Short-time Quantum Dynamics: Classical Simulation, Quantum Advantage, and Characterization

by

Yinchen Liu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization (Quantum Information)

Waterloo, Ontario, Canada, 2026

© Yinchen Liu 2026

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Jonas Haferkamp, Associate Professor Faculty of Computer Science Ruhr University Bochum
Supervisor:	David Gosset, Associate Professor Department of Combinatorics and Optimization University of Waterloo
Internal Member:	Debbie Leung, Professor Department of Combinatorics and Optimization University of Waterloo
Internal Member:	Ashwin Nayak, Professor Department of Combinatorics and Optimization University of Waterloo
Internal-External Member:	Luke Schaeffer, Assistant Professor Cheriton School of Computer Science University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

- I am the sole author of Chapter 1.
- Chapter 2 is based on the paper [1], which is joint work with Sergey Bravyi and David Gosset.
- Chapter 3, except for Sections 3.2 and 3.5, is based on the paper [2], which is joint work with Adam Bene Watts, David Gosset, and Mehdi Soleimanifar.
- Chapter 4 is based on the paper [3], which is joint work with Max McGinley, Thomas Schuster, and David Gosset.
- Chapter 5 is based on the paper [4], which is joint work with James Seddon, Tamara Kohler, Emilio Onorati, and Toby Cubitt.

Abstract

This thesis presents several results on short-time quantum dynamics, with themes spanning classical simulation, quantum advantage, and characterization. A prime example of short-time quantum dynamics is that of a shallow quantum circuit, whose circuit depth is a constant independent of the number of qubits. Another example is the dynamics generated by a single 2-qubit gate, even the identity gate, on a noisy quantum computer.

In Chapter 2, we give classical simulation algorithms for peaked shallow quantum circuits. We say an n-qubit quantum circuit is *peaked* if it has an output probability that is at least inverse-polynomially large as a function of n. We describe a classical algorithm with quasipolynomial $n^{O(\log n)}$ runtime that approximately samples from the output distribution of a peaked shallow quantum circuit. We give even faster algorithms for circuits composed of nearest-neighbour gates on a D-dimensional grid of qubits, with polynomial $n^{O(1)}$ runtime if $D = 2$ and almost-polynomial $n^{O(\log \log n)}$ runtime for $D > 2$. Our sampling algorithms can be used to estimate the output probabilities of shallow quantum circuits to within a given inverse-polynomial additive error, and for cases other than $D = 2$, our algorithms achieve improved runtime scalings compared to prior state-of-the-art algorithms.

In Chapters 3 and 4, we study shallow 1D and 2D random quantum circuits. The output states of these circuits only exhibit entanglement among spatially neighbouring qubits due to locally constrained lightcones. However, entanglement between distant qubits could be generated if we were to measure all other qubits of an output state in the computational basis. For 2D random circuits composed of Haar random 2-qubit gates, Napp et al. [5] conjectured that *long-range measurement-induced entanglement* (MIE) proliferates when the circuit depth exceeds some constant critical value d^*. They also conjectured that this coincides with a *quantum advantage phase transition* in the classical hardness of sampling from their output distributions [5].

In Chapter 3, we first demonstrate numerically and rigorously the emergence of long-range MIE in the output states of 1D and 2D random Clifford circuits in various settings. Then, we prove an unconditional quantum advantage: any distribution of shallow random Clifford circuits satisfying a long-range GHZ-type MIE property cannot be simulated by shallow classical probabilistic circuits with bounded fan-in. This could be viewed as establishing a Clifford/scaled-down version of the quantum advantage phase transition originally conjectured by Napp et al. for shallow 2D Haar random quantum circuits.

In Chapter 4, we concentrate on the computational-basis output distributions of shallow 2D random Clifford circuits with depth $d > d^*$. We propose and support with numerical evidence a conjecture called the Scrooge hypothesis that relates the distribution of pure post-measurement states on a subset of qubits, generated by measuring all other qubits of an output state in the computational basis, to stabilizer Scrooge ensembles, which are natural generalizations of the distribution of uniformly random stabilizer states. We show that the Scrooge hypothesis implies a detailed characterization of the way local and global correlations coexist in the output distribution of a typical shallow 2D random Clifford circuit, corroborating numerically observed behaviours. Building on the structural characterization, we show that a column-by-column algorithm, which only ever keeps track of $O(\log n)$ consecutive columns of qubits in the 2D grid at a time, is able to correctly simulate the marginal output distribution if a contiguous region of $O(\log n)$ qubits is traced out. The correctness

of the column-by-column simulation algorithm implies that any global correlation residing in the output distribution is fragile against local depolarizing noise.

In Chapter 5, we tackle the problem of characterizing the short-time dynamics of a single 2-qubit gate implemented on current quantum computing hardware. Building on prior work, we devise a suite of algorithms utilizing alternating optimization techniques to simultaneously and self-consistently estimate state preparation and measurement (SPAM) errors and best-fit Lindbladian models for a set of gates. We benchmark our algorithms extensively using synthetic 2-qubit gate data under a range of realistic error models and apply them to detect cross-talk on real superconducting-qubit hardware.

Acknowledgements

I would like to thank

- my advisor David Gosset for his guidance throughout my whole graduate career; being able to work with David has been one of my greatest fortunes;
- my examining committee members Jonas Haferkamp, Debbie Leung, Ashwin Nayak, and Luke Schaeffer for taking the time to serve on the committee, reading this thesis, and providing valuable feedback;
- all my collaborators (in rough chronological order): Sergey Bravyi, Giuseppe Carleo, Gregory Bentsen, Bill Fefferman, Soumik Ghosh, Michael Gullans, James Seddon, Tamara Kohler, Emilio Onorati, Toby Cubitt, Adam Bene Watts, Mehdi Soleimanifar, Benjamin Wong, Max McGinley, and Thomas Schuster;
- IQC and the C&O department for fostering a supportive environment throughout the years;
- my parents and Cindy for their continued love and support.

Table of Contents

List of Figures x

List of Tables xii

1 Introduction and Preliminaries 1
- 1.1 Classical simulation of quantum circuits . . . 2
- 1.2 Random quantum circuits . . . 5
- 1.3 Preliminaries . . . 9

2 Classical simulation of peaked shallow quantum circuits 17
- 2.1 General peaked shallow circuits . . . 22
- 2.2 Geometrically local peaked shallow circuits . . . 28
- 2.3 Open questions . . . 42

3 Measurement-induced entanglement in low-depth random Clifford circuits 43
- 3.1 The coarse-grained architectures . . . 45
- 3.2 Long-range MIE in 1D coarse-grained random Clifford circuits . . . 47
- 3.3 Long-range MIE in 2D coarse-grained random Clifford circuits in annulus tripartitions . . . 51
- 3.4 Long-range tripartite MIE in 2D coarse-grained random Clifford circuits . . . 60
- 3.5 Long-range MIE in 2D coarse-grained random Clifford circuits in the row-by-row tripartition . . . 63
- 3.6 Quantum advantage from tripartite MIE . . . 68

4 The stabilizer Scrooge ensemble and shallow 2D random Clifford circuits 76
- 4.1 Numerical investigation of the output distributions of shallow 2D random Clifford circuits . . . 76
- 4.2 The stabilizer Scrooge ensemble . . . 81
- 4.3 Structure of the output distributions of shallow 2D random Clifford circuits . . . 92
- 4.4 Numerical evidence for the Scrooge hypothesis . . . 96
- 4.5 Column-by-column simulation of shallow 2D random Clifford circuits . . . 100

5 Alternating optimization methods for robust Lindbladian estimation 104
- 5.1 Preliminaries . . . 107
- 5.2 Our work . . . 113
- 5.3 Synthetic testing of Alternating Projections and Convex Solve . . . 122
- 5.4 Synthetic testing of Gate Set Flip-Flop . . . 127
- 5.5 Analysis of tomographic data from superconducting qubit hardware . . . 130

5.6 Deferred discussions 148
5.7 Future directions 155

Bibliography **156**

List of Figures

1.1 Circuit diagram of a 20-qubit depth-3 quantum circuit composed of 2-qubit gates . . . 11
1.2 Gate patterns of the 2D brickwork architecture on a 6×6 grid . . . 11
1.3 Circuit diagram for entanglement swapping . . . 14

2.1 Numerical simulation of 2D random Clifford circuits interspersed with $R(\theta)$ rotations . . . 27
2.2 A partition of the grid of qubits into equal-width rectangular regions . . . 32

3.1 An example of an annulus tripartition . . . 44
3.2 Numerical estimation of MIE in an annulus tripartition . . . 45
3.3 Circuit diagram of the 1D coarse-grained architecture . . . 46
3.4 Gate arrangement of a 2D coarse-grained circuit . . . 46
3.5 Circuit diagram for probing MIE in 1D coarse-grained random Clifford circuits. 49
3.6 Example of a cluster . . . 55
3.7 2D coarse-grained row-by-row tripartition iteration . . . 64
3.8 Numerical estimation of the expected number of stabilizers of $|\psi'\rangle$ local on A as a function of L_y . . . 65
3.9 Numerical estimation of $\mathrm{MaxInc}(L_x, 4)$ . . . 66
3.10 Fitting an exponential model to $\mathrm{MaxInc}(L_x, \tau)$ . . . 67

4.1 Examples of typical weight distributions of the Z-type stabilizers of the output states of depth-7 2D brickwork random Clifford circuits . . . 77
4.2 Weight histograms of the Z-type stabilizer generators of the full output states versus reduced output states . . . 80
4.3 Circuit diagram for drawing from a Scrooge ensemble. The $|B|$-qubit unitary U is Haar random. . . . 81
4.4 An example of an ABC tripartition in the set $\mathcal{T}$ . . . 92
4.5 12×12 examples of the ABC tripartitions considered in our numerical experiments . . . 97
4.6 Dependence of the average relative error on width(B) for the row-by-row tripartition . . . 98
4.7 Dependence of the average relative error on width(B) for the row-by-row with two corners tripartition . . . 99
4.8 Dependence of the average relative error on $|\partial A| = |A|$ for the row-by-row tripartition . . . 100

5.1 Circuit diagram for estimating $\mathrm{Tr}((X \otimes X)\mathcal{E}(|1\rangle\langle 1| \otimes |+_i\rangle\langle +_i|))$ where $\mathcal{E}$ corresponds to a noisy CNOT gate. . . . 104

5.2 CNOT gate with coherent ZI, IZ, and ZZ errors and dissipative ZI error . 123
5.3 The same noisy CNOT gate considered in Figure 5.2, at the level of statistical noise induced by 10^4 shots 123
5.4 For the same input E considered in Figure 5.3, the alternating projections algorithm finds a Lindbladian that exponentiates very closely to the input . 124
5.5 Distribution of the $\|e^L - E\|$ (top) and $\|e^L - E^*\|$ (bottom) values for the 20 instances of CNOT with coherent X and dephasing noise tested 126
5.6 Canonical decomposition of a CNOT with coherent Z, amplitude damping, and dephasing noise instance 127
5.7 CNOT with increasing strengths of (a)(b) coherent X, amplitude damping, and dephasing gate noise and (c)(d) overrotation and dephasing gate noise . 128
5.8 Each noise model is tested with 10 levels of increasing noise strengths 129
5.9 Synthetic test results for the Gate Set Flip-Flop algorithm on 1000 randomly generated test cases 130
5.10 Gate Set Flip-Flop figures of merit for data collected from *ibm_perth* 134
5.11 Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth* 135
5.12 Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)^3 = R_{ZX}(1.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth* 136
5.13 Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)^5 = R_{ZX}(2.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth* 137
5.14 Partial qubit layout on *ibm_cairo* 138
5.15 Gate Set Flip-Flop figures of merit for data collected from *ibm_cairo*, tomographing qubits 13 and 14 139
5.16 Comparison of the canonical decompositions of the ideal unitary CNOT gate with the fitted Lindbladians after Gate Set Flip-Flop for the noisy CNOT gate implemented on qubit pair $(13, 14)$ on *ibm_cairo* 140
5.17 The estimated canonical decomposition (Section 5.1.1) of the estimated Lindbladian generators for the CNOT implemented on qubit pair $(13, 14)$ on *ibm_cairo* 141
5.18 Estimated canonical decomposition (Section 5.1.1) of the estimated Lindbladian generators for the CNOT implemented on qubit pair $(13, 14)$ on *ibm_cairo* 142
5.19 Qubit layout on *ibm_perth* 143
5.20 Circuit decomposition for CNOT between qubits 0 and 5 on *ibm_perth*. . . . 144
5.21 Gate Set Flip-Flop figures of merit for the idling experiment on qubits $(0, 1)$ and $(0, 5)$ from *ibm_perth* 145
5.22 Canonical decompositions of the estimated Lindbladians for three variants of the idling process on the adjacent qubit pair $(0, 1)$ on *ibm_perth* 146
5.23 Canonical decompositions of idling processes for the separated qubit pair $(0, 5)$ on *ibm_perth* 147

List of Tables

2.1 Runtime of classical algorithms for approximating output probabilities of n-qubit shallow quantum circuits 19

4.1 Numerical estimation of the distribution of the integer $r = I(A;C|B)$ 79

5.1 Each gate and noise model combination is tested on 20 input instances generated using simulated quantum process tomography with 10^4 shots 125
5.2 Distances between the tomographic data and corresponding ideal transfer matrices from gate set analysis on *ibm_perth* 133
5.3 Distances between the tomographic channel estimates and corresponding ideal transfer matrices from gate cross-talk analysis on *ibm_cairo* 139
5.4 Distances between the tomographic channel estimates and corresponding ideal transfer matrices from idling analysis on *ibm_perth* 144

Chapter 1

Introduction and Preliminaries

Quantum computers promise potential applications throughout science, engineering, and business [6, 7]. Fully realizing these applications requires the deployment of fault-tolerant quantum computers that, even under optimistic forecasts [8], are still years beyond the latest hardware breakthroughs [9–12]. Meanwhile, today's quantum computers, lacking functional error correction, suffer irreversibly from the accumulation of noise the longer they run. The present reality has thus called for the exploration of the potential utilities of short-time quantum dynamics, such as shallow quantum circuits, which when aided by error mitigation [13], have better hopes of yielding significant signals before noise overwhelms the device.

Shallow quantum circuits, whose circuit depth is a constant independent of the number of qubits, have been the subject of extensive research for over a decade, with wide ranging perspectives from the experimental demonstration of variational quantum algorithms [14–18] to the theoretical separation of quantum and classical complexity classes [19]. In Chapters 2 to 4 of this thesis, we report on new research findings that advance our understanding of shallow quantum circuits in aspects motivated by computational complexity theory. We study peaked shallow quantum circuits in Chapter 2 and shallow random Clifford circuits in Chapters 3 and 4. The themes of the results presented span classical simulation (Chapter 2 and Section 4.5), quantum advantage (Section 3.6), and characterization (Sections 3.1 to 3.5 and 4.1 to 4.4).

Chapter 5 tackles a different problem regarding one of the shortest kinds of short-time quantum dynamics, that of a single 2-qubit gate. We consider the practical task of characterizing the noise in a 2-qubit gate implemented on current quantum computing hardware. Absent error correction, to maximize the performance of the device at hand, a detailed understanding of the noise affecting the hardware will be invaluable in tailoring error mitigation and suppression techniques [20–23]. Better error mitigation, in turn, allows clearer signals to be extracted from running shallow quantum circuits on near-term quantum computers.

In this chapter, we first provide a brief introduction to the classical simulation of quantum circuits in Section 1.1, defining a canonical classical simulation task and explaining the main motivations. Section 1.2 includes a brief introduction to random quantum circuits, covering the concepts that are most relevant to the works presented later in Chapters 3 and 4. In Section 1.3, we establish some common notations that will be used throughout this thesis and introduce some concepts and notions that will appear.

1.1 Classical simulation of quantum circuits

The circuit model of quantum computation is the most widely used framework for describing quantum algorithms. In addition to being a common language for algorithm design, quantum circuits have also been studied from the perspective of computational complexity theory. Let $n \geq 1$ be an integer. In a complexity-theoretic context, we consider a set of n-qubit quantum circuits $\mathcal{F}_n$, viewed as a subset of the set of all n qubit quantum circuits, and study the computational power of circuits in $\mathcal{F}_n$ as a whole and as n grows. One potentially insightful question is to ask whether the circuits in $\mathcal{F}_n$ can be efficiently classically simulated.

Let U be an n-qubit quantum circuit with output state $|\psi\rangle = U|0^n\rangle$. We call the distribution $P(x) = |\langle x|\psi\rangle|^2$, $x \in \{0,1\}^n$ the computational-basis output distribution of U. When U is executed on a perfect quantum computer starting from the initial state $|0^n\rangle$ and all qubits of the final state $|\psi\rangle$ are measured in the computational basis, the quantum computer outputs a string sample $x \in \{0,1\}^n$ with probability $P(x)$. The goal of classical simulation is to replicate this input-output relation on a classical computer. More concretely, a canonical classical simulation task can be stated as follows.

Definition 1.1 (Classical simulation of quantum circuits)**.** On input the classical description of an n-qubit depth-d quantum circuit U with $d \leq \text{poly}(n)$ and an error tolerance $\varepsilon \geq 0$, output via a classical randomized algorithm a string $x \in \{0,1\}^n$ drawn according to some probability distribution P' satisfying $\|P' - P\|_1 \leq \varepsilon$ with $P(x) = |\langle x|U|0^n\rangle|^2$, $x \in \{0,1\}^n$ being the computational-basis output distribution of the input circuit U.

Let us dissect Definition 1.1. The classical description of a quantum circuit U could be a listing of the gates in U, with each gate specified by a unitary matrix and a list of qubit indices to which the gate is applied. One reason for the restriction $d \leq \text{poly}(n)$ is allowing n to be viewed as the sole scaling parameter so that the runtimes of algorithms could be stated and compared in terms of just n. Formally, the error tolerance ε should be specified in unary as $1^{1/\varepsilon}$ in the case where $\varepsilon > 0$. This way, a $\text{poly}(n, 1/\varepsilon)$-time algorithm properly runs in time polynomial in the description length of the input. We remark that the difficulty of the problem depends sensitively on ε. On one extreme, demanding $\varepsilon = 0$ may be impossible even for $n = 1$ if the classical algorithm's source of randomness is assumed to be an unlimited stream of independent and unbiased random bits. On the other extreme, a guarantee of $\varepsilon \geq 2$ is uninformative since $\|P' - P\|_1 \leq 2$ is true for any two discrete probability distributions P' and P over the same sample space. Often, it is most sensible to demand an error tolerance $\varepsilon = 1/\text{poly}(n)$ because this choice matches the capability of a fully fault-tolerant quantum computer, which on input the circuit U, can only sample from a distribution $\tilde{P}$ such that $\|\tilde{P} - P\|_1 \leq 1/\text{poly}(n)$ [24–27]. Therefore, requiring $\varepsilon = 1/\text{poly}(n)$ leads to a fair competition between quantum and classical computers.

As a sanity check, recall that it is always possible to solve the problem in Definition 1.1 with negligibly small ε using $2^{O(n)}$ time and memory by performing numerical linear algebra on 2^n dimensional vectors. It is a common belief that some exponential in n time should be the fastest possible if the input is allowed to be an arbitrary n-qubit quantum circuit with depth $d \leq \text{poly}(n)$. In contrast, the problem becomes more interesting if we impose further restrictions on the family of input circuits allowed, making the existence of specialized and faster classical algorithms uncertain. A well-known example of a restricted family of quantum

circuits is the family of Clifford circuits, for which the Gottesman-Knill theorem implies that any n-qubit Clifford circuit with depth $d \leq \text{poly}(n)$ can be classically simulated in $\text{poly}(n)$ time [28, 29].

Whenever one discovers a novel classical simulation algorithm for a restricted circuit family, it often results from some sharpened understanding of the fundamental structures common to circuits in that family. For example, the Gottesman-Knill theorem follows as a corollary of the precise characterization of the algebraic structures of the Clifford group and stabilizer states, and the algorithm serves as a concrete manifestation of that abstract understanding. In Chapter 2, we give classical simulation algorithms for peaked shallow quantum circuits. We say an n-qubit quantum circuit U is peaked if there exists an outcome $x^* \in \{0,1\}^n$ such that $|\langle x^*|U|0^n\rangle|^2 \geq 1/\text{poly}(n)$ and shallow if its circuit depth is a constant independent of n. Our algorithms exploit a Hamming weight concentration property that we show is present in the output distributions of peaked shallow quantum circuits. Section 4.5 of Chapter 4 describes an algorithm for classically simulating shallow 2D brickwork random Clifford circuits, which relies on a detailed understanding of the structure of their output distribution developed earlier in the chapter.

The approach of studying the classical simulability of restricted families of quantum circuits could be viewed as conducting ablation studies on the power of quantum computation. It aims to determine whether removing certain elements from the full generality of quantum computing eliminates their quantumness. This perspective bears out more concretely through the lens of quantum advantage.

Classical simulation is closely related to computational quantum advantage in runtime, which is the only kind of quantum advantage we will discuss. At a high level, a quantum advantage proposal seeks to demonstrate that a quantum computer can perform some computational task substantially faster[1] than the fastest possible classical algorithm for the same task. A clean and well-known candidate for quantum advantage is performing integer factorization using Shor's algorithm [30]. More generally, an end-to-end quantum advantage protocol comprises running selected quantum circuits on a quantum computer, interspersed with some classical pre- and post-processing. Each step run on the quantum computer can always be put into the canonical form of running an n-qubit quantum circuit belonging to some family $\mathcal{F}_n$ and measuring all qubits of the output state in the computational basis [30]. Thus, if it is possible to classically simulate circuits in $\mathcal{F}_n$ in an amount of time comparable to that required by the quantum computer, then one could generically substitute every use of the quantum computer with the classical simulator. In that case, the entire protocol could be simulated using classical computers alone, hence denying quantum advantage. In short, we could think of classical simulation as a "dual problem" of quantum advantage. For example, our algorithm for classically simulating peaked shallow quantum circuits given in Chapter 2 runs in quasipolynomial $n^{O(\log n)}$ time, and this implies that running peaked shallow circuits on a quantum computer could not be the sole basis of an exponential quantum advantage.

In Chapter 2, we will also consider another computational task closely related to Definition 1.1, which we define next.

Definition 1.2 (Output probability estimation)**.** On input the classical description of an n-qubit depth-d quantum circuit U with $d \leq \text{poly}(n)$, a string $x \in \{0,1\}^n$, and an error

[1] The precise definition of "substantially faster" varies case by case.

tolerance $\varepsilon \geq 0$, output a value $p \in \mathbb{R}$ satisfying $|p - |\langle x|U|0^n\rangle|^2| \leq \varepsilon$.

