impagliazzo, Ramamohan Paturi, and Michael E. Saks. Size–depth tradeoffs for threshold circuits. *SIAM Journal on Computing*, 26(3):693–707, 1997.

[269] Daniel M. Kane and Ryan Williams. Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 633–643, New York, NY, USA, 2016. Association for Computing Machinery.

[270] Hussain Anwar, Benjamin J Brown, Earl T Campbell, and Dan E Browne. Fast decoders for qudit topological codes. *New Journal of Physics*, 16(6):063038, 2014.

[271] Louis Paletta, Anthony Leverrier, Alain Sarlette, Mazyar Mirrahimi, and Christophe Vuillot. Robust sparse IQP sampling in constant depth. *Quantum*, 8:1337, May 2024.

[272] Mithuna Yoganathan, Richard Jozsa, and Sergii Strelchuk. Quantum advantage of unitary clifford circuits with magic state inputs. *Proceedings of the Royal Society A*, 475(2225):20180427, 2019.

[273] Atsuya Hasegawa and François Le Gall. Quantum Advantage with Shallow Circuits Under Arbitrary Corruption. In Hee-Kap Ahn and Kunihiko Sadakane, editors, *32nd International Symposium on Algorithms and Computation (ISAAC 2021)*, volume 212 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 74:1–74:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[274] Heribert Vollmer. Relating polynomial time to constant depth. *Theoretical Computer Science*, 207(1):159–170, 1998.

[275] Johan Håstad. *Computational limitations for small depth circuits*. PhD thesis, Massachusetts Institute of Technology, 1986.

[276] Eric W Allender and Klaus W Wagner. Counting hierarchies: polynomial time and constant depth circuits. In *Current trends in theoretical computer science: essays and Tutorials*, pages 469–483. World Scientific, 1993.

[277] Scott Aaronson. BQP and the polynomial hierarchy. In *Proceedings of the forty-second annual ACM symposium on Theory of computing*, pages 141–150, 2010.

[278] Ran Raz and Avishay Tal. Oracle separation of bqp and ph. *ACM Journal of the ACM (JACM)*, 69(4):1–21, 2022.

[279] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, 1990.

[280] Jonathan E. Moussa. Transversal clifford gates on folded surface codes. *Physical Review A*, 94(4), 2016.

[281] Fern H. E. Watson, Hussain Anwar, and Dan E. Browne. Fast fault-tolerant decoder for qubit and qudit surface codes. *Phys. Rev. A*, 92:032309, 2015.

[282] Hussain Anwar. *Towards fault-tolerant quantum computation with higher-dimensional systems*. PhD thesis, UCL (University College London), 2014.

[283] Sergey Bravyi and Jeongwan Haah. Quantum self-correction in the 3d cubic code model. *Physical Review Letters*, 111(20), 2013.

[284] Benjamin Rossman. An entropy proof of the switching lemma and tight bounds on the decision-tree size of ac0, 2017.

[285] Ryan O'Donnell and Rocco A Servedio. New degree bounds for polynomial threshold functions. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 325–334, 2003.

[286] Valentine Kabanets, Daniel M Kane, and Zhenjian Lu. A polynomial restriction lemma with applications. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 615–628, 2017.

[287] Vladimir Podolskii and Nikolay V. Proskurin. Polynomial Threshold Functions for Decision Lists. In Sang Won Bae and Heejin Park, editors, *33rd International Symposium on Algorithms and Computation (ISAAC 2022)*, volume 248 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 52:1–52:12, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[288] Daniel Gottesman. Fault-tolerant quantum computation with higher-dimensional systems. In *NASA International Conference on Quantum Computing and Quantum Communications*, pages 302–313. Springer, 1998.

[289] Daniel Gottesman. Fault-Tolerant Quantum Computation with Constant Overhead. *arXiv e-prints*, page arXiv:1310.2984, October 2013.

[290] Omar Fawzi, Antoine Grospellier, and Anthony Leverrier. Constant overhead quantum fault-tolerance with quantum expander codes. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018.