Similar to the sampling task defined in Definition 1.1, the difficulty of the estimation task also depends sensitively on the error tolerance ε. If we demand an error $\varepsilon \leq 1/2^n$, then it is known that the output probability estimation problem for arbitrary n-qubit quantum circuits with depth $d \leq \text{poly}(n)$ is $\#P$-hard [31], and it is commonly believed that solving $\#P$-hard problems efficiently is beyond the capability of a perfect quantum computer [32]. In contrast, if we demand $\varepsilon = 1/\text{poly}(n)$, then the problem becomes solvable in $\text{poly}(n)$ time using a fault-tolerant quantum computer by sampling from the input circuit's output distribution $\text{poly}(n)$ times and outputting the empirically observed frequency of the input target string x. However, it is worth noting that when we only require $\varepsilon = 1/\text{poly}(n)$, the $|p - |\langle x|U|0^n\rangle|^2| \leq \varepsilon$ accuracy guarantee could become vacuous since $|\langle x|U|0^n\rangle|^2$ could be inverse exponentially small. In which case, $p = 0$ is a valid solution, and $p = 0$ is indeed what the quantum algorithm described above will most likely output because it will most likely never observe x even once. As such, the estimation problem also becomes more interesting if we impose further restrictions on the input circuit. Notice that for $\varepsilon = 1/\text{poly}(n)$, the peakedness condition arises as a natural restriction since assuming $|\langle x|U|0^n\rangle|^2 \geq 1/\text{poly}(n)$ ensures that a valid estimate p must be non-trivial.

What is the relationship between solving the problem defined in Definition 1.2 via a classical algorithm and quantum advantage? Unlike the clean "duality" between the sampling task and quantum advantage, its relationship with the estimation task is more nuanced. For instance, the existence of a $\text{poly}(n)$-time classical algorithm capable of solving the estimation problem with $\varepsilon = 1/\text{poly}(n)$ for a restricted family of n-qubit quantum circuits $\mathcal{F}_n$ is insufficient grounds for ruling out a quantum advantage based on $\mathcal{F}_n$. This is because estimating output probabilities is only one particular way to post-process outputs obtained from a quantum computer. In short, while Definition 1.2 also defines a clean computational task, its significance is more nuanced than the sampling task defined in Definition 1.1.

The notions of classical simulation and quantum advantage that we have discussed thus far are motivated by computational complexity theory and are entirely asymptotic. There is a related field devoted to simulating finite-sized quantum circuits in software. While the discovery of a novel $O(n^{1000})$ time algorithm could constitute a theoretical breakthrough, the algorithm will likely suffer in practice due to the large overhead [33]. That being said, the announcement of a new asymptotic result may inspire followup research into more practical variants [34]. In the opposite direction, a more sophisticated implementation of the algorithms underlying the Gottesman-Knill theorem, such as Stim [35], could be highly valuable in practice with numerous applications, albeit having limited theoretical implications. While the primary motivation of the work in this thesis is to advance our understanding of asymptotic classical simulation and quantum advantage, software packages like Stim have been indispensable tools for a large portion of the research presented in Chapters 3 and 4. Consequently, it is fair to say that advances in numerical simulation software are likewise contributing indirectly to our understanding of quantum computational complexity.

1.2 Random quantum circuits

A random quantum circuit is a quantum circuit whose topological structure is specified by a circuit architecture while the actual unitary transformation applied by each gate is sampled independently and identically according to some probability distribution. If the architecture is made up of k-qubit gates, then the distribution governing each random gate draws from the group of all 2^k-by-2^k unitary matrices. The joint randomness over the individual k-qubit gates of a random quantum circuit induces a probability distribution over quantum circuits, which we refer to as a random circuit distribution. The circuit architecture dictates all the deterministic aspects of a random circuit distribution, including the lightcone structure and the circuit depth. In this thesis, we will focus on geometrically local circuit architectures; for example, the 1D and 2D brickwork architectures that will be described in Section 1.3.3, which are made up of 2-qubit gates.

The gate distributions we consider will either be the continuous uniform distribution over k-qubit unitary matrices, more commonly referred to as the Haar distribution in this context, or the discrete uniform distribution over the k-qubit Clifford group, with an emphasis on the latter. For every integer $k \geq 1$, we will let $\mathcal{H}_k$ denote the k-qubit Haar distribution and let $\mathcal{C}_k$ denote the uniform distribution over the k-qubit Clifford group. In a Haar random quantum circuit, every gate is sampled from $\mathcal{H}_k$, and in a random Clifford circuit, each gate is drawn according to $\mathcal{C}_k$. Let $n \geq 1$ be an integer. For the special n-qubit random quantum circuit U consisting of a single n-qubit gate, we call U a global Haar random unitary if U is drawn from $\mathcal{H}_n$ and a global uniformly random Clifford unitary if U is drawn from $\mathcal{C}_n$. For $U \sim \mathcal{H}_n$, we call $U|0^n\rangle$ a Haar random state, and for $U \sim \mathcal{C}_n$, we call $U|0^n\rangle$ a uniformly random stabilizer state. The Haar distribution $\mathcal{H}_n$ can be characterized by a unitary invariance property, which states that for an arbitrary n-qubit unitary V, if $U \sim \mathcal{H}_n$ is a random variable drawn from $\mathcal{H}_n$, then VU and UV are both random variables distributed according to $\mathcal{H}_n$. An analogous unitary invariance property with V being an arbitrary n-qubit Clifford unitary also characterizes $\mathcal{C}_n$.

Early research since the 1990s began with understanding and applying global Haar random unitaries, Haar random states, and their random Clifford counterparts. For example, the distribution of bipartite entanglement in uniformly random stabilizer states was fully solved in [36, 37]. Early research has demonstrated that "applying a Haar random unitary" could be a useful operation to perform in algorithmic settings [38]. It has also been used as a proof technique in the sense of the probabilistic method: proving non-constructively the existence of unitaries satisfying some desired property by showing that Haar random unitaries satisfy the property with a strictly positive probability or in expectation [39].

The biggest caveat of global Haar random unitaries in algorithmic applications is that they cannot be constructed efficiently on a quantum computer [30, 40]. I believe [41] was the first work that suggested and conjectured that sufficiently deep Haar random quantum circuits could serve as efficiently implementable replacements for global Haar random unitaries whenever the latter are used in an algorithmic protocol. A major line of research pioneered by [42] has proven this conjecture by showing that the statistics produced by poly(n)-depth Haar random quantum circuits, in a very generic sense and for more and more architectures, well-approximate those of global Haar random unitaries [43–49]. More concretely, consider the following definition.

Definition 1.3 (Operator norm ε-approximate unitary t-design)**.** For $\varepsilon \geq 0$ and integer $t \geq 1$, we say an n-qubit random circuit distribution $\mathcal{D}$ forms an ε-approximate unitary t-design if

$$\left\| \mathop{\mathbb{E}}_{U\sim\mathcal{D}} [U^{\otimes t} \otimes (\overline{U})^{\otimes t}] - \mathop{\mathbb{E}}_{U\sim\mathcal{H}_n} [U^{\otimes t} \otimes (\overline{U})^{\otimes t}] \right\|_{\text{op}} \leq \varepsilon \tag{1.1}$$

where $\overline{U}$ is the complex conjugate of U and $\|\cdot\|_{\text{op}}$ is the operator norm.

Definition 1.3 is one way to define an ε-approximate unitary t-design, which suits the discussions of this section. See [49] for a comparison between Definition 1.3 and several alternative definitions of ε-approximate unitary t-design that have been considered in the literature.

From the vectorization isomorphism (see section 1.1.2 of [50]) and using the Cauchy-Schwarz inequality, we obtain the following specialized but more convenient condition.

Corollary 1.1. *If an n-qubit random circuit distribution $\mathcal{D}$ satisfies Definition 1.3, then for every 2^{nt}-by-2^{nt} complex matrix A and B with $\|A\|_F = \|B\|_F = 1$ where $\|\cdot\|_F$ is the Frobenius norm, it holds that*

$$\left| \mathop{\mathbb{E}}_{U\sim\mathcal{D}} [\text{Tr}(BU^{\otimes t} A (U^\dagger)^{\otimes t})] - \mathop{\mathbb{E}}_{U\sim\mathcal{H}_n} [\text{Tr}(BU^{\otimes t} A (U^\dagger)^{\otimes t})] \right| \leq \varepsilon. \tag{1.2}$$

Expressions of the form $\mathbb{E}_{U\sim\mathcal{D}}[\text{Tr}(BU^{\otimes t} A (U^\dagger)^{\otimes t})]$ will appear repeatedly throughout this thesis; in fact, we will only ever encounter special cases with $t = 2$ and $A = |0^{2n}\rangle\langle 0^{2n}|$. As an example, the following theorem has been proven for 1D brickwork Haar random quantum circuits.

Theorem 1.1 ([43, 46])**.** *The distribution of n-qubit 1D brickwork Haar random quantum circuits forms an ε-approximate unitary 2-design in depth $O(\ln(1/\varepsilon))$.*

Due to the favourable scaling in ε, we can achieve an extremely small error, say $\varepsilon = 1/100^n$, in $O(n)$ depth. Suppose we take $A = |0^{2n}\rangle\langle 0^{2n}|$ and $B = \frac{1}{2^n} Z(s) \otimes Z(s)$ for any $s \in \{0,1\}^n \setminus \{0^n\}$ where $Z(s) = \bigotimes_{j=1}^n Z^{s_j}$. It is a well-known fact that (see Lemma 1.2)

$$\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n} [\langle 0^n | U^\dagger Z(s) U | 0^n\rangle^2] = \frac{1}{2^n + 1}. \tag{1.3}$$

Letting $\mathcal{D}_{n,d}$ denote the distribution of n-qubit depth-d 1D brickwork Haar random quantum circuits, Theorem 1.1 implies that $\mathcal{D}_{n,d}$ can achieve

$$\frac{1}{2^n+1} - \frac{1}{100^n} \leq \mathop{\mathbb{E}}_{U\sim\mathcal{D}_{n,d}} [\langle 0^n | U^\dagger Z(s) U | 0^n\rangle^2] \leq \frac{1}{2^n+1} + \frac{1}{100^n} \tag{1.4}$$

in depth linear in n.

It is time to introduce the motivating application of random quantum circuits for this thesis, which is not inherited from the usefulness of global Haar random unitaries. It is the prospect of using random circuit sampling, the task of sampling from the output distributions of Haar random quantum circuits in the sense of Definition 1.1, to establish a quantum advantage [51, 52]. It has been conjectured that if random circuit sampling, in the sense of Definition 1.1, with poly(n)-depth Haar random quantum circuits can be performed

efficiently on a classical computer, then the polynomial hierarchy collapses [32, 53]. A long line of research has tried to prove this conjecture, but it remains unproven (see [31, 54, 55] for the state-of-the-art in this pursuit). Since there is a widely held belief in computational complexity theory that the polynomial hierarchy does not collapse, this conjecture, if proven, would constitute strong complexity theoretic evidence that quantum computation is more powerful than classical computation.

In the past few years, we have witnessed a growing interest in understanding the potential utility of low-depth random quantum circuits, which are random circuits of depth at most $O(\log n)$. Recent breakthroughs in this direction have led to the realization that $\Theta(\log n)$-depth random quantum circuits can approximate many aspects of global Haar random unitaries far better than previously thought [56–60]. To explain the recent developments more concretely, consider the following definition.

Definition 1.4 (Collision probability)**.** For an n-qubit random circuit distribution $\mathcal{D}$, the (expected) collision probability of $\mathcal{D}$ is

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}\left[\Pr_{x,y\sim q}(x=y)\right] \tag{1.5}$$

where $q(z) = |\langle z|U|0^n\rangle|^2$, $z \in \{0,1\}^n$ is the computational-basis output distribution of U and x and y are drawn independently from q.

We can rewrite (1.5) using the definition of q as

$$(1.5) = \mathop{\mathbb{E}}_{U\sim\mathcal{D}}\left[\sum_{x\in\{0,1\}^n}|\langle x|U|0^n\rangle|^4\right] = \mathop{\mathbb{E}}_{U\sim\mathcal{D}}\left[\langle 0^{2n}|(U^\dagger\otimes U^\dagger)O(U\otimes U)|0^{2n}\rangle\right] \tag{1.6}$$

for the Hermitian positive semidefinite observable

$$O = \sum_{x\in\{0,1\}^n}|x\rangle\langle x|\otimes|x\rangle\langle x|. \tag{1.7}$$

Thus, we see that (1.5) is an expression of the form $\mathbb{E}_{U\sim\mathcal{D}}[\mathrm{Tr}(BU^{\otimes t}A(U^\dagger)^{\otimes t})]$ for $t=2$ with $A = |0^{2n}\rangle\langle 0^{2n}|$ and $B = O$. Using the identity

$$\sum_{x\in\{0,1\}^n}|x\rangle\langle x|\otimes|x\rangle\langle x| = \frac{1}{2^n}\sum_{s\in\{0,1\}^n}Z(s)\otimes Z(s) \tag{1.8}$$

together with (1.3) and the linearity of expectation, we find that the collision probability of the Haar distribution is

$$\frac{1}{2^n}\sum_{s\in\{0,1\}^n}\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n}[\langle 0^n|U^\dagger Z(s)U|0^n\rangle^2] = \frac{1}{2^n}+\frac{2^n-1}{2^n(2^n+1)} = \frac{2}{2^n+1}. \tag{1.9}$$

Now, we are ready to define an important property called anti-concentration.

Definition 1.5 (Anti-concentration)**.** We say an n-qubit random circuit distribution $\mathcal{D}$ satisfies the anti-concentration property if the (expected) collision probability of $\mathcal{D}$ (Definition 1.4) can be upper bounded by $c/2^n$ for some constant $c > 0$ independent of n.

From (1.9), we see that Haar random unitaries have the anti-concentration property, and it is straightforward to deduce from (1.4) and (1.9) that 1D brickwork Haar random quantum circuits with depth $O(n)$ satisfy the anti-concentration property. Anti-concentration has been featured prominently as a key ingredient in the leading proof strategy for proving the classical hardness of random circuit sampling [32, 53], but more recently, it is also required for establishing the $\mathrm{poly}(n)$ runtime of a classical algorithm for simulating random circuit sampling with constant-rate depolarizing noise [33]. This dual use has led researchers to rethink whether the anti-concentration property is truly necessary for the classical hardness of random circuit sampling, or if it is just a proof technique.

Before the works of [56, 57], a common belief was that geometrically local D-dimensional Haar random quantum circuits required $\Omega(n^{1/D})$ depth to achieve anti-concentration [44]. However, the surprising results of [56, 57] proved that $O(\log n)$ depth is already sufficient for anti-concentration for the 1D brickwork architecture, while $\Omega(\log n)$ depth is necessary for any architecture. More specifically, again letting $\mathcal{D}_{n,d}$ denote the distribution of n-qubit depth-d 1D brickwork Haar random quantum circuits, the results from [56, 57] imply that the collision probability of $\mathcal{D}_{n,d}$ satisfies the lower and upper bounds

$$\frac{1}{2^n}\left(1+\frac{1}{3}\left(\frac{2}{5}\right)^d\right)^n \leq \mathop{\mathbb{E}}_{U\sim\mathcal{D}_{n,d}}\left[\sum_{x\in\{0,1\}^n}|\langle x|U|0^n\rangle|^4\right] \leq \frac{2}{2^n+1}\left(1+\left(\frac{4}{5}\right)^{d-1}\right)^n. \tag{1.10}$$

We remark that the lower bound in (1.10) holds for an arbitrary circuit architecture composed of 2-qubit gates, not just $\mathcal{D}_{n,d}$. Also, the upper bound shows that the collision probability of $\mathcal{D}_{n,d}$ is approaching that of global Haar random unitaries from above in relative error, which can be made $1/\mathrm{poly}(n)$ by taking $d = O(\log n)$.

Then, the recent breakthrough results of [58,59] could be interpreted as vastly generalizing the relative error convergence in (1.10) with the single observable (1.7) to all Hermitian positive semidefinite observables. Their results imply that for all $\varepsilon = 1/\mathrm{poly}(n)$ and constant $t \geq 1$, it suffices to choose $d = O(\log n)$ depending on ε and t to obtain

$$(1-\varepsilon)\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n}[\mathrm{Tr}(BU^{\otimes t}A(U^\dagger)^{\otimes t})] \leq \mathop{\mathbb{E}}_{U\sim\mathcal{D}_{n,d}}[\mathrm{Tr}(BU^{\otimes t}A(U^\dagger)^{\otimes t})] \leq (1+\varepsilon)\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n}[\mathrm{Tr}(BU^{\otimes t}A(U^\dagger)^{\otimes t})] \tag{1.11}$$

holding for all 2^{nt}-by-2^{nt} Hermitian positive semidefinite matrices A and B. Strictly speaking, this result was established for another architecture called the 1D coarse-grained architecture, which we will introduce and discuss in Chapter 3. The distinction between the two architectures is unimportant for now.

To the best of my knowledge, [5] was the first work that seriously studied constant-depth random quantum circuits. From the lower bound in (1.10), we know for a fact that they cannot serve as potent generic replacements for global Haar random unitaries because even the collision probability of constant-depth random quantum circuits behaves very differently from that of global Haar random unitaries. Then what could constant-depth random quantum circuits be useful for? Napp et al. investigated the computational complexity of classically simulating constant-depth 2D brickwork Haar random quantum circuits in the sense of Definition 1.1 [5]. They arrived at a very counter-intuitive observation that the task could become intractable for classical computers when the circuit depth exceeds some constant critical threshold d^* [5]. Logically, this implies that there is a phase transition in

the complexity of classically simulating constant-depth 2D brickwork Haar random quantum circuits because depth-1 circuits can obviously be efficiently classically simulated, while at the same time, the difficulty should be monotonically non-decreasing in the circuit depth. I consider this a very surprising and bold conjecture. If it were true, then it would clarify that the anti-concentration property is not a fundamental prerequisite for the classical hardness of random circuit sampling.

What justification did Napp et al. provide for their strong conjecture? In the paper [5], they attributed the conjectured phase transition in the hardness of classical simulation to another conjecture regarding the emergence of long-range measurement-induced entanglement (MIE) in constant-depth 2D Haar random quantum circuits after the same critical depth d^*. We will define MIE more concretely in Section 1.3.5. The work [5] provided extensive numerical and analytical evidence supporting their conjectures, but their arguments were not fully rigorous. A few subsequent works further explored the conjectured long-range MIE phase transition in various settings, both numerically and analytically [61, 62], and even experimentally on a 70-qubit superconducting quantum processor [63]. Our paper [2], which Chapter 3 is based on, was the first to rigorously prove the emergence of long-range MIE in 2D Haar random and random Clifford circuits. Unfortunately, we fell short of proving it for a truly constant-depth architecture as it is more accurate to view the 2D coarse-grained architecture we considered to be effectively $O(\log n)$ depth. Despite considerable interest, rigorously proving the existence of long-range MIE in some truly constant-depth architecture of 2D random quantum circuits remained an open problem until the breakthrough work of [64]. We will discuss the main results of [64] more concretely in Section 3.5.

1.3 Preliminaries

1.3.1 Notations

Let $n \geq 1$ be an integer. We adopt the notation $[n] = \{1, 2, \ldots, n\}$. Whenever we have a set of n qubits, the qubits will be labelled 1 through n, so we will use $[n]$ to denote the set of n qubits. Sometimes, we need to partition the n qubits into two or more disjoint subsets; for three subsets, we usually write $[n] = A \cup B \cup C$. In this context, the union is always understood to be disjoint. For every integer $m \geq 1$, let $\mathcal{P}^m = \{I, X, Y, Z\}^{\otimes m}$ denote the set of all m-qubit Pauli operators with $+$ sign. For every set S consisting of Pauli operators, we define $\pm S = S \cup \{-P : P \in S\}$.

For a single-qubit operator O and $s \in \{0, 1\}^n$, we define $O(s) = \bigotimes_{j=1}^n O^{s_j}$ to be the n-qubit operator that applies O to every qubit $j \in [n]$ such that $s_j = 1$ and the identity to all other qubits. In most cases, O will be a 1-qubit Pauli operator X, Y, or Z. We say an n-qubit Pauli operator P with $\pm$ sign is Z-type if $P = Z(s)$ or $P = -Z(s)$ for some $s \in \{0, 1\}^n$. For the projectors onto 1-qubit computational basis states $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$, we instead use $(|0\rangle\langle 0|)_A$ and $(|1\rangle\langle 1|)_A$ to denote applying the projector to each qubit in a subset $A \subseteq [n]$ of qubits and the identity to all the qubits in $[n] \setminus A$.

For a subset $A \subseteq [n]$ and a $|A|$-qubit operator O, we use O_A to denote the n-qubit operator that applies O to the qubits in A and the identity to the qubits in $[n] \setminus A$. For $A = \{j\}$ of size one, we define $O_j = O_{\{j\}}$. Note that for a single-qubit operator O and

$s \in \{0,1\}^n$, we get $O(s) = \prod_{j=1}^n O_j^{s_j}$. Strictly speaking, the "O_A notation" is ambiguous for at least two reasons. Firstly, unlike $s \in \{0,1\}^n$, $A \subseteq [n]$ does not uniquely specify n, so the dimension of O_A is ambiguous. Secondly, A should be an ordered list as opposed to an unordered set. In practice, the dimension of O_A can always be inferred from the context. Throughout this thesis, whenever we take a subset of qubits $A \subseteq [n]$, we will always assume that the qubits in A are ordered from smallest to largest so that the first qubit of A is the minimum element of A.

We adopt the following notation to express pure product states. Suppose $[n] = A \cup B$. For a $|A|$-qubit pure state $|\alpha\rangle$ and a $|B|$-qubit pure state $|\beta\rangle$, we use the notation $|\alpha\rangle_A \otimes |\beta\rangle_B = |\alpha\rangle_A |\beta\rangle_B$ to denote the n-qubit pure product state that is $|\alpha\rangle$ on A and $|\beta\rangle$ on B. We write $|\alpha\rangle_A$ alone to denote the tensor that is $|\alpha\rangle$ on A and the identity on all the other qubits. We use analogous notations to denote the conjugate transposes of pure product states.

1.3.2 Specification of quantum circuits

In this thesis, we work extensively with quantum circuits. Let $n, d \geq 1$ be integers. An n-qubit depth-d quantum circuit U is an n-qubit unitary of the form $U = V_d \cdots V_2 V_1$ where for every $i \in [d]$, V_i is an n-qubit depth-1 quantum circuit, and we call V_i the i-th layer of U. An n-qubit depth-1 quantum circuit W can be specified by a set of gates $\{g_1, \ldots, g_s\}$ where for every $i \in [s]$, the gate g_i can be specified by a subset of qubit indices $\mathrm{Supp}(g_i) \subseteq [n]$ and a k_i-qubit unitary W_i where $k_i = |\mathrm{Supp}(g_i)|$. It is required that for every $i, j \in [s]$, if $i \neq j$, then $\mathrm{Supp}(g_i) \cap \mathrm{Supp}(g_j) = \emptyset$, so the gates $g_1, \ldots, g_s$ act on disjoint subsets of qubits. Using the "O_A notation" defined earlier, we get that $W = \prod_{i=1}^s (W_i)_{\mathrm{Supp}(g_i)}$. Most of the time, all the gates in W will act on the same number of qubits, implying $k_1 = \cdots = k_s$. If for some integer $k \geq 1$, W consists solely of k-qubit gates, then by the disjointness requirement, W can contain at most $s \leq \lfloor \frac{n}{k} \rfloor$ gates. Most of the quantum circuits that will be considered in this thesis will be composed entirely of 2-qubit gates, in which case each depth-1 circuit can contain at most $\lfloor \frac{n}{2} \rfloor$ gates. By a slight abuse of notation, we will refer to a gate g_i in W by just its unitary W_i when the support is clear from context, and we will write $\mathrm{Supp}(W_i)$ in place of $\mathrm{Supp}(g_i)$. With this, we define $\mathrm{Supp}(W) = \bigcup_{i=1}^s \mathrm{Supp}(W_i)$.

We will also use the Supp() notation in another context. For an n-qubit Pauli operator $P \in \pm\mathcal{P}^n$, we define $\mathrm{Supp}(P) \subseteq [n]$ to be the subset of qubits on which P takes a value other than the identity.

1.3.3 Circuit architecture

Let $n, d \geq 1$ be integers. We call an n-qubit depth-d quantum circuit U shallow if $d = O(1)$ is a constant independent of n and low-depth if $d \leq O(\log n)$. Note that for the above definition to make any sense, we really need to talk about a sequence of sets of circuits $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ where $\mathcal{F}_n$ is non-empty for infinitely many $n \in \mathbb{N}$ with $\mathbb{N}$ being the set of natural numbers. For example, there could exist an $N_0 \in \mathbb{N}$ such that $\mathcal{F}_n = \emptyset$ for all $n < N_0$. Then we say $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is shallow if there exists $d^* \in \mathbb{N}$ such that for every $n \in \mathbb{N}$ and every circuit $U \in \mathcal{F}_n$, the depth of U is at most d^*. Similarly, $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is low-depth if there exists $c > 0$ such that for every $n \in \mathbb{N}$ and every circuit $U \in \mathcal{F}_n$, the depth of U is at most $c \log_2(n)$. As such, whenever we say an individual quantum circuit is shallow or low-depth, we are always

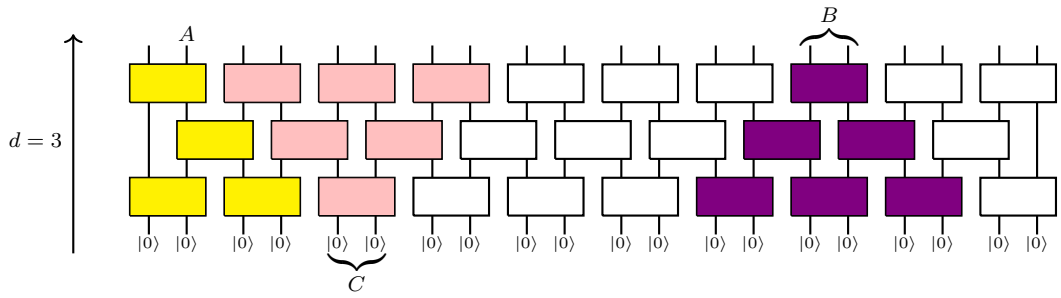


Figure 1.1: Circuit diagram of a 20-qubit depth-3 quantum circuit composed of 2-qubit gates. Circuit evolution time flows upwards. The gates are arranged according to the 1D brickwork architecture. By labelling the qubits 1 through 20 from left to right, the gates shaded in yellow constitute the backward lightcone circuit of $A = \{2\}$ while the gates shaded in violet form the backward lightcone circuit of $B = \{15, 16\}$. The gates shaded in pink correspond to the forward lightcone circuit of $C = \{5, 6\}$. Note that $\mathrm{blc}(A) = \{1, 2, 3, 4\}$ and $\mathrm{blc}(B) = \{13, 14, 15, 16, 17, 18\}$, so $\mathrm{blc}(A) \cap \mathrm{blc}(B) = \emptyset$.

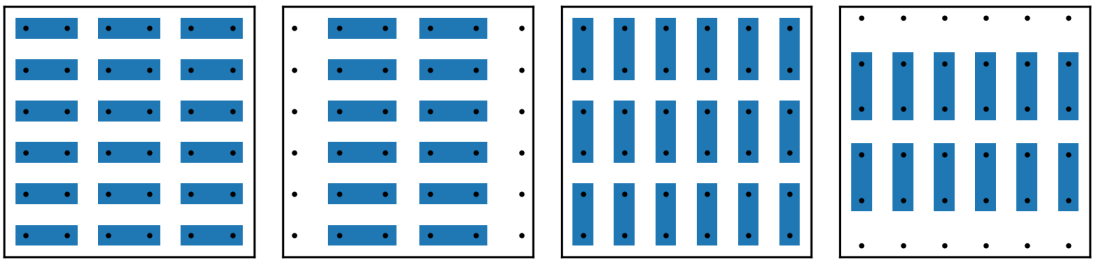


Figure 1.2: Gate patterns of the 2D brickwork architecture on a 6×6 grid. Each black dot represents a qubit, and each blue rectangle represents a 2-qubit gate acting on the qubits covered by the rectangle. The four panels from left to right depict the placement of 2-qubit gates in layers one through four. Circuit layer five cycles back to use the pattern of the first layer (leftmost panel), and so on.

implicitly thinking of a sequence of sets of circuits in the background. Most of the quantum circuits that we will encounter in this thesis will be shallow circuits.

Sometimes, there will be a well-defined sense of geometric locality present among the n qubits. For example, the n qubits may be placed on a horizontal 1D line so that we can call the left-most qubit the first qubit and the right-most qubit the last qubit, and the gates in a quantum circuit will respect the geometry by only ever acting on spatially neighbouring qubits. For a geometrically local quantum circuit, it is common for the supports of the gates to follow a fixed pattern called a circuit architecture. A circuit architecture dictates the topology of the circuit but not the actual unitary transformations applied by each gate. Figure 1.1 depicts the circuit diagram of a 20-qubit depth-3 quantum circuit with the canonical 1D brickwork architecture. Extending beyond 1D, we will also encounter situations where the n qubits are placed on a rectangular 2D grid or even D-dimensional hypercube for some constant $D \geq 3$. Figure 1.2 shows the gate placements of the first

four layers of the 2D brickwork architecture, which is one possible generalization of the 1D brickwork architecture to 2D, with 36 qubits placed on a 6×6 grid.

1.3.4 Lightcones and lightcone separation

Let $n, d \geq 1$ be integers, and let $U = V_d \cdots V_1$ be an n-qubit depth-d quantum circuit. We call the pure state $|\psi\rangle = U|0^n\rangle$ the output state of the circuit U. Let $\emptyset \neq A \subseteq [n]$. The reduced state of $|\psi\rangle$ on A is the mixed state $\rho_A = \mathrm{Tr}_{[n]\setminus A}(|\psi\rangle\langle\psi|)$. By the cyclic property of the partial trace, observe that

$$\begin{aligned} \mathrm{Tr}_{[n]\setminus A}(|\psi\rangle\langle\psi|) &= \mathrm{Tr}_{[n]\setminus A}(V_d \cdots V_1 |0^n\rangle\langle 0^n| V_1^\dagger \cdots V_d^\dagger) && (1.12) \\ &= \mathrm{Tr}_{[n]\setminus A}(V_d' \cdots V_1' |0^n\rangle\langle 0^n| (V_1')^\dagger \cdots (V_d')^\dagger) && (1.13) \\ &= \mathrm{Tr}_{[n]\setminus A}(U' |0^n\rangle\langle 0^n| (U')^\dagger) && (1.14) \end{aligned}$$

for an n-qubit depth-d quantum circuit $U' = V_d' \cdots V_1'$ such that for every $i \in [d]$, V_i' contains only the gates in V_i that cannot be commuted through the partial trace and cancelled. We call the circuit U' the backward lightcone circuit of A and the union of the supports of all the gates in U' the backward lightcone of A which we denote by $\mathrm{blc}(A)$.

Given U and A, we could find U' and $\mathrm{blc}(A)$ using the following inductive procedure. We see that V_d' contains all the gates in V_d that act on a qubit in A. Define $A_d' = \bigcup_{g \in V_d'} \mathrm{Supp}(g)$. Then for every $i \in [d-1]$, V_i' consists of all the gates of V_i that act on a qubit in A_{i+1}' and define $A_i' = A_{i+1}' \cup \left(\bigcup_{g \in V_i'} \mathrm{Supp}(g)\right)$. Finally, we have $\mathrm{blc}(A) = A_1'$. We could also run the above procedure in reverse to find what are called the forward lightcone of A, which we denote by $\mathrm{flc}(A)$, along with the forward lightcone circuit of A. We start by defining V_1'' to be the set of all gates in V_1 that act on a qubit in A and $A_1'' = \bigcup_{g \in V_1''} \mathrm{Supp}(g)$. Then for every $i \in \{2, \ldots, d\}$, define V_i'' to be the set of all gates in V_i that act on a qubit in A_{i-1}'' and $A_i'' = A_{i-1}'' \cup \left(\bigcup_{g \in V_i''} \mathrm{Supp}(g)\right)$. By the end, we have $\mathrm{flc}(A) = A_d''$ and $V_d'' \cdots V_1''$ is the forward lightcone circuit of A. Unlike how we have defined $\mathrm{blc}(A)$ based on the reduced state ρ_A, we are treating $\mathrm{flc}(A)$ as being defined by the above procedure when applied to A.

In Figure 1.1, the backward lightcone circuits of $A = \{2\}$ and $B = \{15, 16\}$ and the forward lightcone circuit of $C = \{5, 6\}$ are shaded in yellow, violet, and pink respectively. For a single qubit $i \in [n]$, we adopt the notational simplification $\mathrm{blc}(i) = \mathrm{blc}(\{i\})$ and $\mathrm{flc}(i) = \mathrm{flc}(\{i\})$.

For two subsets of qubits $A, B \subseteq [n]$, we say A and B are lightcone-separated if $\mathrm{blc}(A) \cap \mathrm{blc}(B) = \emptyset$. We see that for two lightcone-separated subsets of qubits $A, B \subseteq [n]$, by the cyclic property of the partial trace, the reduced state of $A \cup B$ satisfies $\rho_{A\cup B} = \rho_A \otimes \rho_B$, so $\rho_{A\cup B}$ is a product state with no entanglement between A and B. In Figure 1.1, the subsets $A = \{2\}$ and $B = \{15, 16\}$ are lightcone-separated. The following lemma summarizes the d-dependence of the backward lightcones of the individual qubits of a quantum circuit.

Lemma 1.1. *Let U be an n-qubit depth-d quantum circuit composed of k-qubit gates. Let $A = \{i\}$ contain a single qubit $i \in [n]$. Then it holds that $|\mathrm{blc}(A)| \leq k^d$, and A is lightcone-separated from all but at most k^{2d} other qubits.*

Proof. Evidently $|A'_d| \leq k$ and it is easy to see that for every $i \in [d-1]$, it holds that $|A'_i| \leq k|A'_{i+1}|$. Therefore, $|\mathrm{blc}(A)| = |A'_1| \leq k^d$. To prove the second claim, the key observation is that for every $j \in [n]$, $j \in \mathrm{blc}(i)$ if and only if $i \in \mathrm{flc}(j)$. For the forward lightcone, an analogous recursion shows that for every $j \in [n]$, $|\mathrm{flc}(j)| \leq k^d$. Therefore, we deduce the upper bound

$$|\{i' \in [n] : \mathrm{blc}(i) \cap \mathrm{blc}(i') \neq \emptyset\}| \leq \sum_{j \in \mathrm{blc}(i)} |\mathrm{flc}(j)| \leq (k^d)(k^d) = k^{2d}. \tag{1.15}$$

□

For shallow circuits with a constant circuit depth d, the above lemma shows that the backward lightcone of every qubit has a constant size and that every qubit is lightcone-separated from all but a constant number of other qubits. For geometrically local quantum circuits in some $D \geq 1$ dimensions with k-qubit gates, the upper bound on $|\mathrm{blc}(i)|$ for $i \in [n]$ can be improved to $(kd)^D$, which is polynomial in d. This distinction in scaling makes a qualitative difference when $d = \Theta(\log n)$. Without geometric locality, the backward lightcone of each qubit could include all the qubits, yet for a D-dimensional circuit, the upper bound $|\mathrm{blc}(i)| \leq O((\log n)^D)$ for every qubit $i \in [n]$ remains non-trivial.

For a subset $A \subseteq [n]$, we call the distribution $P_A(x) = \langle 0^n | U^\dagger (|x\rangle\langle x|)_A U |0^n\rangle$, $x \in \{0,1\}^{|A|}$ (recall the "O_A notation") the marginal output distribution of A. Note that we can write $P_A(x) = \mathrm{Tr}((|x\rangle\langle x|)_A |\psi\rangle\langle\psi|) = \mathrm{Tr}(|x\rangle\langle x|\rho_A) = \langle x|\rho_A|x\rangle$, so gates in U that are outside the backward lightcone circuit of A do not influence P_A.

1.3.5 Measurement-induced entanglement (MIE)

Let $n \geq 3$ be an integer. Suppose we partition $[n]$ into three non-empty and disjoint subsets $[n] = A \cup B \cup C$. Suppose we have an n-qubit pure state $|\psi\rangle$ whose reduced state on $A \cup C$ is a product state, which means $\mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \rho_A \otimes \rho_C$ for some $|A|$-qubit state ρ_A and $|C|$-qubit state ρ_C. In this case, there is no entanglement between A and C in $|\psi\rangle$. Now, instead of tracing out B, consider measuring all qubits of $|\psi\rangle$ in B in the computational basis and suppose it results in an outcome $x \in \{0,1\}^{|B|}$ and a pure post-measurement state $|\psi'\rangle$ on $A \cup C$. We say there is measurement-induced entanglement (MIE) in $|\psi'\rangle$ if $|\psi'\rangle$ is entangled across the (A, C)-bipartition. For a concrete example, consider the 4-qubit circuit diagram describing entanglement swapping, which is an application of quantum teleportation, shown in Figure 1.3. In this example, the middle two qubits are B, and the state immediately before the measurement is

$$|\psi\rangle = \frac{1}{2} \sum_{x \in \{0,1\}^2} |x\rangle_B \otimes |\Phi_x\rangle_{AC} \tag{1.16}$$

where $\{|\Phi_{00}\rangle, |\Phi_{01}\rangle, |\Phi_{10}\rangle, |\Phi_{11}\rangle\}$ is the Bell basis. If we trace out B, the reduced state on $A \cup C$ is a product of single-qubit maximally mixed states, so $\mathrm{Tr}_B(|\psi\rangle\langle\psi|)$ is unentangled. However, if we perform the measurement and observe an outcome $x \in \{0,1\}^2$, then MIE is present in the post-measurement state, which is the maximally entangled state $|\Phi_x\rangle$ on $A \cup C$.

Next, we briefly explain how MIE is reflected in the stabilizers of a stabilizer state. Let $|\psi\rangle$ be an n-qubit stabilizer state where $[n]$ is partitioned into three disjoint subsets

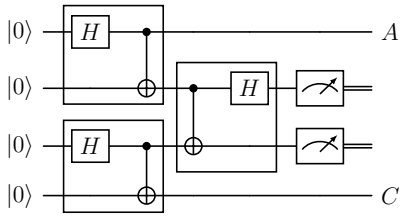


Figure 1.3: Circuit diagram for entanglement swapping (without local Pauli corrections). The three rectangular boxes group the gates for preparing a Bell state.

$[n] = A \cup B \cup C$. Let $S_\psi = \{P \in \pm\mathcal{P}^n : P|\psi\rangle = |\psi\rangle\}$ denote the stabilizer group of $|\psi\rangle$. Suppose we measure the qubits of $|\psi\rangle$ in B in the computational basis and obtain an outcome $x \in \{0,1\}^{|B|}$ and post-measurement state $|\psi'\rangle$ on $A \cup C$. Since the qubits in B are measured in the computational basis and the outcome is x, the stabilizer group of $|\psi'\rangle$ is

$$S_{\psi'} = \{(-1)^{s\cdot x} Q : Q \in \pm\mathcal{P}^{|A|+|C|} \text{ and } Q_{AC} Z(s)_B \in S_\psi \text{ for some } s \in \{0,1\}^{|B|}\}. \tag{1.17}$$

Let

$$S_{\psi'_A} = \{P \in S_{\psi'} : P = Q \otimes I^{\otimes|C|} \text{ for some } Q \in \pm\mathcal{P}^{|A|}\} \tag{1.18}$$

denote the subgroup of $S_{\psi'}$ local on A. In the stabilizer setting, there is bipartite entanglement between A and C in $|\psi'\rangle$ if and only if $|S_{\psi'_A}| < 2^{|A|}$, and the number of ebits of entanglement is exactly $\log_2(2^{|A|}/|S_{\psi'_A}|)$. Notice that the presence or absence of MIE in $|\psi'\rangle$ does not depend on the measurement outcome. We say a stabilizer $P \in S_{\psi'}$ is non-local if P cannot be written as $P = QR$ for some $Q \in S_{\psi'_A}$ and $R \in S_{\psi'_C}$ where $S_{\psi'_C}$ is the subgroup of $S_{\psi'}$ local on C. In terms of the non-local stabilizers, there is entanglement between A and C in $|\psi'\rangle$ if and only if $|\psi'\rangle$ has a non-local stabilizer.

1.3.6 Random Clifford circuits

Recall that for every integer $k \geq 1$, we use $\mathcal{H}_k$ to denote the k-qubit Haar distribution and $\mathcal{C}_k$ to denote the uniform distribution over the k-qubit Clifford group. Let $n \geq 1$ be an integer.

In Chapters 3 and 4 of this thesis, we aim to develop a better understanding of the statistical properties of the output states and the output distributions of low-depth random Clifford circuits that are most relevant to random circuit sampling. The rationales for focusing on the Clifford case are as follows. Firstly, the types of phenomena we wish to understand (the emergence of long-range MIE in a quantum state and long-range correlation in its output distribution) are already captured by the stabilizer formalism, so the Clifford analogs of these problems are non-trivial and could potentially inform the Haar random case. Secondly, restricting to the Clifford setting enables us to investigate these questions numerically because Clifford circuits with thousands of qubits or more can be simulated in software sufficiently quickly using modern simulators like Stim [35]. Moreover, it is a well-known fact that random Clifford unitaries form an exact unitary 2-design[2] with $\varepsilon = 0$ [67, 68], which

[2] In fact, it has been shown that $\mathcal{C}_k$ forms an exact unitary 3-design [65, 66].

more concretely means that for every $k \geq 1$, it holds that

$$\mathop{\mathbb{E}}_{U\sim\mathcal{C}_k}[U^{\otimes 2}\otimes(\overline{U})^{\otimes 2}] = \mathop{\mathbb{E}}_{U\sim\mathcal{H}_k}[U^{\otimes 2}\otimes(\overline{U})^{\otimes 2}], \tag{1.19}$$

and this extends directly to random Clifford circuits composed of random k-qubit Clifford gates. As many of our problems will boil down to analyzing expressions of the form $\mathbb{E}_{U\sim\mathcal{D}}[\mathrm{Tr}(BU^{\otimes 2}A(U^\dagger)^{\otimes 2})]$ for some n-qubit random circuit distribution $\mathcal{D}$ and 2^{2n}-by-2^{2n} matrices A and B, the answers will be exactly the same whether $\mathcal{D}$ is a distribution of random quantum circuits composed of Haar random gates or random Clifford gates. In these cases, performing the analysis with respect to random Clifford circuits constitutes a valid proof technique. Lastly, since a stabilizer state is fully characterized by its stabilizer group, which is a finite set, many things we do will have a clear combinatorial flavour. For example, recall that the collision probability (Definition 1.4) of an n-qubit random circuit distribution $\mathcal{D}$ can be written as

$$\frac{1}{2^n}\mathop{\mathbb{E}}_{U\sim\mathcal{D}}\Big[\sum_{s\in\{0,1\}^n}\langle 0^n|U^\dagger Z(s)U|0^n\rangle^2\Big]. \tag{1.20}$$

When $\mathcal{D}$ is a distribution of random Clifford circuits, for $U\sim\mathcal{D}$, the output state $|\psi\rangle = U|0^n\rangle$ is a stabilizer state. Thus, for every $s\in\{0,1\}^n$, $\langle\psi|Z(s)|\psi\rangle^2$ equals 1 if $Z(s)$ or $-Z(s)$ is a stabilizer of $|\psi\rangle$ and 0 otherwise. As a result, the collision probability gains a combinatorial interpretation

$$(1.20) = \frac{1}{2^n}\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|S_\psi^Z|] \tag{1.21}$$

where we use S_ψ^Z to denote the Z-type stabilizer subgroup of $|\psi\rangle = U|0^n\rangle$. Subsequently, the anti-concentration property (Definition 1.5) can be interpreted as asking whether the output state $|\psi\rangle$ has only a constant number of Z-type stabilizers in total in expectation. The combinatorics of stabilizer groups will be a recurring theme of Chapters 3 and 4.

Lastly, we provide a proof of the following well-known identity using Clifford techniques.

Lemma 1.2. *For every $P\in\mathcal{P}^n\setminus\{I^{\otimes n}\}$, it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n}[\langle 0^n|U^\dagger PU|0^n\rangle^2] = \frac{1}{2^n+1}. \tag{1.22}$$

Proof. Let $P\in\mathcal{P}^n\setminus\{I^{\otimes n}\}$. By (1.19), we get

$$\mathop{\mathbb{E}}_{U\sim\mathcal{H}_n}[\langle 0^n|U^\dagger PU|0^n\rangle^2] = \mathop{\mathbb{E}}_{U\sim\mathcal{C}_n}[\langle 0^n|U^\dagger PU|0^n\rangle^2], \tag{1.23}$$

so we can switch to consider the Clifford case. By viewing the expression $\langle 0^n|U^\dagger PU|0^n\rangle^2$ in the Heisenberg picture, we see that it equals 1 if $U^\dagger PU$ is a Z-type Pauli operator and 0 otherwise. Therefore,

$$\mathop{\mathbb{E}}_{U\sim\mathcal{C}_n}[\langle 0^n|U^\dagger PU|0^n\rangle^2] = \Pr_{U\sim\mathcal{C}_n}\left(U^\dagger PU \text{ is } Z\text{-type}\right). \tag{1.24}$$

Next, we show that for $U\sim\mathcal{C}_n$, $U^\dagger PU$ is a uniformly random non-identity n-qubit Pauli operator with $\pm$ sign. Let $Q, Q'\in\pm(\mathcal{P}^n\setminus\{I^{\otimes n}\})$. Then there exists an n-qubit Clifford

unitary V such that $VQV^\dagger = Q'$. By the unitary invariance property of $\mathcal{C}_n$, observe that

$$\Pr_{U\sim\mathcal{C}_n}(U^\dagger PU = Q) = \Pr_{U\sim\mathcal{C}_n}(P = UQU^\dagger) = \Pr_{U\sim\mathcal{C}_n}(P = (UV)Q(UV)^\dagger) = \Pr_{U\sim\mathcal{C}_n}(U^\dagger PU = Q'). \tag{1.25}$$

Since (1.25) holds for every Q and Q', we deduce that

$$\Pr_{U\sim\mathcal{C}_n}(U^\dagger PU = Q) = \frac{1}{|\pm(\mathcal{P}^n \setminus \{I^{\otimes n}\})|} = \frac{1}{2(4^n-1)}. \tag{1.26}$$

Therefore,

$$\Pr_{U\sim\mathcal{C}_n}\big(U^\dagger PU \text{ is } Z\text{-type}\big) = \frac{2^n-1}{4^n-1} = \frac{1}{2^n+1}. \tag{1.27}$$

□

Chapter 2

Classical simulation of peaked shallow quantum circuits

This chapter is based on the paper [1]. Throughout this chapter, all quantum circuits are assumed to be composed entirely of 2-qubit gates.