[291] Ehsan Zahedinejad, Joydip Ghosh, and Barry C. Sanders. Designing high-fidelity single-shot three-qubit gates: A machine-learning approach. *Phys. Rev. Appl.*, 6:054005, 2016.

[292] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69:062311, 2004.

[293] Gilles Brassard, Anne Broadbent, and Alain Tapp. Recasting mermin's multi-player game into the framework of pseudo-telepathy. *arXiv preprint quant-ph/0408052*, 2004.

[294] Antonio deMarti iOlius, Patricio Fuentes, Román Orús, Pedro M. Crespo, and Josu Etxezarreta Martinez. Decoding algorithms for surface codes. *Quantum*, 8:1498, October 2024.

[295] Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965.

[296] Hussain Anwar, Benjamin J Brown, Earl T Campbell, and Dan E Browne. Fast decoders for qudit topological codes. *New Journal of Physics*, 16(6):063038, 2014.

[297] Adrian Hutter, Daniel Loss, and James R Wootton. Improved hdrg decoders for qudit and non-abelian quantum error correction. *New Journal of Physics*, 17(3):035017, 2015.

[298] Jacob Van Den Berg and Harry Kesten. Inequalities with applications to percolation and reliability. *Journal of applied probability*, 22(3):556–569, 1985.

[299] Nadish de Silva. Efficient quantum gate teleportation in higher dimensions. *Proceedings of the Royal Society A*, 477(2251):20200865, 2021.

[300] Shawn X. Cui, Daniel Gottesman, and Anirudh Krishna. Diagonal gates in the clifford hierarchy. *Physical Review A*, 95(1), January 2017.

[301] Mark Howard and Jiri Vala. Qudit versions of the qubit $\pi/8$ gate. *Phys. Rev. A*, 86:022316, August 2012.

[302] Craig Gidney and Austin G. Fowler. Efficient magic state factories with a catalyzed $|CCZ>$ to $2|T>$ transformation. *Quantum*, 3:135, April 2019.

[303] Michael Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. Lower bounds on the non-clifford resources for quantum computations. *Quantum Science and Technology*, 5(3):035009, May 2020.

[304] Kishor Bharti and Rahul Jain. On the power of geometrically-local classical and quantum circuits. *arXiv e-prints*, page arXiv:2310.01540, 2023.

[305] Ian Parberry. *Circuit Complexity and Neural Networks*. Foundations of Computing. MIT Press, London, England, 1994.

[306] John S. Shawe-Taylor, Martin H.G. Anthony, and Walter Kern. Classes of feedforward neural networks and their circuit complexity. *Neural Networks*, 5(6):971–977, 1992.

[307] Paul Smolensky. *Mathematical Perspectives on Neural Networks*. Psychology Press, 2013.

[308] Carlo Mereghetti and Beatrice Palano. Threshold circuits for iterated matrix product and powering. *RAIRO - Theoretical Informatics and Applications*, 34(1):39–46, 2000.

[309] Raymond Greenlaw, H James Hoover, and Walter L Ruzzo. *A compendium of problems complete for P*. Citeseer, 1991.

[310] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[311] Michael Hahn. Theoretical limitations of self-attention in neural sequence models. *Transactions of the Association for Computational Linguistics*, 8:156–171, 2020.

[312] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2048–2057, Lille, France, July 2015. PMLR.

[313] Gamaleldin F. Elsayed, Simon Kornblith, and Quoc V. Le. *Saccader: improving accuracy of hard attention models for vision*. Curran Associates Inc., Red Hook, NY, USA, 2019.

[314] Lena Strobl. Average-Hard Attention Transformers are Constant-Depth Uniform Threshold Circuits. *arXiv e-prints*, page arXiv:2308.03212, August 2023.

[315] Jorge Pérez, Pablo Barceló, and Javier Marinkovic. Attention is turing-complete. *Journal of Machine Learning Research*, 22(75):1–35, 2021.

[316] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17(9):1013–1017, 2021.

[317] Eric R Anschuetz and Xun Gao. Arbitrary polynomial separations in trainable quantum machine learning. *arXiv:2402.08606*, 2024.