We say an n-qubit quantum circuit U is peaked if there exists an $x^* \in \{0,1\}^n$ such that $|\langle x^*|U|0^n\rangle|^2 \geq 1/\mathrm{poly}(n)$ and shallow if its circuit depth is a constant independent of n.[1] The peakedness condition ensures that if one runs a peaked quantum circuit U on a perfect quantum computer and samples from the output distribution $\mathrm{poly}(n)$ times, a peaked outcome x satisfying $|\langle x|U|0^n\rangle|^2 \geq 1/\mathrm{poly}(n)$ will be observed with high probability. This feature naturally suggests a following kind of quantum advantage proposal. A quantum challenger designs a quantum circuit with a known but obfuscated peaked outcome and first demonstrates that the peaked outcome could be found reliably using their quantum computer. Then, the quantum challenger publishes the description of the challenge circuit and challenges the public to solve for the special peaked string via classical simulation. A quantum advantage is declared and stands until the challenge circuit is successfully classically simulated. Approaches along this line were first conceptualized by Aaronson and Zhang [69, 70] and subsequently carried out on Quantinuum's experimental quantum hardware in 2025 [71]. A team from IBM reported successful numerical simulation of the largest challenge circuits of [71] in 2026 [72].

In this chapter, we consider the task of classically simulating peaked shallow (constant-depth) quantum circuits in the sense of Definition 1.1 with error tolerance $\varepsilon = 1/\mathrm{poly}(n)$. Note that if we drop the peakedness assumption, there is strong complexity theoretic evidence that exactly sampling ($\varepsilon = 0$) from the output distributions of arbitrary depth-3 quantum circuits should be classically intractable [5, 73, 74]. On the other hand, assuming peakedness without restricting the circuit depth is uninstructive, as sampling from the output distributions of arbitrary peaked quantum circuits is evidently BQP-hard.

As our main results, we present an algorithm that runs in quasipolynomial $n^{O(\log n)}$ time for classically simulating peaked shallow quantum circuits without assuming any geometric locality. We give a different algorithm that runs in polynomial $n^{O(1)}$ time for geometrically-local 2D peaked shallow circuits. We also have a third almost-polynomial $n^{O(\log\log n)}$ time algorithm for simulating peaked shallow circuits that are geometrically-local in D dimensions for any constant $D \geq 3$. Prior to our work, no non-trivial classical simulation algorithms

[1] Formally, the peaked and shallow conditions need to be defined on a sequence of sets of n-qubit circuits $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$; see Section 1.3.3.

were known for any of these tasks (general architecture, 2D, or 3D and higher).

Next, we describe the input-output behaviour of our algorithms more formally. For an integer $a \geq 1$, we say that an n-qubit quantum circuit U is a-peaked if

$$\max_{x \in \{0,1\}^n} |\langle x|U|0^n\rangle|^2 \geq n^{-a}. \tag{2.1}$$

The precise input-output relations and runtimes of our algorithms are stated in the following theorem.

Theorem 2.1 (**Classical simulation of peaked shallow quantum circuits**). *There exists a classical randomized algorithm which takes as input a positive integer $a \geq 1$, the classical description of an n-qubit depth-d quantum circuit U, and a positive integer $c \geq 1$ specifying an error tolerance $\varepsilon = n^{-c}$. If the algorithm succeeds, then it outputs a sample drawn from a probability distribution P' over n-bit strings such that $\|P' - P\|_1 \leq \varepsilon$ where P is the computational-basis output distribution of the input circuit U; otherwise, it outputs an error flag. The algorithm is guaranteed to succeed if U is a-peaked. The runtime of the algorithm is $n^{O(\log n)}$ where the $O()$ notation hides factors depending on a, c, and d.*

There exist a second and a third classical randomized algorithms with the same input-output behaviour as stated above except that the input circuit U is required to be geometrically-local in either 2D or D-dimensions for some $D \geq 3$. The runtime of the second algorithm is $n^{O(1)}$ where the $O()$ notation hides factors depending on a, c, and d, while the runtime of the third algorithm is $n^{O(\log\log n)}$ where the $O()$ notation hides factors depending on a, c, d, and D.

Some remarks about the algorithms' behaviour are in order. We are mainly interested in applying our algorithms to sequences of input instances of the form $\{(a, U_n, c)\}_{n\in\mathbb{N}}$ where each U_n is an n-qubit depth-d quantum circuit for some fixed d and a and c are fixed integers, so the runtimes are expressed to reflect the algorithms' scaling in n.[2] Suppose a user tries to run one of our algorithms on some n-qubit shallow circuit U. Note that the user does not need to know or supply a peaked outcome $x^* \in \{0,1\}^n$ such that $|\langle x^*|U|0^n\rangle|^2 \geq n^{-a}$. In fact, in order to guarantee success, the user only needs the knowledge that some constant integer $a \geq 1$ exists such that the circuit U is a-peaked and does not require knowing the value of a beforehand. The user could try to run the algorithm starting at $a = 1$ and increment the value of a by one each time the algorithm raises an error flag. The search will terminate once a large enough value of a has been reached, and the overall asymptotic runtime scaling of the whole procedure is the same as that of just one run of the algorithm.[3] Our algorithms may succeed even if the input circuit U is not a-peaked, so our algorithms cannot be used to decide if a circuit is a-peaked. However, receiving an error flag does certify that the input circuit U is not a-peaked.

[2] A closer look at our runtime analysis shows that for example, the runtime of our general architecture algorithm, as a function of n, d, a, c, satisfies $T(n, d, a, c) \leq n^{O(2^{2d}(\ln(n)+\ln(n^{\max\{a,c\}})))}$ for all integers n, d, a, c where the O notation hides absolute constants independent of n, d, a, c. Therefore, if the input integers a and c are formally specified in unary as 1^{n^a} and 1^{n^c} respectively, then the algorithm's runtime becomes properly quasipolynomial in the description length of the input when the input circuit is a shallow circuit.

[3] Since this work mainly focuses on the distinction between different classes of runtime scalings such as $\mathrm{poly}(n)$ vs. $n^{O(\log n)}$.

	Prior work	Our results
2-dimensional	$\mathrm{poly}(n)$ [75]	$\mathrm{poly}(n)$
D-dimensional, $D \geq 3$	$n^{O(\mathrm{poly}(\log n))}$ [76, 77]	$n^{O(\log\log n)}$
General	$e^{O(\sqrt{n \log n})}$ [75]	$n^{O(\log n)}$

Table 2.1: Runtime of classical algorithms for approximating output probabilities of n-qubit shallow quantum circuits to within a given $1/\mathrm{poly}(n)$ additive error. In contrast with prior work, the algorithms presented here solve the more general problem of sampling from the output distributions of peaked shallow circuits to within a given $1/\mathrm{poly}(n)$ error in total variation distance.

As a straightforward application of our sampling algorithms, we obtain classical randomized algorithms for estimating output probabilities of shallow circuits (that are not necessarily peaked) to within a given $1/\mathrm{poly}(n)$ additive error in the sense of Definition 1.2. Suppose we have an n-qubit depth-d shallow circuit U, a string $x \in \{0,1\}^n$, and a constant integer $c \geq 1$ specifying an error tolerance $\varepsilon = n^{-c}$, and we aim to compute an estimate p satisfying $|p - |\langle x|U|0^n\rangle|^2| \leq \varepsilon$. To do so, we can call one of the algorithms from Theorem 2.1 on input circuit U with $a = c$ and error tolerance specified by $c+1$. We repeat the sampling algorithm $\lceil n^{2c+2} \ln(n) \rceil$ times. If the sampler ever returns an error flag, we set $p = 0$. In this case, the circuit U is not c-peaked, which implies that $p = 0$ satisfies $|p - |\langle x|U|0^n\rangle|^2| \leq \varepsilon$. Otherwise, we set p to be the fraction of times that the outcome x is obtained. In this case, letting P' denote the distribution sampled by the algorithm from Theorem 2.1, by Hoeffding's inequality, we get that with high probability,

$$|p - |\langle x|U|0^n\rangle|^2| \leq |p - P'(x)| + |P'(x) - |\langle x|U|0^n\rangle|^2| \leq n^{-(c+1)} + n^{-(c+1)} \leq \varepsilon \tag{2.2}$$

for sufficiently large n. The asymptotic runtime of the output probability estimation algorithm described above is clearly the same as that of the underlying sampling algorithm from Theorem 2.1 up to $\mathrm{poly}(n)$ factors. This approach shows that an efficient blackbox reduction from estimating output probabilities to the statements of Theorem 2.1 is possible. In fact, it will become evident from the inner workings of the sampling algorithms responsible for Theorem 2.1 that repeated sampling is unnecessary; subroutines of the sampling algorithms include the ability to estimate output probability values up to $1/\mathrm{poly}(n)$ additive error in one go with zero failure probability. Our algorithm for general shallow circuits is sufficiently simple that we were able to implement it in software to simulate circuits with up to 56 qubits and depth 5; see Section 2.1.1 for details.

Our shallow circuit output probability estimation algorithms improve upon previously known methods. The runtimes of the best previously known algorithms for this task and our algorithms are compared in Table 2.1. The work [75] showed that this problem admits a $\mathrm{poly}(n)$-time classical algorithm for shallow circuits that are geometrically local in 2D. This

was extended to a quasipolynomial $n^{O(\mathrm{poly}(\log n))}$ time algorithm for 3D circuits in [76] and for D-dimensional circuits with constant $D \geq 4$ in [77]. For general shallow circuits lacking geometric locality, the best previously known algorithm had a scaling that is exponential in $\sqrt{n}$ [75]. In our work, we obtain a quasipolynomial $n^{O(\log n)}$ time classical algorithm for output probability estimation for general shallow quantum circuits, an exponential improvement in runtime over prior work. We also improve the asymptotic scaling of the best prior algorithms for D-dimensional shallow circuits with constant $D \geq 3$ from $n^{O(\mathrm{poly}(\log n))}$ to $n^{O(\log \log n)}$.

A simple application of output probability estimation is to additively approximate the magnitude of the expected values of tensor product observables. Let U be an n-qubit shallow quantum circuit with depth d, and we do not assume that U is peaked. Given single-qubit observables $O_1, O_2, \ldots O_n$ such that $\|O_j\| \leq 1$ for all $j \in [n]$, consider the squared expected value

$$|\langle 0^n|U^\dagger(O_1 \otimes O_2 \otimes \cdots \otimes O_n)U|0^n\rangle|^2. \tag{2.3}$$

If each observable O_j is unitary, then (2.3) is an output probability of a circuit with depth $2d+1$, so it can be estimated to within $1/\mathrm{poly}(n)$ additive error by directly applying our algorithms for output probability estimation. Otherwise, if some or all of the observables O_j are not unitary, we can still express the mean value (2.3) as an output probability of a depth $2d+1$ quantum circuit acting on at most $2n$ qubits. This follows from the simple fact that for any single-qubit operator O satisfying $\|O\| \leq 1$, we can always find a 2-qubit unitary B such that $(I \otimes \langle 0|)B(I \otimes |0\rangle) = O$; see for example, Lemma 5 of the Supplementary Material of [75]. Consequently, we obtain an $n^{O(\log n)}$ time algorithm to additively estimate the magnitude of any Pauli operator at the output of a shallow quantum circuit. Pauli expected values are key quantities needing to be estimated in variational quantum algorithms [14, 15]. However, we do not know how to use our algorithms (or the ones from [76, 77]) to compute the overall $\pm$ sign of the Pauli expected value. For 2D circuits, the efficient algorithm from [75] is capable of computing expected values of tensor product observables, including the sign.

We will see that the main novelty in proving Theorem 2.1 lies in discovering the algorithms themselves, with the runtime analysis being more or less straightforward. To understand the proof of Theorem 2.1 for general peaked shallow circuits, first consider the very special case in which the output distribution of U is a product distribution $P(x) = p_1(x_1)p_2(x_2)\cdots p_n(x_n)$, where each p_j is a distribution over a single bit $x_j \in \{0, 1\}$. Let us suppose P is of this form and in addition (2.1) holds. For simplicity, let us assume that all-zeros is a peaked outcome, so $|\langle 0^n|U|0^n\rangle|^2 \geq n^{-a}$. Then the expected number of zeros in a string x sampled from P is (using the arithmetic-geometric mean inequality)

$$\sum_{i=1}^{n} p_i(0) \geq n\left(p_1(0)p_2(0)\cdots p_n(0)\right)^{\frac{1}{n}} = nP(0^n)^{\frac{1}{n}} \geq n^{1-a/n} = n - O\left(\log n\right). \tag{2.4}$$

That is, we expect x to have $O(\log n)$ ones. This argument can be further strengthened using a Chernoff bound to show that the support of P is concentrated on binary strings of Hamming weight at most $O(\log n)$. The classical simulation algorithm is then based on restricting our attention to the subspace of Hilbert space corresponding to these $n^{O(\log n)}$ computational basis vectors. For general shallow circuits, the bits in the output distribution are not independent random variables. However, dependency is highly constrained: due to lightcone separation, each bit is independent of all but a constant number of other bits. In the proof

of Theorem 2.1 for general peaked shallow circuits, we show that the above concentration of measure argument can be extended to the setting of shallow quantum circuits. The proof uses a large deviation bound for partially dependent random variables due to Janson [78]. In this way, we also establish that the output distribution of a peaked shallow circuit is concentrated in a Hamming ball of diameter $O(\log n)$. In contrast, for shallow quantum circuits that may not be peaked, the Hamming weight is only guaranteed to be concentrated in an interval of width $O(\sqrt{n})$ around its mean value[4] [79]. Our classical simulation algorithm combines the Hamming weight concentration result with a simple circuit-to-Hamiltonian mapping. By restricting a local Hamiltonian $\sum_{j=1}^{n} U|0\rangle\langle 0|_j U^\dagger$ associated with the circuit U onto the Hamming ball of diameter $O(\log n)$ and computing an eigenvector of the restricted Hamiltonian with the largest eigenvalue, we obtain a good approximation of the output state $U|0^n\rangle$, which is specified by its $n^{O(\log n)}$ nonzero amplitudes in the computational basis. If the largest eigenvalue of the restricted Hamiltonian is below a certain threshold value, the algorithm infers that the circuit U is not a-peaked and outputs an error flag. This yields a refined version of Theorem 2.1 for general peaked shallow circuits.

Theorem 2.2 (**Output state of peaked shallow circuits**). *There exists a classical deterministic algorithm which takes as input a positive integer $a \geq 1$, the classical description of an n-qubit depth-d quantum circuit U, and a positive integer $c \geq 1$ specifying an error tolerance $\varepsilon = n^{-c}$. If the algorithm succeeds, then it outputs the classical description of an n-qubit state $|\phi\rangle$ satisfying $|\langle\phi|U|0^n\rangle|^2 \geq 1 - \varepsilon$ specified by its $n^{O(\log n)}$ nonzero amplitudes in the computational basis; otherwise, it outputs an error flag. The algorithm is guaranteed to succeed if U is a-peaked. The runtime of the algorithm is $n^{O(\log n)}$ where the $O()$ notation hides factors depending on a, c, and d.*

Theorem 2.1 for general peaked shallow circuits is a direct corollary of Theorem 2.2 obtained by using the relationship between fidelity and trace distance and the fact that trace distance does not increase under CPTP maps; in this case, measuring all qubits in the computational basis.

While Theorem 2.2 applies to general shallow quantum circuits in which 2-qubit gates may be applied to any pair of qubits, the proof of Theorem 2.1 for geometrically local peaked shallow circuits is based on a different technique. It uses some ideas from [76] but deploys them in a novel way. The key technique is a kind of dissection along one side of the grid that we now explain. For concreteness, consider the case of a shallow 2D circuit U on a $\sqrt{n} \times \sqrt{n}$ square grid of qubits that satisfies $|\langle 0^n|U|0^n\rangle|^2 \geq n^{-a}$. In the first step of the proof, we show how to find $O(\sqrt{n})$ *heavy slices* arranged in a somewhat regular pattern. By definition, each heavy slice is a tall rectangular strip of qubits of size $\sqrt{n} \times O(1)$ that has a good probability, say at least 0.99, of being measured to be in the all-zeros state. Furthermore, tracing out any heavy slice leaves the two remaining regions on the left and on the right in a tensor product state. These heavy slices are arranged in a regular pattern so that the maximum distance between heavy slices is at most $O(\log n)$. Imagine taking the true output state of the circuit $|\psi\rangle = U|0^n\rangle$ and swapping all the heavy slices with fresh qubits prepared in the

[4] It is again instructive to consider the special case of product distributions, which illustrates optimality in this setting as well.

all-zeros state. The result is a mixed state

$$\sigma' = \sigma^1 \otimes \sigma^2 \otimes \cdots \otimes \sigma^T \tag{2.5}$$

where each σ^i coincides with a reduced density matrix of the output state $|\psi\rangle$ on some quasi-one-dimensional rectangular strip of qubits of size $\sqrt{n} \times O(\log n)$, along with some additional qubits in the $|0\rangle$ state. The output distribution of σ' can be efficiently sampled using matrix product state techniques. Unfortunately, the measurement statistics of σ' do not in general approximate those of $|\psi\rangle$ because our heavy slices only come with a weak guarantee (0.99 probability of measuring all zeros). We show that this kind of weak approximation can be improved using an inclusion-exclusion method; the result is an approximation ρ to $|\psi\rangle$ that is a *real linear combination*, or "pseudomixture", of simple quantum states of the form (2.5). Although the linear combination contains exponentially many terms, its structure is such that marginals like $\mathrm{Tr}(\rho(|x\rangle\langle x|)_S)$ for $S \subseteq [n]$ and $x \in \{0,1\}^{|S|}$ can be expressed as a matrix-vector product and computed efficiently. To approximately sample from the output distribution of $|\psi\rangle$, we use a standard sampling-to-marginal reduction from [80] that is robust to error in the 1-norm. For D-dimensional circuits with $D \geq 3$, a similar dissection technique is used inductively.

The remainder of this chapter is structured as follows. In Section 2.1, we describe the Hamming weight concentration properties of peaked shallow quantum circuits and give the proofs of Theorem 2.2 and Theorem 2.1 for general peaked shallow quantum circuits. In Section 2.1.1, we present data from a software implementation of the general architecture algorithm, which illustrates the range of its utility in practice. In Section 2.2, we prove Theorem 2.1 for geometrically local peaked shallow quantum circuits. Some open questions can be found in Section 2.3.

2.1 General peaked shallow circuits

In this section, we prove Theorem 2.2 and Theorem 2.1 for general peaked shallow quantum circuits following the strategy outlined in the introduction. As a first step, we establish a certain Hamming weight concentration property for probability distributions over n-bit strings such that each bit is independent of all but a constant number of other bits.

Lemma 2.1. *Let $P(x)$ be a probability distribution on n-bit strings such that for every $j \in [n]$, the variable x_j is independent of all but at most K of the other bits and $\mathbb{E}_{x\sim P}[x_j] \leq 1/2$. Define*

$$P_{\max} = \max_{x \in \{0,1\}^n} P(x) \tag{2.6}$$

and let $\delta \leq P_{\max}^{18/25}$ be a given error tolerance. Then it holds that

$$\Pr_{x\sim P}\left[\sum_{j=1}^{n} x_j \geq \frac{25}{6}(K+1)\ln(1/\delta)\right] \leq \delta. \tag{2.7}$$

Proof. For every $x \in \{0,1\}^n$, define $w(x) = \sum_{j=1}^n x_j$ to be the Hamming weight of x. We first argue that

$$\mathop{\mathbb{E}}_{x\sim P}[w(x)] \leq (K+1)\ln(1/P_{\max}). \tag{2.8}$$

For every $j \in [n]$, define $m_j = \mathbb{E}_{x\sim P}[x_j]$. Assume for the sake of contradiction that (2.8) is false. Then by the linearity of expectation, we get

$$\mathop{\mathbb{E}}_{x\sim P}[w(x)] = \sum_{j=1}^{n} m_j > (K+1)\ln(1/P_{\max}). \tag{2.9}$$

Define a dependency graph $G = (V, E)$ with n vertices such that each vertex of G represents a variable x_j and the edges of G represent dependencies among the variables. More precisely, we assume that a variable x_j is independent of a subset of variables $S \subseteq V \setminus \{j\}$ whenever the graph G has no edges between j and S. By assumption, the degree of every vertex in G is at most K. Using a simple greedy argument, we can show that G has an independent set $S \subseteq V$ such that

$$\sum_{j\in S} m_j \geq \frac{1}{K+1}\sum_{j=1}^{n} m_j > \ln\left(1/P_{\max}\right). \tag{2.10}$$

Let $h \in \{0,1\}^n$ be any bit string such that $P(h) = P_{\max}$. Since all variables x_j with $j \in S$ are independent, we get

$$P_{\max} = P(h) \leq \Pr_{x\sim P}[x_j = h_j \text{ for all } j \in S] = \prod_{j\in S\,:\,h_j=1} m_j \prod_{j\in S\,:\,h_j=0} (1-m_j). \tag{2.11}$$

For every $j \in [n]$, since we assumed that $m_j \leq 1/2$, we have $m_j \leq 1 - m_j$. Thus

$$P_{\max} \leq \prod_{j\in S}(1-m_j) \leq \exp\left[-\sum_{j\in S} m_j\right] < P_{\max} \tag{2.12}$$

where the last inequality follows from (2.10). Clearly, we have a contradiction, so (2.8) holds. Define

$$R = (K+1)\ln\left(1/P_{\max}\right). \tag{2.13}$$

Note that $\mathbb{E}_{x\sim P}[w(x)] \leq R$ due to (2.8). A large deviation bound for a sum of random variables whose dependency graph has a bounded degree was obtained in [78]. Applying Theorem 2.3 of [78] gives for every $t \geq 0$,

$$\Pr_{x\sim P}\left[w(x) \geq R + t\right] \leq \Pr_{x\sim P}\left[w(x) \geq \mathop{\mathbb{E}}_{x\sim P}[w(x)] + t\right] \leq \exp\left(-\frac{8t^2}{25(K+1)(S+t/3)}\right), \tag{2.14}$$

where $S = \sum_{j=1}^{n} \mathrm{Var}(x_j) = \sum_{j=1}^{n} m_j - m_j^2 \leq \sum_{j=1}^{n} m_j \leq R$. We are going to choose $t = (\lambda - 1)R$ with

$$\lambda = \frac{25\ln\left(1/\delta\right)}{6\ln\left(1/P_{\max}\right)}. \tag{2.15}$$

The assumption $\delta \leq P_{\max}^{18/25}$ implies that $\lambda \geq 3$. With this choice, we get

$$R + t = \lambda R = \frac{25}{6}(K+1)\ln(1/\delta) \tag{2.16}$$

and

$$\Pr_{x\sim P}[w(x) \geq \lambda R] \leq \exp\left(-\frac{24(\lambda-1)^2R^2}{25(K+1)(2+\lambda)R}\right) \leq \exp\left(-\frac{6\lambda R}{25(K+1)}\right) = \delta \tag{2.17}$$

where the last inequality follows from the bound $(\lambda-1)^2/(2+\lambda) \geq \lambda/4$ which holds since $\lambda \geq 3$. Therefore, it holds that

$$\Pr_{x\sim P}\left[w(x) \geq \frac{25}{6}(K+1)\ln(1/\delta)\right] \leq \delta. \tag{2.18}$$

□

Suppose we are given an integer $a \geq 1$, an n-qubit shallow quantum circuit U with depth d, and an integer $c \geq 1$ specifying an error tolerance $\varepsilon = n^{-c}$. We do not assume that U is geometrically local. Let $|\psi\rangle = U|0^n\rangle$ denote the output state of U, and let $P(x) = |\langle x|\psi\rangle|^2, x \in \{0,1\}^n$ denote the output distribution of U. Define

$$P_{\max} = \max_{x\in\{0,1\}^n} P(x). \tag{2.19}$$

Below, we describe classical algorithms that, if they succeed, (a) output the description of an n-qubit state $|\phi\rangle$ such that $|\langle\phi|\psi\rangle|^2 \geq 1-\varepsilon^2/4$ and (b) sample an n-bit string x from a distribution $P'(x)$ such that $\|P-P'\|_1 \leq \varepsilon$. If the algorithms do not succeed, they output an error flag. The algorithms are guaranteed to succeed if $P_{\max}$ is sufficiently large. The precise peakedness condition we require is that

$$\varepsilon \leq 2n^{1/2}P_{\max}^{9/50}, \tag{2.20}$$

which, for sufficiently large n, is satisfied if $P_{\max} \geq n^{-a}$ and $c \geq a$. It will become evident that the algorithms have a quasipolynomial time and space complexity scaling

$$T(n,\varepsilon) = n^{O(\log n + \log(1/\varepsilon))}. \tag{2.21}$$

By Lemma 1.1 from Section 1.3, every qubit of U is lightcone-separated from all but $K = 2^{2d}$ other qubits, which is a constant independent of n. This implies that the independence condition of Lemma 2.1 is satisfied for P with $K = 2^{2d}$. Furthermore, expected values $\mathbb{E}_{x\sim P}[x_j]$ can be efficiently computed exactly by restricting the circuit to the backward lightcone of a single qubit $j \in [n]$ since $\mathbb{E}_{x\sim P}[x_j]$ is dictated by the marginal output distribution of the single qubit j and $d = O(1)$. By relabelling bit values 0 and 1, if necessary, one can ensure that $\mathbb{E}_{x\sim P}[x_j] \leq 1/2$ for all $j \in [n]$. Define

$$\delta = \frac{\varepsilon^4}{64n^2}. \tag{2.22}$$

The peakedness condition (2.20) implies that $\delta \leq P_{\max}^{18/25}$ holds. Define

$$W = \frac{25}{6}(K+1)\ln(1/\delta). \tag{2.23}$$

Therefore, by applying Lemma 2.1 to P, we get that

$$\sum_{x\in\{0,1\}^n\,:\,w(x)\le W} P(x) \ge 1-\delta \tag{2.24}$$

where $w(x)=\sum_{j=1}^n x_j$ is the Hamming weight of x.

The state $|\psi\rangle$ can be described as the unique largest eigenvector with the eigenvalue one of a parent local Hamiltonian

$$H=\frac{1}{n}\sum_{j=1}^{n} U|0\rangle\langle 0|_j U^\dagger \tag{2.25}$$

where for every $j\in[n]$, $|0\rangle\langle 0|_j$ is the projector $|0\rangle\langle 0|$ acting on the j-th qubit tensored with the identity on all remaining qubits. For every $r\in[2^n]$, let $\lambda_r(H)$ denote the r-th largest eigenvalue of H. Since H has the same spectrum as the diagonal Hamiltonian $\frac{1}{n}\sum_{j=1}^n |0\rangle\langle 0|_j$ whose set of eigenvalues is $\{\frac{m}{n} : m\in\{0,\dots,n-1,n\}\}$, we have

$$\lambda_1(H)=1 \quad \text{and} \quad \lambda_2(H)=1-\frac{1}{n}. \tag{2.26}$$

Lemma 2.2. *Let*

$$\Pi=\sum_{x\in\{0,1\}^n\,:\,w(x)\le W} |x\rangle\langle x| \tag{2.27}$$

be the projector onto the Hamming ball of radius W centred at 0^n where W is the cutoff value defined in (2.23). Consider the projected parent Hamiltonian $G=\Pi H\Pi$ and let $|\phi\rangle$ be an eigenvector of G with the largest eigenvalue $\lambda_1(G)$. If this eigenvalue satisfies

$$\lambda_1(G)\ge 1-2\sqrt{\delta}, \tag{2.28}$$

then we have

$$|\langle\phi|\psi\rangle|^2\ge 1-2n\sqrt{\delta}. \tag{2.29}$$

Moreover, (2.28) holds whenever (2.20) does.

Proof. Define $|\psi_1\rangle=\Pi|\psi\rangle$ and $|\psi_0\rangle=(I-\Pi)|\psi\rangle$, so we have $|\psi\rangle=|\psi_0\rangle+|\psi_1\rangle$ and $\langle\psi_0|\psi_1\rangle=0$. Observe that

$$\lambda_1(G)\ge\langle\psi|G|\psi\rangle=\langle\psi_1|H|\psi_1\rangle=\langle\psi|H|\psi\rangle+\langle\psi_0|H|\psi_0\rangle-2\mathrm{Re}(\langle\psi_1|H|\psi_0\rangle) \tag{2.30}$$

$$\ge 1-2\|\psi_0\|. \tag{2.31}$$

If (2.20) holds, then from (2.24), we get $\|\psi_1\|^2=\langle\psi|\Pi|\psi\rangle\ge 1-\delta$, so $\|\psi_0\|^2=1-\|\psi_1\|^2\le\delta$. Therefore, (2.20) implies

$$\lambda_1(G)\ge 1-2\sqrt{\delta}. \tag{2.32}$$

To complete the proof, we now show that (2.32) implies (2.29). Define

$$H^\perp=H-|\psi\rangle\langle\psi|, \tag{2.33}$$

so

$$H^\perp|\psi\rangle=0 \quad \text{and} \quad 0\le H^\perp\le(1-\frac{1}{n})I \tag{2.34}$$

since the set of eigenvalues of $H^\perp$ is $\{\frac{m}{n} : m \in \{0, \ldots, n-2, n-1\}\}$. Let $|\phi\rangle$ be an eigenvector of G with the eigenvalue $\lambda_1(G)$. It can be written as $|\phi\rangle = \alpha|\psi\rangle + |\psi^\perp\rangle$ for some $\alpha \geq 0$ and vector $|\psi^\perp\rangle$ orthogonal to $|\psi\rangle$. The normalization of $|\phi\rangle$ gives $\|\psi^\perp\|^2 = 1 - \alpha^2$. Clearly, $\Pi|\phi\rangle = |\phi\rangle$. Thus,

$$\lambda_1(G) = \langle\phi|G|\phi\rangle = \langle\phi|H|\phi\rangle = \alpha^2 + \langle\phi|H^\perp|\phi\rangle = \alpha^2 + \langle\psi^\perp|H^\perp|\psi^\perp\rangle \tag{2.35}$$

$$\leq \alpha^2 + (1 - \alpha^2)(1 - \frac{1}{n}) \tag{2.36}$$

where the fourth equality and the last inequality use (2.34). Combining (2.32) and (2.36), we get

$$1 - 2\sqrt{\delta} \leq \lambda_1(G) \leq \alpha^2 + (1 - \alpha^2)(1 - \frac{1}{n}). \tag{2.37}$$

Expanding the RHS and rearranging gives

$$|\langle\phi|\psi\rangle|^2 = \alpha^2 \geq 1 - 2n\sqrt{\delta}. \tag{2.38}$$

□

Now let us describe our algorithms. The first step is to redefine $c := a$ and $\varepsilon := n^{-a}$ if the input c is less than a. Then, we compute the values of K, δ, and W using their respective definitions. Next, we compute the largest eigenvalue $\lambda_1(G)$ of G with an associated eigenvector $|\phi\rangle$. To this end, consider the Hilbert space $\mathcal{H}$ whose basis vectors are n-bit strings with Hamming weight at most W. We have

$$D = \dim(\mathcal{H}) = \sum_{j=0}^{W} \binom{n}{j} \leq n^W = n^{O(\log(1/\delta))} = n^{O(\log n + \log(1/\varepsilon))}. \tag{2.39}$$

To compute $\lambda_1(G)$ with an associated eigenvector $|\phi\rangle$ of $G = \Pi H \Pi$, we can restrict G onto $\mathcal{H}$. Any matrix element of G in the chosen basis of $\mathcal{H}$ coincides with some matrix element of H in the standard basis. Since $d = O(1)$, H is a local Hamiltonian with n local terms and any matrix element of H can be computed in $O(n)$ time. After constructing G, we can use standard linear algebra algorithms to compute $\lambda_1(G)$ and $|\phi\rangle$ in time $O(nD^3)$.

The algorithm for approximating the output state proceeds as follows. If $\lambda_1(G)$ satisfies (2.28), then we output the state $|\phi\rangle$; otherwise, we output an error flag. By Lemma 2.2, whenever we output the state $|\phi\rangle$, it satisfies

$$|\langle\phi|\psi\rangle|^2 \geq 1 - 2n\sqrt{\delta} = 1 - \frac{\varepsilon^2}{4}. \tag{2.40}$$

To approximately sample from the output distribution, we first construct the above approximation $|\phi\rangle$ to the output state (or output an error flag). If this step succeeds, then we continue by computing all probabilities $P'(x) = |\langle x|\phi\rangle|^2, x \in \mathcal{H}$, which enables sampling from P' in $O(D)$ time. The standard inequality between trace distance and fidelity gives

$$\|P' - P\|_1 = \sum_{x \in \{0,1\}^n} |P'(x) - P(x)| \leq 2\sqrt{1 - |\langle\phi|\psi\rangle|^2} \leq \varepsilon \tag{2.41}$$

where the last inequality uses (2.40).

Note that the above arguments prove a slightly stronger version of Theorem 2.2 stated below.

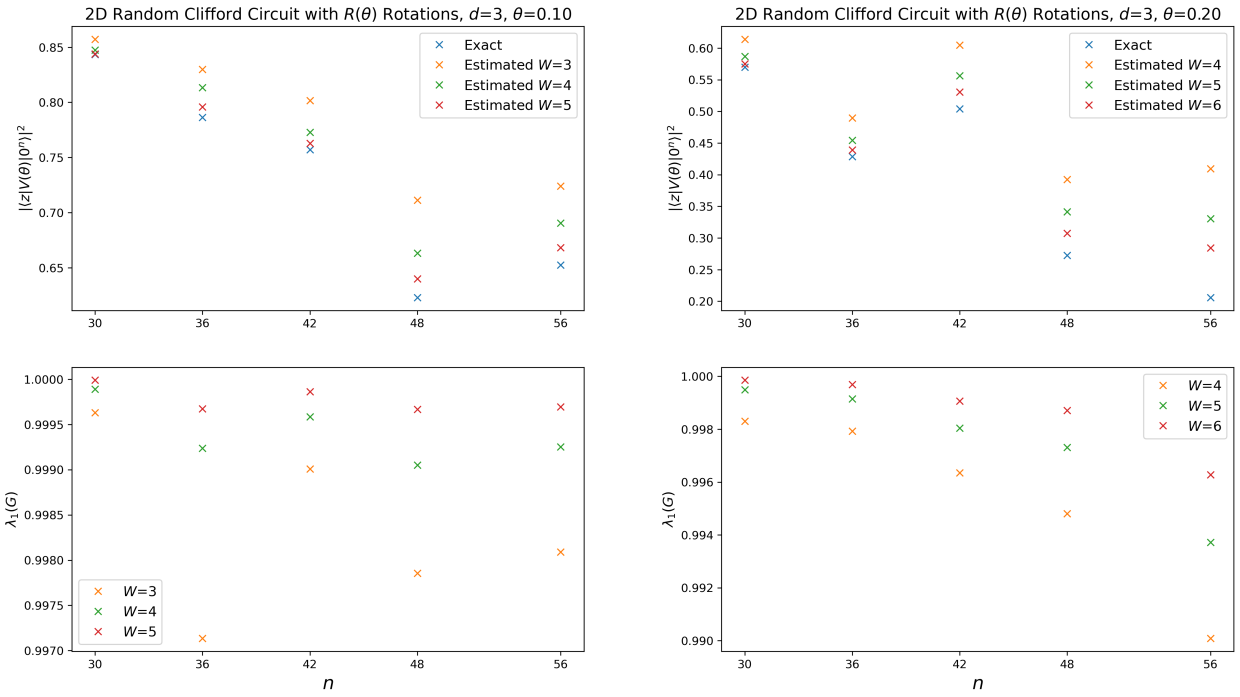


Figure 2.1: Numerical simulation of 2D random Clifford circuits interspersed with $R(\theta)$ rotations; the number of qubits n ranges from 30 to 56, and all simulated $U(\theta)$ circuits have depth $d = 3$ (so the $V(\theta)$'s all have depth 5); $\lambda_1(G)$ is the largest eigenvalue of the G matrix defined in Lemma 2.2, and W is the radius of Hamming ball centred at z; the algorithm is self-certifying since the simulation error is inversely proportional to $\lambda_1(G)$.

Theorem 2.3. *There exists a classical deterministic algorithm which takes as input the classical description of an n-qubit depth-d quantum circuit U and an error tolerance $\varepsilon > 0$. If the algorithm succeeds, then it outputs the classical description of an n-qubit state $|\phi\rangle$ satisfying $|\langle\phi|U|0^n\rangle|^2 \geq 1 - \frac{\varepsilon^2}{4}$ specified by its $n^{O(\log n + \log(1/\varepsilon))}$ nonzero amplitudes in the computational basis; otherwise, it outputs an error flag. The algorithm is guaranteed to succeed if U is sufficiently peaked such that $\varepsilon \leq 2n^{1/2}P_{\max}^{9/50}$ where*

$$P_{\max} = \max_{x \in \{0,1\}^n} |\langle x|U|0^n\rangle|^2. \tag{2.42}$$

The runtime of the algorithm is $n^{O(\log n + \log(1/\varepsilon))}$ where the $O()$ notation hides factors depending on d.

Theorem 2.2 is obtained as a special case of Theorem 2.3 since the algorithm for Theorem 2.2 always ensures $c \geq a$, and $P_{\max} \geq n^{-a}$ and $\varepsilon = n^{-c}$ for $c \geq a$ imply the condition $\varepsilon \leq 2n^{1/2}P_{\max}^{9/50}$ for sufficiently large n. We shall use Theorem 2.3 later on in Section 2.2.

2.1.1 Software implementation

In this subsection, we investigate the practical performance of the peaked shallow circuit simulation algorithm from the previous section. In particular, we implement our algorithm

to simulate n-qubit 2D random Clifford circuits interspersed with Z-rotations. We consider quantum circuits $U(\theta)$ composed of 2-qubit gates arranged in the 2D brickwork circuit architecture described in Figure 1.2 of Section 1.3. Every gate of $U(\theta)$ is a uniformly random 2-qubit Clifford gate followed by single-qubit gates $R(\theta) \otimes R(\theta)$ where

$$R(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}. \tag{2.43}$$

Define $V(\theta) = U^\dagger(\theta) Z^{\otimes n} U(\theta)$. Since $U(0)$ is a Clifford circuit, $V(0)$ is an n-qubit Pauli operator, so there exists $z \in \{0,1\}^n$ such that $|\langle z|V(0)|0^n\rangle| = 1$. Hence, for small values of θ, $V(\theta)$ will be a peaked circuit (whose peak is z), and $V(\theta)$ becomes less peaked as θ increases. We implement our algorithm to simulate instances of $V(\theta)$ for $\theta \in \{0.1, 0.2\}$, n ranging from 30 to 56, and using increasing values of the Hamming ball radius W. The algorithm outputs an estimate $p = P'(z)$ of the probability $|\langle z|V(\theta)|0^n\rangle|^2$. Recall that $P'(z) = |\langle z|\phi\rangle|^2$ where $|\phi\rangle$ is the state approximating $V(\theta)|0^n\rangle$ constructed in Lemma 2.2. Let G be the projected Hamiltonian defined in Lemma 2.2. Numerical results from Figure 2.1 confirm that the algorithm's simulation error is controlled by $\lambda_1(G)$: the closer $\lambda_1(G)$ is to 1, the closer the approximate values of $|\langle z|V(\theta)|0^n\rangle|^2$ outputted by the algorithm are to the exact values. Indeed, the proof of Lemma 2.2 implies the upper bound

$$|p - |\langle z|V(\theta)|0^n\rangle|^2| = |P'(z) - P(z)| \leq \|P' - P\|_1 \leq 2\sqrt{n(1-\lambda_1(G))}. \tag{2.44}$$

Although the algorithm enjoys favourable scaling in n, its scaling in W makes it prohibitive to try $W \geq 7$ for simulating circuits with $n = 56$. The observed simulation error is also very sensitive to how peaked the circuits are; compared to the results for $\theta = 0.1$, the error margins widen for $\theta = 0.2$ where the circuits being simulated are not as peaked. Nonetheless, the value of $\lambda_1(G)$ can still be used to gauge simulation accuracy. In addition, it is worth noting that even for the largest $n = 56$ circuits considered here, the amplitude $\langle z|V(\theta)|0^n\rangle$ (including the sign) can be computed exactly using off-the-shelf tensor network simulators in seconds [81]. However, our algorithm finds at once a whole distribution close to the exact one (i.e., an eigenvector corresponding to $\lambda_1(G)$), and it could be much slower to use tensor network based methods to construct a distribution-encoding vector one entry at a time. Therefore, it may be advantageous to adopt our algorithm for applications that require full distributional information (for example, approximate sampling or computation of entanglement measures).

2.2 Geometrically local peaked shallow circuits

In this section, we prove Theorem 2.1 for geometrically local peaked shallow quantum circuits. We will rephrase the algorithm for 2D and the algorithm for 3D and higher as a single algorithm that satisfies different runtime guarantees depending on the geometric locality of the input circuit.

Theorem 2.4. *There exists a classical randomized algorithm which takes as input a positive integer $a \geq 1$, the classical description of an n-qubit depth-d quantum circuit U that is geometrically-local in D-dimensions for some $D \geq 2$, and a positive integer $c \geq 1$ specifying an error tolerance $\varepsilon = n^{-c}$. If the algorithm succeeds, then it outputs a sample drawn from a probability distribution P' over n-bit strings such that $\|P' - P\|_1 \leq \varepsilon$ where P is the computational-basis output distribution of the input circuit U; otherwise, it outputs an error flag. The algorithm is guaranteed to succeed if U is a-peaked. The runtime of the algorithm is $n^{O(1)}$ for $D = 2$ where the $O()$ notation hides factors depending on a, c, and d and $n^{O(\log \log n)}$ for $D \geq 3$ where the $O()$ notation hides factors depending on a, c, d, and D.*

Suppose we are given input $a, c \geq 1$ and an n-qubit shallow quantum circuit U with depth-d that is geometrically local in D dimensions for some constant $D \geq 2$. In the following, we treat a and c as $O(1)$ constants independent of n. We assume the n qubits are arranged in an $n^{1/D} \times n^{1/D} \times \cdots \times n^{1/D}$ grid. Let $|\psi\rangle = U|0^n\rangle$ denote the output state of U, and let $P(x) = |\langle x|\psi\rangle|^2, x \in \{0,1\}^n$ denote the computational-basis output distribution of U. As in the previous section, we shall assume without loss of generality that for every $j \in [n]$,

$$\mathop{\mathbb{E}}_{x\sim P}[x_j] = \Pr_{x\sim P}(x_j = 1) \leq \frac{1}{2} \tag{2.45}$$

where x_j is the j-th bit of x. We shall be interested in the case where the circuit U is a-peaked, which means there exists $h \in \{0,1\}^n$ such that

$$P(h) \geq n^{-a}. \tag{2.46}$$

We do not assume h is given. However, (2.45), (2.46), and Lemma 2.1 (setting $\delta = n^{-2a}$) imply that the Hamming weight $w(h) = \sum_{j=1}^n h_j$ satisfies

$$w(h) \leq \frac{50a}{6}(K+1)\ln(n) \tag{2.47}$$

where K is a circuit architecture dependent constant.

Our strategy is to establish the following theorem for $\varepsilon = n^{-c}$. Recall the notations we have established in Section 1.3 for expressing product states and operators acting on a subset of qubits. Let $\mathcal{M}(\cdot)$ be the CPTP map corresponding to measuring all qubits in the computational basis.

Theorem 2.5. *If (2.46) holds, then there exists a Hermitian matrix τ satisfying*

$$\|\mathcal{M}(\tau) - \mathcal{M}(|\psi\rangle\langle\psi|)\|_1 = \sum_{y\in\{0,1\}^n} |\langle y|\tau|y\rangle - |\langle y|\psi\rangle|^2| \leq \varepsilon \tag{2.48}$$

and such that there exists a classical algorithm to exactly compute "marginal probabilities" $\mathrm{Tr}(\tau(|x\rangle\langle x|)_S)$ for every $S \subseteq [n]$ and $x \in \{0,1\}^{|S|}$. The runtime of the algorithm is $n^{O(1)}$ if $D = 2$ and $n^{O(\log\log n)}$ if $D \geq 3$. If the algorithm is run on an input circuit U that does not satisfy (2.46), it will either succeed, in which case (2.48) holds, or it will output an error flag.

Theorem 2.4 follows from Theorem 2.5 along with the fact that the standard linear-time reduction from sampling to computing marginals ("bit-by-bit") is robust to error in the 1-norm, even when the approximating vector can take negative values, as shown in [80].

Lemma 2.3 (Lemma 10 of [80]). *Let p be a probability distribution over n-bit strings. Suppose $p' : \{0,1\}^n \to \mathbb{R}$ satisfies $\|p' - p\|_1 \leq \delta$ for some $\delta < 1$. If there is a classical algorithm with runtime $T(n)$ that, given $0 \leq j \leq n$ and $x \in \{0,1\}^j$, exactly computes*

$$\sum_{y \in \{0,1\}^{n-j}} p'(xy), \tag{2.49}$$

then there is a classical algorithm with runtime $nT(n)$ that samples from a probability distribution q such that $\|q - p\|_1 \leq 4\delta/(1-\delta)$.

In the proof of Theorem 2.5 for $D = 2$, we will construct τ explicitly as a pseudomixture (real linear combination of density matrices) of tensor products of matrix product states. For $D \geq 3$, τ is constructed as a pseudomixture of tensor products of states on subgrids each consisting of $O(\log^D n)$ qubits. The remainder of this section is concerned with establishing Theorem 2.5.