[318] Eric R Anschuetz, Hong-Ye Hu, Jin-Long Huang, and Xun Gao. Interpretable quantum advantage in neural sequence learning. *PRX Quantum*, 4(2):020338, 2023.

[319] Joseph Bowles, Victoria J Wright, Máté Farkas, Nathan Killoran, and Maria Schuld. Contextuality and inductive bias in quantum machine learning. *arXiv:2302.01365*, 2023.

[320] Xun Gao, Eric R. Anschuetz, Sheng-Tao Wang, J. Ignacio Cirac, and Mikhail D. Lukin. Enhancing generative models via quantum correlations. *Phys. Rev. X*, 12:021037, 2022.

[321] Zhihan Zhang, Weiyuan Gong, Weikang Li, and Dong-Ling Deng. Quantum-classical separations in shallow-circuit-based learning with and without noises. *arXiv:2405.00770*, 2024.

[322] Hsin-Yuan Huang, Yunchao Liu, Michael Broughton, Isaac Kim, Anurag Anshu, Zeph Landau, and Jarrod R McClean. Learning shallow quantum circuits. *arXiv:2401.10095*, 2024.

[323] Xun Gao and Lu-Ming Duan. Efficient representation of quantum many-body states with deep neural networks. *Nature communications*, 8(1):662, 2017.

[324] Amira Abbas, David Sutter, Christa Zoufal, Aurélien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, 2021.

[325] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Shan You, and Dacheng Tao. Learnability of quantum neural networks. *PRX Quantum*, 2:040337, 2021.

[326] Yongxin Song, Liberto Beltrán, Ilya Besedin, Michael Kerschbaum, Marek Pechal, François Swiadek, Christoph Hellings, Dante Colao Zanuz, Alexander Flasby, Jean-Claude Besse, and Andreas Wallraff. Realization of Constant-Depth Fan-Out with Real-Time Feedforward on a Superconducting Quantum Processor. *arXiv e-prints*, page arXiv:2409.06989, September 2024.

[327] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[328] Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. Complexity theory column 89: The polynomial hierarchy, random oracles, and boolean circuits. *SIGACT News*, 46(4):50–68, dec 2015.

[329] Sergey Bravyi, Isaac Kim, Alexander Kliesch, and Robert Koenig. Adaptive constant-depth circuits for manipulating non-abelian anyons. *arXiv:2205.01933*, 2022.

[330] Merrick Furst, James B Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical systems theory*, 17(1):13–27, 1984.

[331] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[332] Robert Raussendorf and Hans J. Briegel. Computational model underlying the one-way quantum computer. *Quantum Info. Comput.*, 2(6):443–486, oct 2002.

[333] Maarten Van den Nest, Akimasa Miyake, Wolfgang Dür, and Hans J. Briegel. Universal resources for measurement-based quantum computation. *Phys. Rev. Lett.*, 97:150504, Oct 2006.

[334] Robert Ivan Booth. *Measurement-based quantum computation beyond qubits*. Theses, Sorbonne Université, February 2022.

[335] Dan Browne, Elham Kashefi, and Simon Perdrix. Computational depth complexity of measurement-based quantum computation. In Wim van Dam, Vivien M. Kendon, and Simone Severini, editors, *Theory of Quantum Computation, Communication, and Cryptography*, pages 35–46, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[336] Zhenning Liu and Alexandru Gheorghiu. Depth-efficient proofs of quantumness. *Quantum*, 6:807, September 2022.

[337] Oded Goldreich. On teaching the approximation method for circuit lower bounds. *Eletronical Colloquium on Computational Complexity*, 2023.

[338] Ryuhei Mori. Periodic fourier representation of boolean functions. *arXiv preprint arXiv:1803.09947*, 2018.

[339] Filipa C. R. Peres. The Pauli-based model of quantum computation with higher dimensional systems. *arXiv e-prints*, page arXiv:2302.13702, 2023.