2.2.1 Dissection along a side

Our proof technique uses some of the ideas from [76]. In particular, the method is based on choosing one of the sides of the D-dimensional grid and identifying *heavy slices*, which are subgrids of qubits of constant thickness along the chosen side that have a high chance of being measured in the all-zeros state. However, our method differs substantially from [76] in the way the heavy slices are used in the algorithm. Whereas [76] uses a heavy slice to divide the grid into left and right subproblems (reducing to problems of half the size), here we show how to use a 1D regular arrangement of heavy slices to reduce our D-dimensional problem to one that is quasi-$(D-1)$-dimensional. More precisely, the type of quantum states we will encounter will be reduced density matrices of $|\psi\rangle = U|0^n\rangle$ of the following type.

Definition 2.1. We say ρ is an (n, k, D)-state if it is equal to the reduced density matrix of $|\psi\rangle$ on a D-dimensional rectangular grid of qubits with k sides of length $n^{\frac{1}{D}}$ and $D - k$ sides of length $O(\log n)$.

First note that $|\psi\rangle\langle\psi|$ is itself an (n, D, D)-state. Suppose ρ is an (n, k, D)-state with $1 \leq k \leq D$. For every $j \in [D]$, let S_j denote the length of the j-th side of the grid of qubits, so we can identify the qubits in the support of ρ by a coordinate $(i_1, i_2, \ldots, i_D)$ where for every $j \in [D]$, $1 \leq i_j \leq S_j$. Let us suppose without loss of generality that the first side has length $n^{1/D}$, so $S_1 = n^{1/D}$. It will be convenient to refer to subsets of the qubits of the form

$$\{(i_1, i_2, \ldots, i_D) : \alpha \leq i_1 \leq \beta \text{ and } 1 \leq i_j \leq S_j \text{ for } j \in \{2, \ldots, D\}\} \tag{2.50}$$

which we call a *rectangular region* of width $\beta - \alpha$.

Definition 2.2. Suppose ρ is an (n, k, D)-state. A heavy slice s is any rectangular region of width $w = 2d + 1$ satisfying

$$\mathrm{Tr}(\rho(|0\rangle\langle 0|)_s) \geq 0.99. \tag{2.51}$$

Note that the definition of a heavy slice depends on the support of the reduced density matrix ρ as well as the circuit U, though this dependence is suppressed in our terminology. A heavy slice s partitions the support of ρ into three rectangular regions consisting of s and the regions L and R to the left and right of s, respectively. The condition that the width of the slice satisfies $w = 2d+1$ ensures that the backward lightcones of L and R are disjoint. In particular, this means that when we trace out any heavy slice s, we obtain a tensor product

$$\mathrm{Tr}_s(\rho) = \sigma_L \otimes \sigma_R \tag{2.52}$$

for some states σ_L, σ_R on the left and right rectangular regions, respectively.

If s is a heavy slice, then there is a good (at least 99%) chance of measuring 0 for all the qubits in s. Therefore, measuring the qubits in s in the computational basis is approximated by simply tracing out the qubits in s and replacing them with all zeros, as we now show. For any subset of qubits $S \subseteq [n]$, define CPTP maps

$$M_S(\rho) = \sum_{z\in\{0,1\}^{|S|}} (\langle z|_S)\rho(|z\rangle_S) \otimes |z\rangle\langle z| \quad \text{and} \quad E_S(\rho) = \mathrm{Tr}_S(\rho) \otimes |0^{|S|}\rangle\langle 0^{|S|}|. \tag{2.53}$$

Lemma 2.4. *Suppose ρ is an (n, k, D)-state. Let s be a heavy slice. Then*

$$\|M_s(\rho) - E_s(\rho)\|_1 \leq 0.02. \tag{2.54}$$

Proof. Using the definitions, we can write

$$M_s(\rho) - E_s(\rho) = \sum_{z\in\{0,1\}^{|s|}, z\neq 0^{|s|}} (\langle z|_s)\rho(|z\rangle_s) \otimes (|z\rangle\langle z| - |0^{|s|}\rangle\langle 0^{|s|}|). \tag{2.55}$$

Since $\||z\rangle\langle z| - |0^{|s|}\rangle\langle 0^{|s|}|\|_1 = 2$ whenever $z \neq 0^{|s|}$, we have

$$\begin{aligned}
\|M_s(\rho) - E_s(\rho)\|_1 &\leq \sum_{z\in\{0,1\}^{|s|}, z\neq 0^{|s|}} 2\|(\langle z|_s)\rho(|z\rangle_s)\|_1 && (2.56)\\
&= 2 \sum_{z\in\{0,1\}^{|s|}, z\neq 0^{|s|}} \mathrm{Tr}(\rho(|z\rangle\langle z|)_s) && (2.57)\\
&= 2\mathrm{Tr}(\rho(I - |0^{|s|}\rangle\langle 0^{|s|}|)_s) && (2.58)\\
&\leq 0.02 && (2.59)
\end{aligned}$$

where in the last inequality, we used the fact that s is a heavy slice. □

We now show that (2.45) and (2.46) imply the existence of such heavy slices. In the following lemma, we show how to find a large number of them, arranged in a certain regular pattern. To describe this pattern, we will refer to a partition of the grid into $2T$ equal-width rectangular regions $V^1, H^1, \ldots, V^T, H^T$, as shown in Figure 2.2 for the special case $D = 2$. Below, we will choose the width of the rectangular regions to satisfy

$$W = \Theta(d \log n) \tag{2.60}$$

so that the total number of regions is

$$T = \frac{n^{1/D}}{2W} = \Theta\left(\frac{n^{1/D}}{d\log(n)}\right). \tag{2.61}$$

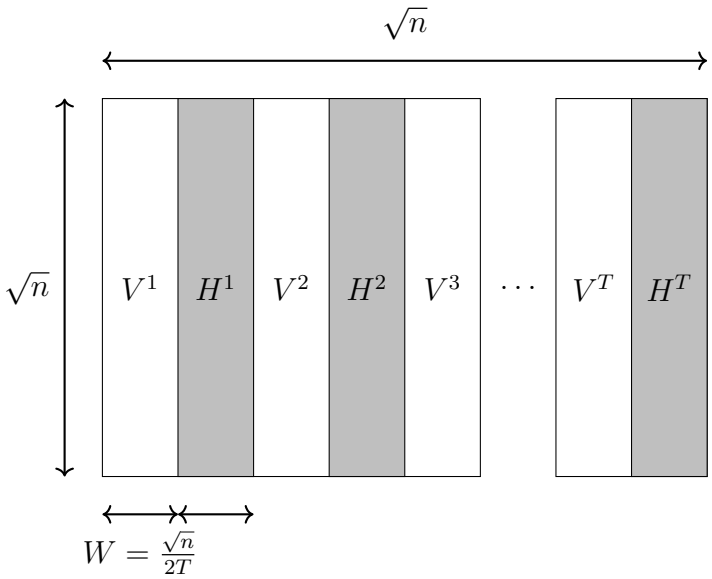


Figure 2.2: A partition of the grid of qubits into equal-width rectangular regions, depicted here in the two-dimensional case. Lemma 2.5 describes how we can find lightcone-separated heavy slices within each region H^j for $j \in [T]$.

Lemma 2.5. *Suppose (2.46) holds and that ρ is an (n, k, D)-state. Consider partitioning the qubits into equal-sized rectangular regions $V^1, H^1, \ldots, V^T, H^T$ as shown in Figure 2.2. Then there exist $\ell = \Theta(\log n)$ heavy slices $s_1^j, s_2^j, \ldots, s_\ell^j$ within each region H^j for $j \in [T]$ where*

$$\ell \geq \ln(4n^{1/D}\varepsilon^{-1}). \tag{2.62}$$

The heavy slices are all pairwise lightcone-separated. If $D = k = 2$, then there exists an efficient classical algorithm that outputs the heavy slices.

Proof. Let $W = \frac{n^{1/D}}{2T}$ be the width of each region H^j. Recall that the heavy slices we aim to find must have width $2d + 1$ and must be lightcone-separated from other heavy slices. Let us fix $N = \Theta(W/d)$ candidate heavy slices $C_1^j, C_2^j, \ldots, C_N^j \subseteq H^j$, all of which satisfy these two properties (required width and separation from the other candidates). Recall that h is the "heavy string" satisfying (2.46). For each $j \in [T]$ and $i \in [N]$, let h_{ij} be the substring of h consisting of the bits in region C_i^j. Since all the candidates are lightcone-separated we have, for each $j \in [T]$,

$$\prod_{i=1}^{N} \mathrm{Tr}(\rho(|h_{ij}\rangle\langle h_{ij}|)_{C_i^j}) = \langle\psi|\left(\prod_{i=1}^{N}(|h_{ij}\rangle\langle h_{ij}|)_{C_i^j}\right)|\psi\rangle \geq P(h) \geq \frac{1}{n^a} \tag{2.63}$$

where we used (2.46). To ease notation, let us fix $j \in [T]$ and write $p_i = \mathrm{Tr}(\rho(|h_{ij}\rangle\langle h_{ij}|)_{C_i^j})$. Then taking the log of the above gives

$$\frac{1}{N}\sum_{i=1}^{N} \ln\left(\frac{1}{p_i}\right) \leq \frac{a \ln n}{N}. \tag{2.64}$$

By Markov's inequality, there is a subset $S \subseteq [N]$ of size $|S| \geq N/2$ such that

$$\ln\left(\frac{1}{p_i}\right) \leq \frac{2a \ln n}{N} \qquad \text{for all } i \in S. \tag{2.65}$$

Our heavy slices within region H^j will be the candidates C_i^j with $i \in S$. We choose $W = \Theta(d \log n)$ large enough so that $N = \Theta(W/d)$ satisfies $\frac{2a \ln n}{N} \leq \ln(1/0.99)$ and $N/2 \geq \ln(4n^{1/D}\varepsilon^{-1})$. The first condition ensures that $p_i \geq 0.99$. From (2.45), we see that this implies that h_{ij} is the all-zeros string whenever $i \in S$, i.e.,

$$\mathrm{Tr}(\rho(|0\rangle\langle 0|)_{C_i^j}) \geq 0.99 \quad \text{for all} \quad i \in S, \tag{2.66}$$

which shows that our candidates C_i^j with $i \in S$ are heavy slices. The condition $N/2 \geq \ln(4n^{1/D}\varepsilon^{-1})$ ensures that the number of heavy slices $\ell = |S| \geq N/2$ within each region H^j satisfies (2.62).

So far, we have shown the existence of enough heavy slices among our candidates. If $D = k = 2$, then we can efficiently check whether or not a given candidate is a heavy slice, since the circuit composed of all the gates in the backward lightcone of a given candidate is a depth-d circuit on a rectangular strip of size $\sqrt{n} \times O(d)$, and can be simulated efficiently using matrix product state techniques. □

From Lemma 2.5, we infer the existence of lots of lightcone-separated heavy slices. We will use them along with the following extension of Lemma 2.4 (whose proof is quite similar).

Lemma 2.6. *Suppose ρ is an (n, k, D)-state. Let $s_1, s_2, \ldots, s_m$ be any set of lightcone-separated heavy slices. For any $C \subseteq [m]$, we have*

$$\left\| \prod_{k \in [m]\setminus C} M_{s_k} \prod_{i \in C} (M_{s_i} - E_{s_i})(\rho) \right\|_1 \leq 0.02^{|C|}. \tag{2.67}$$

(The above product should be interpreted as the composition of commuting superoperators.)

Proof. We have

$$\prod_{k \in [m]\setminus C} M_{s_k} \prod_{i \in C} (M_{s_i} - E_{s_i})(\rho) \tag{2.68}$$

$$= \sum_{z_1, z_2, \ldots z_m} (\otimes_{i=1}^m \langle z_i|_{s_i}) \, \rho \, (\otimes_{i=1}^m |z_i\rangle_{s_i}) \bigotimes_{i \in C} \left(|z_i\rangle\langle z_i| - |0^{|s_i|}\rangle\langle 0^{|s_i|}|\right) \bigotimes_{k \in [m]\setminus C} (|z_k\rangle\langle z_k|). \tag{2.69}$$

Therefore,

$$\left\| \prod_{k \in [m]\setminus C} M_{s_k} \prod_{i \in C} (M_{s_i} - E_{s_i})(\rho) \right\|_1 \leq \sum_{\substack{z_1, z_2, \ldots, z_m \text{ s.t.} \\ z_i \neq 0^{|s_i|} \text{ for } i \in C}} \| (\otimes_{i=1}^m \langle z_i|_{s_i}) \, \rho \, (\otimes_{i=1}^m |z_i\rangle_{s_i}) \|_1 2^{|C|} \tag{2.70}$$

$$= 2^{|C|} \mathrm{Tr}\left(\rho \prod_{i \in C} (I - |0^{|s_i|}\rangle\langle 0^{|s_i|}|)_{s_i} \right) \tag{2.71}$$

$$\leq 0.02^{|C|}. \tag{2.72}$$

□

Now, for each $j \in [T]$, define a superoperator that describes the measurement (in the computational basis) of all qubits in all the heavy slices in region H^j from Lemma 2.5

$$M^j = \prod_{i=1}^{\ell} M_{s_i^j}. \tag{2.73}$$

Write

$$M^j = A^j + B^j \tag{2.74}$$

where

$$B^j = \prod_{i=1}^{\ell} \left(M_{s_i^j} - E_{s_i^j} \right) \tag{2.75}$$

and

$$A^j = M^j - B^j = \sum_{\omega \subseteq [\ell], \omega \neq \emptyset} (-1)^{|\omega|+1} \prod_{i \in [\ell] \setminus \omega} M_{s_i^j} \prod_{i \in \omega} E_{s_i^j}. \tag{2.76}$$

Lemma 2.7. *Suppose ρ is an (n, k, D)-state, and that $s_1^j, s_2^j, \ldots, s_\ell^j$ for $j \in [T]$ are heavy slices with the properties described in Lemma 2.5. Then, for any $\Omega \subseteq [T]$, we have*

$$\left\| \left(\prod_{j \in \Omega} A^j - \prod_{j \in \Omega} M^j \right)(\rho) \right\|_1 \leq \frac{\varepsilon}{2}. \tag{2.77}$$

Proof. Let $\Omega \subseteq [T]$. By the definition of A^j, we have

$$\prod_{j \in \Omega} A^j(\rho) = \prod_{j \in \Omega} (M^j - B^j)(\rho) = \prod_{j \in \Omega} M^j(\rho) + \sum_{S \subseteq \Omega, S \neq \emptyset} (-1)^{|S|} \prod_{j \in \Omega \setminus S} M^j \prod_{k \in S} B^k(\rho). \tag{2.78}$$

Note that a direct application of Lemma 2.6 gives, for every $S \subseteq \Omega$,

$$\left\| \prod_{j \in \Omega \setminus S} M^j \prod_{k \in S} B^k(\rho) \right\|_1 \leq 0.02^{\ell |S|}. \tag{2.79}$$

Therefore, using the triangle inequality, we arrive at

$$\left\| \left(\prod_{j \in \Omega} A^j - \prod_{j \in \Omega} M^j \right)(\rho) \right\|_1 \leq \sum_{S \subseteq \Omega, S \neq \emptyset} 0.02^{\ell |S|} = (1 + 0.02^\ell)^{|\Omega|} - 1 \leq (1 + 0.02^\ell)^T - 1. \tag{2.80}$$

We can further upper bound the above by

$$(2.80) = 0.02^\ell \sum_{j=0}^{T-1} (1 + 0.02^\ell)^j \leq 0.02^\ell T (1 + 0.02^\ell)^T \leq T \cdot 0.02^\ell e^{T \cdot 0.02^\ell}. \tag{2.81}$$

Since ℓ satisfies (2.62), we have $0.02^\ell \leq e^{-\ell} \leq \varepsilon/(4n^{1/D})$. We also have $T \leq n^{1/D}$ (see Figure 2.2). Plugging these into (2.81), we see that

$$\left\| \left(\prod_{j \in \Omega} A^j - \prod_{j \in \Omega} M^j \right)(\rho) \right\|_1 \leq \frac{\varepsilon}{4} e^{\varepsilon/4}. \tag{2.82}$$

Since $e^{\varepsilon/4} \leq e^{1/2} \leq 2$, we are done. □

2.2.2 Approximating the output state

Suppose that (2.46) holds and that ρ is an (n, k, D)-state. We shall make use of the heavy slices whose existence is guaranteed by Lemma 2.5 and the following matrix

$$\sigma = A^1 A^2 \cdots A^T(\rho) \tag{2.83}$$

that is defined in terms of them. Note that the operators $A^1, \ldots, A^T$ depend on the location of the heavy slices, so we will have to compute this information (i.e., the guarantee that they exist is not enough); we will account for the runtime required to do so later. We now describe some properties of σ.

Noting from (2.76) that each A^j is a real linear combination of CPTP maps, we see that σ is Hermitian. For any Hermitian matrix Q, we have $\|Q\|_1 \geq \sum_y |\langle y|Q|y\rangle|$. Applying this along with Lemma 2.7, we get

$$\frac{\varepsilon}{2} \geq \left\|(A^1 A^2 \cdots A^T - M^1 M^2 \cdots M^T)(\rho)\right\|_1 \geq \sum_{y \in \{0,1\}^n} |\langle y|\sigma|y\rangle - \langle y|\rho|y\rangle| \tag{2.84}$$

where in the last inequality, we used the fact that $M^1 M^2 \cdots M^T(\rho)$ has the same diagonal entries as ρ since each M^j is a CPTP map corresponding to measuring some of the qubits in the computational basis. Thus, the matrix σ satisfies the error guarantee (2.48) from Theorem 2.5 with $\varepsilon \leftarrow \varepsilon/2$.

From the definition (2.76), we see that A^j is a linear combination of $L = 2^\ell - 1$ CPTP maps. Since $\ell = \Theta(\log n)$, we have $L \leq n^{O(1)}$. Moreover, σ is a pseudomixture

$$\sigma = \sum_{a_1, a_2, \ldots, a_T = 1}^{L} f_{a_1} f_{a_2} \cdots f_{a_T} \cdot \mathcal{R}_{\vec{a}}(\rho) \tag{2.85}$$

where $\{f_a\}_a$ are efficiently computable ± 1-valued coefficients, and each $\mathcal{R}_{\vec{a}}$ is a tensor product of CPTP maps, each of which either measures a heavy slice in the computational basis or traces out a heavy slice and replaces it with all-zeros. Crucially, due to the fact that $\omega \neq \emptyset$ in (2.76), the CPTP map $\mathcal{R}_{\vec{a}}$ always traces out at least one heavy slice in each of the regions H^j for $j \in [T]$. Therefore, we may write

$$\mathcal{R}_{\vec{a}}(\rho) = \sigma^1_{a_1} \otimes \sigma^2_{a_1 a_2} \otimes \sigma^3_{a_2 a_3} \otimes \cdots \otimes \sigma^T_{a_{T-1} a_T} \tag{2.86}$$

where $\sigma^1_{a_1}$ is a state whose support overlaps regions V^1 and H^1, $\sigma^2_{a_1 a_2}$ is a state with support overlapping the three regions H^1, V^2, and H^2, and etc. Note that the subsystems separated by the tensor product in (2.86) depend on $a_1, a_2, \ldots, a_T$. For notational convenience, in the following we shall write $\sigma^1_{a_0 a_1} = \sigma^1_{a_1}$ for all a_0.

Each state $\sigma^j_{a_{j-1} a_j}$ can be prepared by starting with a reduced state of $|\psi\rangle$ on a rectangular region of width $O(\log n)$ (an $(n, k-1, D)$-state), adjoining some fresh qubits in the all-zeros state, and then measuring some qubits. The latter operation does not alter the measurement statistics in the computational basis. In particular, the measurement statistics of $\sigma^j_{a_{j-1} a_j}$ are those of an $(n, k-1, D)$ state in tensor product with some ancilla qubits in the all-zeros state.

Finally, let us describe a method for computing "marginals" of the form $\mathrm{Tr}(\sigma(|x\rangle\langle x|)_S)$. In particular, we describe a polynomial-time reduction from this task (computing marginals of σ) to the task of computing marginals of the states $\sigma^1_{a_1}, \sigma^2_{a_1a_2}, \ldots$ (this is nontrivial because the sum (2.85) includes a superpolynomial number of terms). As discussed above, this latter task reduces to computing marginals of $(n, k-1, D)$-states.

For any subset $S \subseteq [n]$ in the support of ρ, bit string $x \in \{0,1\}^{|S|}$, and indices $a_1, a_2, \ldots, a_T \in [L]$, we can efficiently compute projectors $\Pi^1_{a_1}, \Pi^2_{a_1,a_2}, \ldots, \Pi^T_{a_{T-1},a_T}$ diagonal in the computational basis, and each of the form $|y\rangle\langle y| \otimes I$ for a suitable substring y of x, such that

$$\mathrm{Tr}(\mathcal{R}_{\vec{a}}(\rho)(|x\rangle\langle x|)_S) = \mathrm{Tr}(\sigma^1_{a_1}\Pi^1_{a_1})\mathrm{Tr}(\sigma^2_{a_1a_2}\Pi^2_{a_1a_2}) \cdots \mathrm{Tr}(\sigma^T_{a_{T-1}a_T}\Pi^T_{a_{T-1}a_T}). \tag{2.87}$$

Therefore,

$$\mathrm{Tr}(\sigma(|x\rangle\langle x|)_S) = \sum_{a_1,a_2,\ldots,a_T=1}^{L} (f_{a_1}\mathrm{Tr}(\sigma^1_{a_1}\Pi^1_{a_1}))(f_{a_2}\mathrm{Tr}(\sigma^2_{a_1a_2}\Pi^2_{a_1a_2})) \cdots (f_{a_T}\mathrm{Tr}(\sigma^T_{a_{T-1}a_T}\Pi^T_{a_{T-1}a_T})). \tag{2.88}$$

The RHS is a matrix-vector product $vG^2G^3\cdots G^Tw$ where v is $1 \times L$, $G^2, G^3, \ldots, G^T$ are $L \times L$, and w is the $L \times 1$ all-ones vector. Each entry of each of these matrices and vectors is a marginal of one of the states $\sigma^1_{a_1}, \sigma^2_{a_1a_2}, \ldots$, and we need to compute all $O(TL^2) \leq n^{O(1)}$ such entries. Then, we can compute the RHS using matrix-vector multiplication with runtime $O(TL^2) \leq n^{O(1)}$.

We are now ready to prove Theorem 2.5. The proof uses the approximation σ to the output state but proceeds slightly differently for $D = 2$ and $D \geq 3$ respectively.

Proof of Theorem 2.5. During the course of the algorithm described below, we will need to compute certain sets of heavy slices whose existence is guaranteed by Lemma 2.5 as long as (2.46) holds. If we run the algorithm on an input circuit U that does not satisfy (2.46), then we may be unable to find them, in which case we output an error flag. Our algorithm also uses, as a subroutine, the classical simulation algorithm for general peaked shallow circuits from Section 2.1; this algorithm may also output an error flag if (2.46) does not hold. For simplicity, we assume (2.46) holds in the remainder of the proof.

First consider the case $D = 2$. For 2D circuits ($D = k = 2$), the heavy slices from Lemma 2.5 can be computed efficiently (as stated in the Lemma). In this case, our approximation to the output state of the circuit is the Hermitian matrix σ defined in (2.83). We have already established the error bound (2.48) in (2.84) by taking $\rho = |\psi\rangle\langle\psi|$. All that remains to show is that marginals of the form $\mathrm{Tr}(\sigma(|x\rangle\langle x|)_S)$ can be computed efficiently. Above, we showed that this task reduces in $\mathrm{poly}(n)$ time to computing marginals of the states $\{\sigma^j_{a_{j-1}a_j}\}$. Moreover, this latter task is equivalent to computing marginals of $(n, 1, 2)$-states, i.e., a state prepared by a constant-depth circuit in a quasi-1D rectangular region of width $O(\log(n))$ and height $\sqrt{n}$. Such a circuit prepares a matrix product state of $\mathrm{poly}(n)$ bond dimension, and marginals of such states can be computed efficiently using standard techniques.

Next suppose $D \geq 3$. Here, U is a depth-d circuit composed of geometrically local gates between nearest-neighbour qubits on a D-dimensional grid such that each side has length $n^{1/D}$. By definition, the state $|\psi\rangle\langle\psi|$ of interest is an (n, D, D)-state. We will establish the theorem for (n, k, D)-states ρ. The proof is by induction on k for $k = 0, \ldots, D$.

We first establish a technical lemma. Consider the reduced density matrix ρ of an n-qubit shallow circuit describing a subset of qubits $A \subseteq [n]$ such that $\langle h|\rho|h\rangle \geq \alpha$ for some heavy string $h \in \{0,1\}^{|A|}$ and peakedness $\alpha > 0$. The following lemma shows, in particular, that ρ admits a purification that is the output state of a shallow circuit with peakedness α^2. For technical reasons, we shall use the slightly more general statement below, which describes a family of shallow circuits that each purify ρ, and at least one of which has peakedness α^2.

Lemma 2.8 (Peaked shallow purification)**.** *Suppose V is an n-qubit depth-d quantum circuit, and suppose the qubits are partitioned as $[n] = A \cup B$. Let $\rho = \mathrm{Tr}_B\left(V|0^n\rangle\langle 0^n|V^\dagger\right)$ be the reduced density matrix of the output state on subsystem A. Suppose further that for a given $h \in \{0,1\}^{|A|}$, we have*

$$\langle h|\rho|h\rangle \geq \alpha. \tag{2.89}$$

For each $z \in \{0,1\}^{|A|}$, let $X(z)$ be a $|A|$-qubit Pauli operator such that $X(z)|0^{|A|}\rangle = |z\rangle$. Define $m = 2|A| + B$ and define m-qubit unitaries $Q(z)$ acting on three registers $AA'B$, where $|A'| = |A|$, by

$$Q(z) = V^\dagger_{A'B} V_{AB} X(z)_{A'}. \tag{2.90}$$

Notice that $Q(z)$ acts on $m \leq 2n$ qubits and has depth $2d$. Moreover, it holds that

$$\max_{x\in\{0,1\}^m} |\langle x|Q(h)|0^m\rangle|^2 \geq \alpha^2 \tag{2.91}$$

and

$$\mathrm{Tr}_{A'B}(Q(z)|0^m\rangle\langle 0^m|Q(z)^\dagger) = \rho \quad \text{for all} \quad z \in \{0,1\}^{|A|}. \tag{2.92}$$

Proof. For every $z \in \{0,1\}^{|A|}$, using (2.90) and the cyclic property of the trace, we get

$$\mathrm{Tr}_{A'B}(Q(z)|0^m\rangle\langle 0^m|Q(z)^\dagger) = \mathrm{Tr}_{A'B}\left(V_{AB}(|0^n\rangle\langle 0^n|)_{AB}V^\dagger_{AB}(|z\rangle\langle z|)_{A'}\right) = \rho. \tag{2.93}$$

Now write the Schmidt decomposition

$$V|0^n\rangle = \sum_k \sqrt{p_k}|\alpha_k\rangle \otimes |\beta_k\rangle \tag{2.94}$$

so that $\rho = \sum_k p_k|\alpha_k\rangle\langle\alpha_k|$. Then we have

$$(\langle h|_A\langle 0^n|_{A'B})Q(h)|0^m\rangle = \left(\sum_k \sqrt{p_k}\langle h|_A\langle\alpha_k|_{A'}\langle\beta_k|_B\right)\left(\sum_j \sqrt{p_j}|\alpha_j\rangle_A|\beta_j\rangle_B|h\rangle_{A'}\right) \tag{2.95}$$

$$= \sum_k p_k|\langle h|\alpha_k\rangle|^2 = \langle h|\rho|h\rangle \geq \alpha, \tag{2.96}$$

which completes the proof. □

Base case The base case of the induction is an $(n, 0, D)$-state ρ with error tolerance $\varepsilon = n^{-c}$. Let $J \subseteq [n]$ be the support of ρ, and let J' be the backward lightcone of J with respect to the circuit U. From the definition of an $(n, 0, D)$-state, we have $|J| = O(\log^D n)$,

and since U has constant depth, we also have $|J'| = O(1) \cdot |J| = O(\log^D n)$. Let $U(J)$ denote the backward lightcone circuit of J. We have

$$\rho = \mathrm{Tr}_{[n]\setminus J}\left(U|0^n\rangle\langle 0^n|U^\dagger\right) = \mathrm{Tr}_{J'\setminus J}\left(U(J)|0^{|J'|}\rangle\langle 0^{|J'|}|U(J)^\dagger\right). \tag{2.97}$$

The circuit $U(J)$ that prepares ρ acts on $N = O(\log^D n)$ qubits in total and has depth d. A naive classical simulation of this circuit would expend runtime $e^{O(\log^D n)}$, but we show below how to do better than this by exploiting (2.98) and (2.99).

First note that (2.46) implies

$$\langle h_J|\rho|h_J\rangle \geq n^{-a} \tag{2.98}$$

where h_J is the substring of h consisting of bits in J. By (2.47), the Hamming weight of h_J is upper bounded as

$$w(h_J) \leq w(h) \leq r\ln(n) \tag{2.99}$$

where r is a known constant from (2.47). Define

$$\mathcal{C} = \{z \in \{0,1\}^{|J|} : w(z) \leq r\ln(n)\}. \tag{2.100}$$

From (2.99), we know $h_J \in \mathcal{C}$. By (2.98), (2.99), and applying Lemma 2.8 to $U(J)$, we can compute circuits $\{Q(z)\}_{z\in\mathcal{C}}$, each of which has depth $2d$ and acts on $m \leq 2N = O(\log^D n)$ qubits, and find a subset of qubits $S \subseteq [m]$ such that

$$\mathrm{Tr}_S(Q(z)|0^m\rangle\langle 0^m|Q(z)^\dagger) = \rho,\ \forall z \in \mathcal{C} \quad \text{and} \quad \max_{x\in\{0,1\}^m, z\in\mathcal{C}} |\langle x|Q(z)|0^m\rangle|^2 \geq n^{-2a}. \tag{2.101}$$

The algorithm proceeds as follows. First, we redefine $\varepsilon := n^{-2a}$ if $c < 2a$. Then, we construct the set of circuits $\{Q(z)\}_{z\in\mathcal{C}}$ and the set S satisfying (2.101) described above. The next step involves an exhaustive search over all the circuits in $\{Q(z)\}_{z\in\mathcal{C}}$. For each $z \in \mathcal{C}$, we invoke the algorithm from Theorem 2.3 of Section 2.1 on input circuit $Q(z)$ and error tolerance ε. The algorithm may output an error flag if it fails on the input circuit $Q(z)$. We terminate the exhaustive search over $z \in \mathcal{C}$ as soon as the algorithm succeeds. (2.101) shows that there is at least one $\bar{z} \in \mathcal{C}$ such that $Q(\bar{z})$ has peakedness at least n^{-2a}. Forcing $\varepsilon \leq n^{-2a}$ ensures that the condition $\varepsilon \leq 2m^{1/2}P_{\max}^{9/50}$ of Theorem 2.3 is satisfied for $Q(\bar{z})$ with $P_{\max} \geq n^{-2a}$.[5] Therefore, the search must terminate at some point. If the search terminates at some $z^* \in \mathcal{C}$, it outputs the classical description of an m-qubit state $|\phi\rangle$ described by a list of its nonzero coefficients in the computational basis that satisfies $|\langle\phi|Q(z^*)|0^m\rangle|^2 \geq 1-\varepsilon^2/4$. The state $\tau = \mathrm{Tr}_S(|\phi\rangle\langle\phi|)$ then satisfies the conditions of Theorem 2.5. In particular,

$$\begin{aligned}
\|\tau - \rho\|_1 &= \|\mathrm{Tr}_S(|\phi\rangle\langle\phi|) - \mathrm{Tr}_S(Q(z^*)|0^m\rangle\langle 0^m|Q(z^*)^\dagger)\|_1 && (2.102)\\
&\leq \||\phi\rangle\langle\phi| - Q(z^*)|0^m\rangle\langle 0^m|Q(z^*)^\dagger\|_1 && (2.103)\\
&= 2\sqrt{1 - |\langle\phi|Q(z^*)|0^m\rangle|^2} && (2.104)\\
&\leq \varepsilon. && (2.105)
\end{aligned}$$

[5] Note that $Q(\bar{z})$ is most likely not a peaked circuit since n is much bigger than $m = O(\log^D n)$.

Next, we analyze the runtime of the algorithm described above. The runtime of one run of the algorithm from Theorem 2.3 is

$$T(m, \varepsilon) = m^{O(\log m)+O(\log(1/\varepsilon))}. \tag{2.106}$$

Since $m = O(\log^D n)$ with a constant D, the term $O(\log m) = O(\log\log n)$ is negligible compared with $O(\log(1/\varepsilon)) = O(\log n)$. Therefore, one run of the algorithm takes

$$T(m, \varepsilon) = m^{O(\log(1/\varepsilon))} = n^{O(\log\log n)} \tag{2.107}$$

time. The total number of circuits in the set is $|\mathcal{C}| \leq N^{c\ln(n)} = n^{O(\log\log n)}$. Therefore, the total runtime of the exhaustive search can be upper bounded by $|\mathcal{C}|n^{O(\log\log n)} = n^{O(\log\log n)}$. Marginals of the found state τ can be computed in time polynomial in the number of nonzero entries of $|\phi\rangle$ in the standard basis, which is also $n^{O(\log\log n)}$. Therefore, the whole algorithm runs in $n^{O(\log\log n)}$ time.

Induction step Now suppose that Theorem 2.5 holds for all $(n, k-1, D)$-states. In the following, we show this implies that it holds for all (n, k, D)-states. Let ρ be an (n, k, D)-state, and let $\varepsilon = n^{-c}$ be given. Recall that Lemma 2.5 guarantees the existence of heavy slices arranged in a certain regular pattern. Moreover, whether or not a given rectangular region of width $O(d)$ is a heavy slice is determined by the marginal probability of measuring all-zeros in an $(n, k-1, D)$-state. Thus, by our inductive hypothesis, there exists an algorithm with runtime $n^{O(\log\log n)}$ to check whether or not a given rectangular region of width $O(d)$ is a heavy slice. In this way, we can compute the heavy slices from Lemma 2.5 with runtime $n^{O(\log\log n)}$.

Once we have computed the heavy slices, we use them to specify the Hermitian matrix σ from (2.83). We have already shown in (2.84) that

$$\|\mathcal{M}(\sigma) - \mathcal{M}(\rho)\|_1 \leq \frac{\varepsilon}{2}. \tag{2.108}$$

Let $J \subseteq [n]$ denote the support of ρ. Below, we show using our inductive hypothesis that

$$\|\mathcal{M}(\tau) - \mathcal{M}(\sigma)\|_1 \leq \frac{\varepsilon}{2} \tag{2.109}$$

for some Hermitian matrix τ such that marginals $\mathrm{Tr}(\tau(|x\rangle\langle x|)_S)$ for every $S \subseteq J$ and $x \in \{0,1\}^{|S|}$ can be computed in $n^{O(\log\log n)}$ time. Combining (2.108) and (2.109) with the triangle inequality gives the desired error bound (2.48).

To construct τ from σ, recall that each of the states $\sigma^j_{a_{j-1}a_j}$ in the tensor product (2.86) has the same measurement statistics as an $(n, k-1, D)$-state, in tensor product with some fresh qubits in the all-zeros state. That is,

$$\mathcal{M}(\sigma^j_{a_{j-1}a_j}) = \mathcal{M}(\tilde{\sigma}^j_{a_{j-1}a_j} \otimes |0^r\rangle\langle 0^r|) \tag{2.110}$$

for some $r \geq 0$ and $(n, k-1, D)$-state $\tilde{\sigma}^j_{a_{j-1}a_j}$. We apply the inductive hypothesis to the $(n, k-1, D)$-state $\tilde{\sigma}^j_{a_{j-1}a_j}$ with an error tolerance $\delta = 1/\mathrm{poly}(n)$ to obtain a Hermitian matrix $\tilde{\tau}^j_{a_{j-1}a_j}$. Letting $\tau^j_{a_{j-1}a_j} = \tilde{\tau}^j_{a_{j-1}a_j} \otimes |0^r\rangle\langle 0^r|$ and using (2.110), we get that

$$\|\mathcal{M}(\sigma^j_{a_{j-1}a_j}) - \mathcal{M}(\tau^j_{a_{j-1}a_j})\|_1 \leq \delta \tag{2.111}$$

and that there exists a classical algorithm with runtime $n^{O(\log\log n)}$ to exactly compute marginals $\mathrm{Tr}(\tau^j_{a_{j-1}a_j}(|x\rangle\langle x|)_S)$ for every S and $x \in \{0,1\}^{|S|}$. Our matrix τ is obtained by replacing each $\sigma^j_{a_{j-1}a_j}$ in (2.85) and (2.86) with $\tau^j_{a_{j-1}a_j}$, i.e.,

$$\tau = \sum_{\vec{a}} \bigotimes_{j=1}^{T} f_{a_j} \tau^j_{a_{j-1}a_j}. \tag{2.112}$$

We have already shown above that computing marginals of such a state τ can be reduced in poly(n)-time (via matrix multiplication) to computing marginals of the individual states $\tau^j_{a_{j-1}a_j}$, which takes runtime $n^{O(\log\log n)}$. To complete the proof, it remains to establish (2.109), which is obtained from the following Lemma by setting $\delta = (8TL^2)^{-1}\varepsilon$.

Lemma 2.9. *If (2.111) holds for every $j \in [T]$ and $a_j \in [L]$ for some $\delta \leq \frac{1}{2TL^2}$, then $\|\mathcal{M}(\tau) - \mathcal{M}(\sigma)\|_1 \leq 4TL^2\delta$.*

Proof. For reference, we restate the definition of σ from (2.83), (2.85), and (2.86):

$$\sigma = A^1 A^2 \cdots A^T(\rho) = \sum_{a_i=1 : i\in[T]}^{L} \bigotimes_{j=1}^{T} f_{a_j} \sigma^j_{a_{j-1}a_j} \tag{2.113}$$

where $A^1, \ldots, A^T$ are defined in (2.76). Let $\Delta^j_{a_{j-1}a_j} = \tau^j_{a_{j-1}a_j} - \sigma^j_{a_{j-1}a_j}$, so that (2.111) becomes

$$\|\mathcal{M}(\Delta^j_{a_{j-1}a_j})\|_1 \leq \delta. \tag{2.114}$$

For every $S \subseteq [T] \setminus \{\emptyset\}$, define

$$\overline{S} = \Big(\bigcup_{j\in S} \{j-1, j\} \Big) \setminus \{0\}. \tag{2.115}$$

We can write

$$\tau - \sigma = \sum_{\vec{a}} \bigotimes_{j=1}^{T} f_{a_j}(\sigma^j_{a_{j-1}a_j} + \Delta^j_{a_{j-1}a_j}) - \sum_{\vec{a}} \bigotimes_{j=1}^{T} f_{a_j}\sigma^j_{a_{j-1}a_j} \tag{2.116}$$

$$= \sum_{\vec{a}} \sum_{S\subseteq[T]\setminus\{\emptyset\}} \Big(\bigotimes_{j\in S} f_{a_j}\Delta^j_{a_{j-1}a_j} \Big) \otimes \Big(\bigotimes_{j\notin S} f_{a_j}\sigma^j_{a_{j-1}a_j} \Big) \tag{2.117}$$

$$= \sum_{S\subseteq[T]\setminus\{\emptyset\}} \sum_{a_i=1:i\in\overline{S}}^{L} \Big(\prod_{j\in\overline{S}\setminus S} f_{a_j} \bigotimes_{j\in S} f_{a_j}\Delta^j_{a_{j-1}a_j} \Big) \otimes \Big(\sum_{a_i=1:i\notin\overline{S}}^{L} \Big(\bigotimes_{j\in\overline{S}\setminus S} \sigma^j_{a_{j-1}a_j} \Big) \otimes \Big(\bigotimes_{j\notin\overline{S}} f_{a_j}\sigma^j_{a_{j-1}a_j} \Big) \Big).$$

Let $S \subseteq [T] \setminus \{\emptyset\}$. Suppose we fix the value of $a_i \in [L]$ for every $i \in \overline{S}$ and consider

$$\sum_{a_i=1:i\in[T]\setminus\overline{S}}^{L} \Big(\bigotimes_{j\in\overline{S}\setminus S} \sigma^j_{a_{j-1}a_j} \Big) \otimes \Big(\bigotimes_{j\in[T]\setminus\overline{S}} f_{a_j}\sigma^j_{a_{j-1}a_j} \Big). \tag{2.118}$$

Recall that for every $j \in [T]$, the index $a_j \in [L]$ specifies a subset $\omega \subseteq [\ell]$ corresponding to one term $\prod_{i\in[\ell]\setminus\omega} M_{s_i^j} \prod_{i\in\omega} E_{s_i^j}$ in the definition of A^j (see (2.76)). For every $j \in \overline{S}$, let $\mathcal{R}^j_{a_j}$ denote the term in the definition of A^j specified by a_j, so that

$$\mathcal{R}_{\vec{a}} = \mathcal{R}^1_{a_1} \otimes \mathcal{R}^2_{a_2} \cdots \otimes \mathcal{R}^T_{a_T}, \tag{2.119}$$

where $\mathcal{R}_{\vec{a}}$ is the CPTP map from (2.85) and (2.86).

By definition, we have

$$\left(\bigotimes_{j\in\overline{S}} \mathcal{R}^j_{a_j}\right)(\rho) = \left(\bigotimes_{j\in S} \sigma^j_{a_{j-1}a_j}\right) \otimes \tilde{\rho} \tag{2.120}$$

for some density matrix $\tilde{\rho}$. For every $j \in S$, let $\mathrm{Tr}^j_{a_{j-1}a_j}$ denote the operation of tracing out all the qubits in $\sigma^j_{a_{j-1}a_j}$. Note that

$$\tilde{\rho} = \mathcal{E}(\rho) \quad \text{with} \quad \mathcal{E} = \prod_{j\in S} \mathrm{Tr}^j_{a_{j-1}a_j}\left(\bigotimes_{j\in\overline{S}} \mathcal{R}^j_{a_j}\right) \tag{2.121}$$

and

$$\prod_{j\in[T]\setminus\overline{S}} A^j(\tilde{\rho}) = \sum_{a_i=1:i\in[T]\setminus\overline{S}}^{L} \left(\bigotimes_{j\in\overline{S}\setminus S} \sigma^j_{a_{j-1}a_j}\right) \otimes \left(\bigotimes_{j\in[T]\setminus\overline{S}} f_{a_j}\sigma^j_{a_{j-1}a_j}\right) \tag{2.122}$$

which is precisely (2.118). Now,

$$\left\| \prod_{j\in[T]\setminus\overline{S}} A^j(\tilde{\rho}) - \prod_{j\in[T]\setminus\overline{S}} M^j(\tilde{\rho}) \right\|_1 = \left\| \mathcal{E}\left(\prod_{j\in[T]\setminus\overline{S}} A^j(\rho) - \prod_{j\in[T]\setminus\overline{S}} M^j(\rho) \right) \right\|_1 \tag{2.123}$$

$$\leq \left\| \prod_{j\in[T]\setminus\overline{S}} A^j(\rho) - \prod_{j\in[T]\setminus\overline{S}} M^j(\rho) \right\|_1 \tag{2.124}$$

where in the second line, we used the fact that $\|\mathcal{E}(Q)\|_1 \leq \|Q\|_1$ for all Hermitian matrices Q, since $\mathcal{E}$ is a CPTP map. We can use Lemma 2.7 to upper bound the RHS of (2.124) by $\varepsilon/2$. Then, by the reverse triangle inequality,

$$\left\| \prod_{j\in[T]\setminus\overline{S}} A^j(\tilde{\rho}) \right\|_1 \leq \left\| \prod_{j\in[T]\setminus\overline{S}} M^j(\tilde{\rho}) \right\|_1 + \frac{\varepsilon}{2} = 1 + \frac{\varepsilon}{2} \leq 2. \tag{2.125}$$

Therefore,

$$\|\mathcal{M}(\tau) - \mathcal{M}(\sigma)\|_1 \leq \sum_{S\subseteq[T]\setminus\{\emptyset\}} \sum_{a_i=1:i\in\overline{S}}^{L} \left\| \mathcal{M}\left(\prod_{j\in\overline{S}\setminus S} f_{a_j} \bigotimes_{j\in S} f_{a_j}\Delta^j_{a_{j-1}a_j} \right) \otimes \mathcal{M}\left(\prod_{j\in[T]\setminus\overline{S}} A^j(\tilde{\rho}) \right) \right\|_1$$

$$\leq \sum_{S\subseteq[T]\setminus\{\emptyset\}} L^{|\overline{S}|}\delta^{|S|}2 \tag{2.126}$$

where we used (2.114), (2.125), and the fact that $\|\mathcal{M}(Q)\|_1 \leq \|Q\|_1$ for any Hermitian matrix Q. Continuing from (2.126), we arrive at

$$\|\mathcal{M}(\tau) - \mathcal{M}(\sigma)\|_1 \leq 2 \sum_{k=1}^{T} \binom{T}{k} L^{2k} \delta^k \tag{2.127}$$

$$= 2\left((1 + L^2\delta)^T - 1\right) \tag{2.128}$$

$$\leq 2TL^2\delta e^{TL^2\delta} \tag{2.129}$$

$$\leq 4TL^2\delta \tag{2.130}$$

where in the last line, we used $\delta \leq (2TL^2)^{-1}$ to infer $e^{TL^2\delta} \leq e^{1/2} \leq 2$. □

□

2.3 Open questions

Our work leaves several open questions.

1. Is there a poly(n)-time classical algorithm for simulating general peaked shallow circuits?

2. Can we approximate the output state $|\psi\rangle$ of a peaked shallow circuit, including the overall phase? Note that our algorithms only aim to approximate the projector $|\psi\rangle\langle\psi|$ onto the output state. As a consequence, our algorithms are not capable of recovering the sign $\pm$ of the expected values of Pauli observables, which are of great interest for variational quantum algorithms.