[340] Jelena Mackeprang, Daniel Bhatti, Matty J Hoban, and Stefanie Barz. The power of qutrits for non-adaptive measurement-based quantum computing. *New Journal of Physics*, 25(7):073007, jul 2023.

[341] Bao Luong. *Fourier analysis on finite Abelian groups*. Springer Science & Business Media, 2009.

[342] Ryan Babbush, Dominic W. Berry, Robin Kothari, Rolando D. Somma, and Nathan Wiebe. Exponential quantum speedup in simulating coupled classical oscillators. *Phys. Rev. X*, 13:041041, Dec 2023.

[343] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[344] Jeremy L O'Brien. Optical Quantum Computing. *Science*, 318(5856):1567–1570, 2007.

[345] Anne Broadbent, Joseph Fitzsimons, and Elham Kashefi. Universal blind quantum computation. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 517–526, 2009.

[346] Will Simmons. Relating measurement patterns to circuits via pauli flow. *arXiv preprint arXiv:2109.05654*, 2021.

[347] Monireh Houshmand, Mahboobeh Houshmand, and Joseph F. Fitzsimons. Minimal qubit resources for the realization of measurement-based quantum computation. *Phys. Rev. A*, 98:012318, Jul 2018.

[348] Raphael Dias da Silva, Einar Pius, and Elham Kashefi. Global quantum circuit optimization. *arXiv preprint arXiv:1301.0351*, 2013.

[349] Jisho Miyazaki, Michal Hajdušek, and Mio Murao. Analysis of the trade-off between spatial and temporal resources for measurement-based quantum computation. *Phys. Rev. A*, 91:052302, May 2015.

[350] Measurement-induced entanglement and teleportation on a noisy quantum processor. *Nature*, 622(7983):481–486, 2023.

[351] Emma Deist, Yue-Hui Lu, Jacquelyn Ho, Mary Kate Pasha, Johannes Zeiher, Zhenjie Yan, and Dan M. Stamper-Kurn. Mid-circuit cavity measurement in a neutral atom array. *Phys. Rev. Lett.*, 129:203602, Nov 2022.

[352] Jin Ming Koh, Shi-Ning Sun, Mario Motta, and Austin J Minnich. Experimental realization of a measurement-induced entanglement phase transition on a superconducting quantum processor. *arXiv preprint arXiv:2203.04338*, 2022.

[353] Matthew DeCross, Eli Chertkov, Megan Kohagen, and Michael Foss-Feig. Qubit-reuse compilation with mid-circuit measurement and reset. *arXiv preprint arXiv:2210.08039*, 2022.

[354] Stefanie Barz, Joseph F Fitzsimons, Elham Kashefi, and Philip Walther. Experimental verification of quantum computation. *Nature physics*, 9(11):727–731, 2013.

[355] Matty J Hoban, Joel J Wallman, and Dan E Browne. Generalized Bell-inequality experiments and computation. *Phys. Rev. A*, 84(6):62107, dec 2011.

[356] Alessandro Luongo, Antonio Michele Miti, Varun Narasimhachar, and Adithya Sireesh. Measurement-based uncomputation of quantum circuits for modular arithmetic. *arXiv e-prints*, page arXiv:2407.20167, July 2024.

[357] Sergey Bravyi, David Gosset, Robert Koenig, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020.

[358] Francisca Vasconcelos and Hsin-Yuan Huang. Learning shallow quantum circuits with many-qubit gates. *arXiv e-prints*, page arXiv:2410.16693, October 2024.

[359] Anne Broadbent, Joseph Fitzsimons, and Elham Kashefi. Universal Blind Quantum Computation. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 517–526, 2009.

[360] Atul Singh Arora, Kishor Bharti, Alexandru Cojocaru, and Andrea Coladangelo. A computational test of quantum contextuality, and even simpler proofs of quantumness. *arXiv e-prints*, page arXiv:2405.06787, May 2024.

[361] Matilde Baroni, Quoc-Huy Vu, Boris Bourdoncle, Eleni Diamanti, Damian Markham, and Ivan Šupić. Quantum bounds for compiled XOR games and d-outcome CHSH games. *arXiv e-prints*, page arXiv:2403.05502, March 2024.

[362] Yael Kalai, Alex Lombardi, Vinod Vaikuntanathan, and Lisa Yang. Quantum advantage from any non-local game. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 1617–1628, 2023.

[363] Anand Natarajan and Tina Zhang. Bounding the quantum value of compiled nonlocal games: from chsh to bqp verification. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1342–1348. IEEE, 2023.

[364] Nasir Ahmed and Kamisetty Ramamohan Rao. Walsh-hadamard transform. In *Orthogonal transforms for digital signal processing*, pages 99–152. Springer, 1975.

[365] Philip Feinsilver and Jerzy Kocik. *Krawtchouk Polynomials and Krawtchouk Matrices*, pages 115–141. Recent Advances in Applied Probability. Springer US, Boston, MA, 2005.

[366] Philip Feinsilver and Rene Schott. Krawtchouk transforms and convolutions. *Bulletin of Mathematical Sciences*, pages 1–19, 2018.

[367] M. Stobińska, A. Buraczewski, M. Moore, W. R. Clements, J. J. Renema, S. W. Nam, T. Gerrits, A. Lita, W. S. Kolthammer, A. Eckstein, and I. A. Walmsley. Quantum interference enables constant-time quantum information processing. *Science Advances*, 5(7):eaau9674, 2019.

Part III
Appendices

Appendix A
Mathematical transformations

A.1 Walsh-Hadamard transform

The Walsh-Hadamard transform [364, 42] $\mathcal{F}$ is a Fourier transform that has their characters $X_S(x)$ defined as

$$X_S(x) = \begin{cases} \prod_{i \in S} x_i, & S \subseteq ([n] \setminus \emptyset) \\ 1, & S = \emptyset \end{cases} , \tag{A.1.1}$$

with $x \in \{1, -1\}^n$ and $S \subseteq [n]$. These characters are the real parity functions that are used to define the multi-linear polynomials (Definition Definition 2). Moreover, the matrix that defines the transformation is composed of all possible characters $X_S(x)$ that can be computed based on the size of the input string $S \subseteq [n]$ evaluated over all possible input values $x \in \{1, -1\}^n$,

$$\mathcal{F} = \begin{bmatrix} X_{S_1}(x_1) & X_{S_2}(x_1) & X_{S_3}(x_1) & ... & X_{S_{2^n}}(x_1) \\ X_{S_1}(x_2) & X_{S_2}(x_2) & X_{S_3}(x_2) & ... & X_{S_{2^n}}(x_2) \\ X_{S_1}(x_3) & X_{S_2}(x_3) & X_{S_3}(x_3) & ... & X_{S_{2^n}}(x_3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ X_{S_1}(x_{2^n}) & X_{S_2}(x_{2^n}) & X_{S_2}(x_{2^n}) & ... & X_{S_{2^n}}(x_{2^n}) \end{bmatrix} . \tag{A.1.2}$$

Furthermore, the transposed matrix of the Walsh-Hadamard transform applied to the column vector containing the truth table of the Boolean function generates the Fourier coefficients,

$$\frac{1}{2^n} \begin{bmatrix} X_{S_1}(x_1) & X_{S_1}(x_2) & X_{S_1}(x_2) & ... & X_{S_1}(x_{2^n}) \\ X_{S_2}(x_1) & X_{S_2}(x_2) & X_{S_2}(x_2) & ... & X_{S_2}(x_{2^n}) \\ X_{S_3}(x_1) & X_{S_3}(x_2) & X_{S_3}(x_2) & ... & X_{S_3}(x_{2^n}) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ X_{S_{2^n}}(x_1) & X_{2^n}(x_2) & X_{2^n}(x_2) & ... & X_{2^n}(x_{2^n}) \end{bmatrix} * \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_{2^n}) \end{bmatrix} = \begin{bmatrix} \widehat{f}(S_1) \\ \widehat{f}(S_2) \\ \widehat{f}(S_3) \\ \vdots \\ \widehat{f}(S_{2^n}) \end{bmatrix} . \tag{A.1.3}$$

Given that, every element of the resulting column vector is derived from the equation

$$\widehat{f}(S) = \sum_{x_i \in \{0,1\}^n} f(x_i) X_S(g^{-1}(x_i)) , \tag{A.1.4}$$

which computes the Fourier coefficient of the Boolean function [193].

A.2 Krawtchouk transform

The Krawtchouk transform [365, 366, 367], $\mathcal{K}$ is defined by a matrix with entries specified by the following expressions,

$$K_{i,j} = \begin{cases} \sum_{k=0}^{i}(-1)^k\binom{j}{k}\binom{n-j}{i-k}, & i < j \; and \; i+j <= n \\ \sum_{k=0}^{j}(-1)^k\binom{j}{k}\binom{n-j}{i-k}, & i >= j \; and \; i+j < n \\ \sum_{k=j+i-n}^{j}(-1)^k\binom{j}{k}\binom{n-j}{i-k}, & i >= j \; and \; i+j >= n \\ \sum_{k=0}^{n-j}(-1)^{i-k}\binom{j}{i-k}\binom{n-j}{k}, & i < j \; and \; i+j > n \end{cases} . \tag{A.2.1}$$

The matrix has, therefore, the following form,

$$\mathcal{K}^{[n]} = \begin{bmatrix} \binom{n}{0} & \binom{n}{1} & \dots & \binom{n}{i} & \dots & \binom{n}{n} \\ \binom{n-1}{0} & -\binom{n-1}{0}+\binom{n-1}{1} & \dots & -\binom{n-1}{1}+\binom{n-1}{2} & \dots & -\binom{n-1}{n-1} \\ \binom{n-2}{0} & -2\binom{n-2}{0}+\binom{n-2}{1} & \dots & \binom{n-2}{0}-2\binom{n-2}{1}+\binom{n-2}{2} & \dots & \binom{n-2}{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \binom{n-j}{0} & \dots & \dots & rf_{i,j} = \sum_{k=0}^{j}(-1)^{j-k}\binom{j}{k}\binom{n-j}{i-k} & \dots & (-1)^j\binom{n-j}{n-j} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (-1)^0\binom{n}{n} & (-1)^1\binom{n}{n-1} & \dots & (-1)^i\binom{n}{n-i} & \dots & (-1)^{i=j}\binom{n}{0} \end{bmatrix} \tag{A.2.2}$$

with dimension $n+1 \times n+1$.

Appendix B
Concentration Inequalities

B.1 Hoeffding's Inequality

We use the Hoeffding's inequality for $X_1, \ldots, X_m$ independent random variables with $a_i \leq X_i \leq b_i$. Consider $S_m = X_1 + \cdots + X_m$. Then it holds:

$$\forall t > 0 : \mathsf{Pr}[S_m - \mathbb{E}[S_m] \geq t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{m}(b_i - a_i)^2}\right) \tag{B.1.1}$$

If we substitute $t = m\tilde{t}$ we obtain

$$\mathsf{Pr}\left[\frac{S_m}{m} - \frac{\mathbb{E}[S_m]}{m} \geq \tilde{t}\right] \leq \exp\left(-\frac{2m^2\tilde{t}^2}{\sum_{i=1}^{m}(b_i - a_i)^2}\right), \tag{B.1.2}$$

because $b_i = 1$ and $a_i = 0$. Now we utilize the fact that X_i are i.i.d. binary random variables. Therefore

$$\mathbb{E}[S_m] = m\mathbb{E}[X] = mp \tag{B.1.3}$$

where p is the probability of getting 1 as result. Also: $(b_i - a_i)^2 = 1$. At the same time we rename $\frac{S_m}{m} = z$ to obtain

$$\mathsf{Pr}[z - p \geq t] \leq \exp\big(-2mt^2\big). \tag{B.1.4}$$

Finally, rearranging gives us

$$\mathsf{Pr}[p \leq z - t] \leq \exp\big(-2mt^2\big), \tag{B.1.5}$$

which is the Hoeffding's inequality for m repeated biased coin flips.