3. What is the complexity of simulating $\Theta(\log n)$-depth peaked circuits? Our analysis breaks down for such circuits since the backward lightcone of a single output qubit may span all n input qubits.

4. Can some of our techniques be adapted to study problems related to a local Hamiltonian with a peaked ground state?

Chapter 3

Measurement-induced entanglement in low-depth random Clifford circuits

In this chapter, we first demonstrate either numerically or rigorously the emergence of long-range measurement-induced entanglement (MIE) in the output states of low-depth random Clifford circuits in various settings. Recall that we say an n-qubit depth-d quantum circuit is low-depth if $d \leq O(\log n)$. Then in Section 3.6, we investigate the implications of long-range MIE for the computational complexity of classically simulating sampling from their output distributions. Some of the results presented in this chapter are from the paper [2].

Recall the setting of MIE described in Section 1.3.5, where we partition a set of n qubits into three non-empty and disjoint subsets, namely $[n] = A \cup B \cup C$. For the sake of discussion, consider $|\psi\rangle$ to be the output state of a 2D brickwork random Clifford circuit on a $\sqrt{n} \times \sqrt{n}$ grid, and suppose we have measured the qubits of $|\psi\rangle$ in B in the computational basis and obtained a pure post-measurement state $|\psi'\rangle$ on $A \cup C$. Recall that since $|\psi'\rangle$ is a stabilizer state, the presence or absence of MIE in $|\psi'\rangle$ does not depend on the measurement outcome. Let $\mathrm{dist}(A, C) = \min_{x \in A, y \in C} \mathrm{dist}(x, y)$ denote the minimum L1 distance between $x \in A$ and $y \in C$ on the 2D grid. Then, the presence of MIE should not be surprising when $|B|$ is very small since in that case, $\mathrm{dist}(A, C)$ cannot be that large. Intuitively, the amount of entanglement between A and C in $|\psi'\rangle$ should be monotonically decreasing in $\mathrm{dist}(A, C)$. Long-range MIE captures the situation where the entanglement between A and C can be lower bounded by some amount independent of $\mathrm{dist}(A, C)$.

To be more concrete, we need to look at a specific ABC tripartition. For the 2D case, we will consider two kinds of ABC tripartitions in this chapter. We are going to focus on the first kind for now, where we take A to be a single qubit and B to be a square region of qubits shielding A from the rest of the qubits forming the set $C \neq \emptyset$; see Figure 3.1 for an example where A is at the centre of the grid. We call this kind of tripartition an annulus tripartition because the region B has the shape of a square annulus. In an annulus tripartition, since $|A| = 1$, A is entangled with C in $|\psi'\rangle$ if and only if $\rho = \mathrm{Tr}_C(|\psi'\rangle\langle\psi'|)$ is the single-qubit maximally mixed state. Notice that with randomness over the choice of the circuit, the expected purity of ρ satisfies

$$\mathbb{E}[\mathrm{Tr}(\rho^2)] = 1 - \frac{1}{2} \Pr(A \text{ is entangled with } C \text{ in } |\psi'\rangle); \tag{3.1}$$

therefore, the closeness of $\mathbb{E}[\mathrm{Tr}(\rho^2)]$ to $1/2$ quantifies the likelihood of having MIE in $|\psi'\rangle$.

A major benefit of working in the Clifford setting is that we can simulate finite instances of the setup we wish to understand in software [35]. Figure 3.2 shows the results of numerically

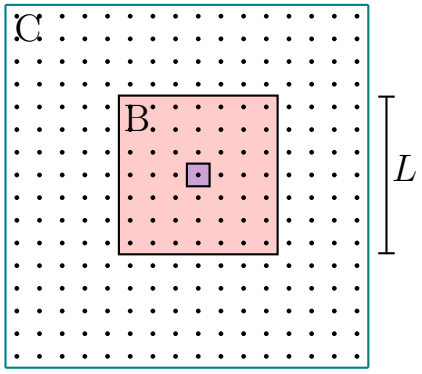


Figure 3.1: An example of an annulus tripartition. A 2D grid of qubits is partitioned into three non-empty and disjoint regions A, B, and C. The purple region indicates the single qubit in A. The pink region indicates the square shielding region B with side length L which in this example is $L = 7$. All the other qubits are in C.

estimating $\mathbb{E}[\mathrm{Tr}(\rho^2)]$ as a function of the grid side length $\sqrt{n}$ for 2D brickwork random Clifford circuits with circuit depth $d \in \{4, 5, 6, 7\}$, where we take A to contain the qubit at the centre and C to contain only the outermost ring of qubits, so $\mathrm{dist}(A, C)$ is maximal and grows as $\sqrt{n}$ increases. At $d \in \{4, 5\}$, we see that as the grid side length $\sqrt{n}$ increases from 11 to 101, with $\mathrm{dist}(A, C)$ increasing proportionally, $\mathbb{E}[\mathrm{Tr}(\rho^2)]$ crawls towards one, implying that the probability of observing MIE in $|\psi'\rangle$ decays to zero. However, a qualitatively different behaviour is observed at $d \in \{6, 7\}$. Surprisingly, for $d \in \{6, 7\}$, the numerics indicate that $\mathbb{E}[\mathrm{Tr}(\rho^2)]$ stays flat, so MIE between A and C in $|\psi'\rangle$ persists, even as $\sqrt{n}$ increases. The numerics predict that for this specific architecture and tripartition combination, a long-range MIE phase transition occurs at $d^* = 6$.

Our numerical findings presented in Figure 3.2 motivate the following definition.

Property 3.1 (Annulus tripartition long-range MIE)**.** A distribution $\mathcal{D}$ of n-qubit 2D random Clifford circuits satisfies the long-range MIE property with respect to the annulus tripartition if there exists a constant $c \in (1/2, 1)$ such that

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[\mathrm{Tr}(\rho^2)] < c \tag{3.2}$$

holds for every annulus tripartition $[n] = A \cup B \cup C$, where for $U \sim \mathcal{D}$, $|\psi'\rangle$ is the post-measurement state obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in B in the computational basis and $\rho = \mathrm{Tr}_C(|\psi'\rangle\langle\psi'|)$.

We remark that requiring c to be a constant independent of n is the best possible (e.g., requiring $c = 1/2 + 1/\mathrm{poly}(n)$ is impossible) for a constant depth-d circuit composed of 2-qubit gates, as there is some constant probability depending only on d that all the gates acting on the qubit in A are the identity gate. Thus, for an annulus tripartition, there is no MIE in $|\psi'\rangle$ with at least some probability depending on d.

Given the evidence in Figure 3.2, we conjecture that 2D brickwork random Clifford circuits satisfy Property 3.1 whenever the circuit depth is at least 7. Unfortunately, we did not succeed in proving this conjecture for the 2D brickwork architecture. Instead, we are

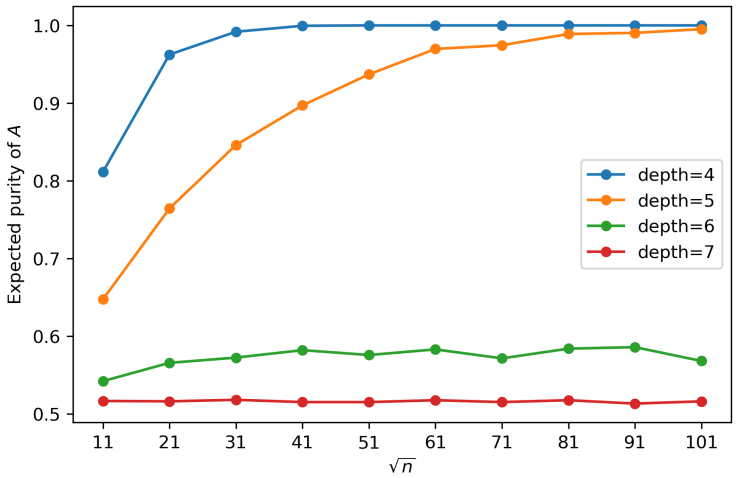


Figure 3.2: Numerical estimation of MIE in the output states of 2D brickwork random Clifford circuits (Figure 1.2) in an annulus tripartition. Here, A consists of only the qubit at the centre of the grid, C contains all the qubits on the boundary of the grid, and B contains all the other qubits. Each data point is estimated from 1040 circuit instance samples. We observe a qualitative change in the behaviour of MIE, transitioning from decaying to long-range as the circuit depth d increases beyond a critical value of $d^* = 6$.

able to prove that random Clifford circuits drawn according to a certain 2D coarse-grained architecture satisfy Property 3.1. We introduce the coarse-grained architectures in the next section.

3.1 The coarse-grained architectures

There has been a common recognition that many properties of geometrically local low-depth random quantum circuits, including long-range MIE, could become significantly easier to analyze with respect to a coarse-grained architecture [58, 59, 64]. To see the idea, consider the circuit diagram of the 1D coarse-grained architecture shown below in Figure 3.3. The 1D coarse-grained architecture consists of exactly two layers of coarse-grained gates, each acting on k qubits for some even $k \geq 2$, arranged like the depth-2 1D brickwork architecture. While the nominal circuit depth d of a 1D coarse-grained circuit is 2 by definition, it will become clear that the main idea is to let the gate size parameter k play the operational role of the circuit depth in a usual circuit composed of 2-qubit gates. For example, k dictates the lightcone structure of coarse-grained architectures, just like d does for architectures composed of 2-qubit gates.

The relationship between the 1D coarse-grained architecture with gate size k and the 1D brickwork architecture with circuit depth $d = O(k)$ can be made more concrete. In the Clifford case, it is known that an arbitrary k-qubit Clifford unitary can be implemented exactly in the 1D brickwork architecture with depth $7k - 4$ [82, 83]. Thus, by compiling each k-qubit gate in the 1D coarse-grained architecture down to an $O(k)$-depth 1D circuit composed of 2-qubit gates, an entire 1D coarse-grained circuit can be implemented in the 1D

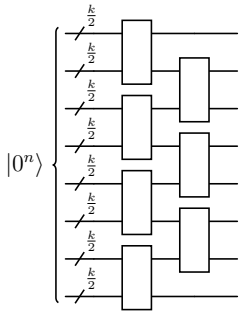


Figure 3.3: Circuit diagram of the 1D coarse-grained architecture. The circuit acts on n qubits, and each gate acts on k qubits for some even $k \geq 2$. The topological arrangement of the gates are the same as the depth-2 1D brickwork architecture.

brickwork architecture with depth $O(k)$. However, the 1D brickwork circuit obtained in this way no longer qualifies as a random quantum circuit because each 2-qubit gate is no longer sampled independently from some gate distribution. Since the intention of considering 1D coarse-grained random quantum circuits is to facilitate our understanding of 1D brickwork random quantum circuits, we need a different compilation procedure. An alternative is to replace each random k-qubit Clifford gate with an $O(k)$-depth 1D brickwork random Clifford circuit realizing an ε-approximate unitary 2-design in the sense of Definition 1.3 for some $\varepsilon = 1/\exp(k)$. This compilation procedure converts a 1D coarse-grained random Clifford circuit with gate size k to an $O(k)$-depth 1D brickwork random Clifford circuit with some gates deleted. For problems that reduce to bounding expressions of the form $\mathbb{E}_{U\sim\mathcal{D}}[\mathrm{Tr}(BU^{\otimes 2}A(U^\dagger)^{\otimes 2})]$, this connection can be formalized.

The 1D coarse-grained architecture can be generalized naturally to 2D, with each gate acting on a $k = (\tau \times \tau)$ patch of qubits where τ is even. Figure 3.4 below depicts an example on a grid of $n = 16 \times 16$ qubits, with each coarse-grained gate acting on a square of $k = \tau^2 = 16$ qubits.

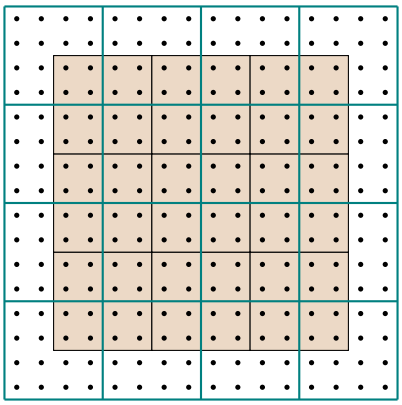


Figure 3.4: Gate arrangement of a 2D coarse-grained circuit, which has two layers of coarse-grained gates. In this example, $\tau = 4$. The gates with green boundaries represent one layer of 16 gates. The shaded region denotes the support of another layer of 9 gates.

Notice that in the caption of Figure 3.4, it does not specify which layer of gates is the first layer. It turns out that we will consider both cases later in this chapter, depending on the scenario. Partially as a warm up exercise, we will demonstrate the technical advantage of working with coarse-grained architectures in the next section by studying long-range MIE in 1D coarse-grained random Clifford circuits.

3.2 Long-range MIE in 1D coarse-grained random Clifford circuits

In this section, we will prove the following theorem. Let $n \geq 1$ be an integer.

Theorem 3.1. *Let $\mathcal{D}$ denote the distribution of n-qubit 1D coarse-grained random Clifford circuits with gate size k. Define the tripartition $[n] = A \cup B \cup C$ where A contains the first $k/2$ qubits, C contains the last $k/2$ qubits, and B contains all the other qubits. For $U \sim \mathcal{D}$, let $|\psi'\rangle$ denote the post-measurement state obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in B in the computational basis. Let E denote the number of ebits of entanglement between A and C in $|\psi'\rangle$. For $k = c\log_2(n)$, it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[E] \geq \begin{cases} 0 & \text{if } c < 2, \\ \log_2(\log_2(n)) & \text{if } c = 2, \\ \left(\frac{c}{2} - 1\right)\log_2(n) & \text{if } c > 2. \end{cases} \tag{3.3}$$

Theorem 3.1 predicts that a sharp transition in the behaviour of MIE in this setup occurs at $k = 2\log_2(n)$. The remainder of this section is devoted to proving Theorem 3.1.

We first recall a few definitions from Section 1.3. For every integer $m \geq 1$, let $\mathcal{P}^m = \{I, X, Y, Z\}^{\otimes m}$ denote the set of all m-qubit Pauli operators with $+$ sign. For every set S consisting of Pauli operators, we define $\pm S = S \cup \{-P : P \in S\}$. For an n-qubit Pauli operator $P \in \pm\mathcal{P}^n$, we define $\mathrm{Supp}(P) \subseteq [n]$ to be the subset of qubits on which P takes a value other than the identity.

For every n-qubit stabilizer state $|\psi\rangle$, we use

$$S_\psi = \{P \in \pm\mathcal{P}^n : P|\psi\rangle = |\psi\rangle\} \tag{3.4}$$

to denote the stabilizer group of $|\psi\rangle$, and for every $A \subseteq [n]$, we use

$$S_{\psi_A} = \{P \in S_\psi : \mathrm{Supp}(P) \subseteq A\} \tag{3.5}$$

to denote the subgroup of S_ψ local on A. We first establish two useful technical lemmas.

Lemma 3.1. *Let $\mathcal{D}$ denote the uniform distribution over n-qubit stabilizer states. It holds that*

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}}[|S_{\psi_A}|] = 1 + \frac{4^{|A|} - 1}{2^n + 1} \leq \begin{cases} 1 + 2^{-\frac{n}{2}} & \text{if } |A| = \frac{n}{4}, \\ 2 & \text{if } |A| = \frac{n}{2}, \\ 1 + 2^{\frac{n}{2}} & \text{if } |A| = \frac{3n}{4}. \end{cases} \tag{3.6}$$

Proof. This follows from

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}}[|S_{\psi_A}|] = \sum_{P\in\mathcal{P}^{|A|}} \Pr_{|\psi\rangle\sim\mathcal{D}}(\pm P_A \in S_\psi) = \sum_{P\in\mathcal{P}^{|A|}} \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}}[\langle\psi|P_A|\psi\rangle^2] = 1 + \frac{4^{|A|}-1}{2^n+1} \tag{3.7}$$

where for every $P \in \mathcal{P}^n \setminus \{I^{\otimes n}\}$,

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}}[\langle\psi|P|\psi\rangle^2] = \frac{1}{2^n+1} \tag{3.8}$$

is a standard fact (see Lemma 1.2). □

Lemma 3.2. *Suppose we start with a tripartite stabilizer state $|\psi\rangle$ over subsystems A, B, and C, apply a uniformly random $|B|$-qubit Clifford unitary U to B, and then measure all qubits in B in the computational basis to get a pure post-measurement state $|\psi'\rangle$ on A and C. It holds that*

$$\mathop{\mathbb{E}}_{U}[|S_{\psi'_A}|] \le |S_{\psi_A}| + \frac{2^{|A|}}{2^{|C|}}|S_{\psi_C}|. \tag{3.9}$$

In particular, if $|A| = |C|$, then

$$\mathop{\mathbb{E}}_{U}[|S_{\psi'_A}|] \le |S_{\psi_A}| + |S_{\psi_C}|. \tag{3.10}$$

Proof. Since we measure all the qubits in B in the computational basis, we have

$$\mathop{\mathbb{E}}_{U}[|S_{\psi'_A}|] \le \sum_{P\in\mathcal{P}^{|A|}} \sum_{s\in\{0,1\}^{|B|}} \mathop{\mathbb{E}}_{U} \left[\langle\psi|(I\otimes U^\dagger\otimes I)(P\otimes Z(s)\otimes I)(I\otimes U\otimes I)|\psi\rangle^2\right]. \tag{3.11}$$

We can isolate the part conjugated by the random unitary on B

$$\sum_{s\in\{0,1\}^{|B|}} \mathop{\mathbb{E}}_{U} \left[(U^\dagger\otimes U^\dagger)(Z(s)\otimes Z(s))(U\otimes U)\right] \tag{3.12}$$

$$= 2^{|B|} \sum_{x\in\{0,1\}^{|B|}} \mathop{\mathbb{E}}_{U} \left[(U^\dagger\otimes U^\dagger)(|x\rangle\langle x|\otimes|x\rangle\langle x|)(U\otimes U)\right] \tag{3.13}$$

$$= 4^{|B|} \mathop{\mathbb{E}}_{U} \left[(U^\dagger\otimes U^\dagger)(|0^{|B|}\rangle\langle 0^{|B|}|\otimes|0^{|B|}\rangle\langle 0^{|B|}|)(U\otimes U)\right] \tag{3.14}$$

$$= \frac{4^{|B|}}{4^{|B|}+2^{|B|}}(I + \mathrm{SWAP}_{BB'}) \tag{3.15}$$

where the last equality uses the formula for the second moment of Haar random states (see for example Lemma 2.3 of [84]) and the fact that the uniform distribution over the set of all $|B|$-qubit stabilizer states forms an exact state 2-design. Thus,

$$\mathop{\mathbb{E}}_{U}[|S_{\psi'_A}|] \le \sum_{P\in\mathcal{P}^{|A|}} \langle\psi|P_A|\psi\rangle^2 + (\langle\psi|_{ABC}\langle\psi|_{A'B'C'})(P_A P_{A'}\mathrm{SWAP}_{BB'})(|\psi\rangle_{ABC}|\psi\rangle_{A'B'C'}) \tag{3.16}$$

$$= |S_{\psi_A}| + \sum_{P\in\mathcal{P}^{|A|}} (\langle\psi|_{ABC}\langle\psi|_{A'B'C'})(P_A P_{A'}\mathrm{SWAP}_{BB'})(|\psi\rangle_{ABC}|\psi\rangle_{A'B'C'}). \tag{3.17}$$

To handle the second term with SWAP, we will make use of the 2-qubit identity

$$\mathrm{SWAP} = \frac{1}{2}(I \otimes I + X \otimes X + Y \otimes Y + Z \otimes Z). \tag{3.18}$$

Using the above identity, we get

$$\sum_{P \in \mathcal{P}^{|A|}} (\langle\psi|_{ABC}\langle\psi|_{A'B'C'})(P_A P_{A'} \mathrm{SWAP}_{BB'})(|\psi\rangle_{ABC}|\psi\rangle_{A'B'C'}) \tag{3.19}$$

$$= 2^{|A|}(\langle\psi|_{ABC}\langle\psi|_{A'B'C'})(\mathrm{SWAP}_{AA'}\mathrm{SWAP}_{BB'})(|\psi\rangle_{ABC}|\psi\rangle_{A'B'C'}) \tag{3.20}$$

$$= 2^{|A|}(\langle\psi|_{ABC}\langle\psi|_{A'B'C'})\mathrm{SWAP}_{CC'}(|\psi\rangle_{ABC}|\psi\rangle_{A'B'C'}). \tag{3.21}$$

Using the identity again, we get

$$(3.21) = \frac{2^{|A|}}{2^{|C|}} \sum_{P \in \mathcal{P}^{|C|}} \langle\psi|P_C|\psi\rangle^2 = \frac{2^{|A|}}{2^{|C|}}|S_{\psi_C}|. \tag{3.22}$$

□

Next, we prove Theorem 3.1. Our strategy is to apply Lemma 3.2 iteratively in the following way. Recall the setup where we have an n-qubit 1D coarse-grained random Clifford circuit with exactly two layers of coarse-grained gates, each acting on k qubits. We call the top $k/2$ qubits A and the bottom $k/2$ qubits C. We apply the circuit and then measure all qubits in $[n] \setminus (A \cup C)$ in the computational basis to obtain a k-qubit pure post-measurement state $|\psi'\rangle$ on A and C. Figure 3.5 is a circuit diagram depicting the setup.

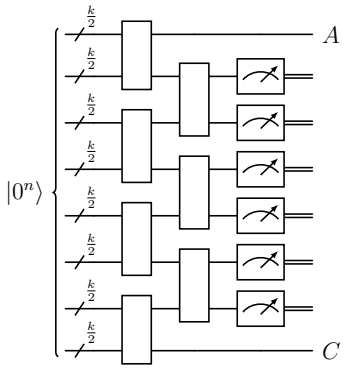


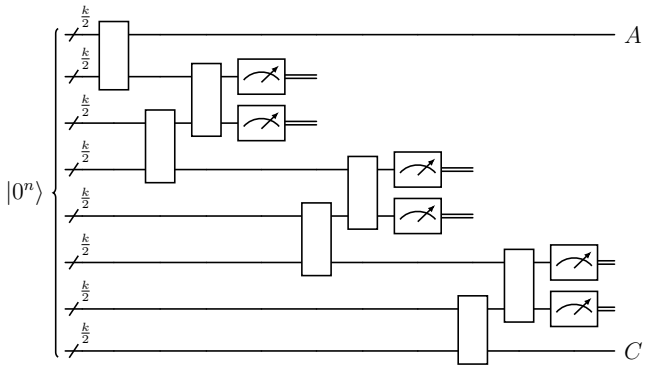
Figure 3.5: Circuit diagram for probing MIE in 1D coarse-grained random Clifford circuits.

We are interested in understanding the amount of MIE between A and C in $|\psi'\rangle$ as a function of n and k. Recall that we use $S_{\psi'_A}$ to denote the stabilizer subgroup of $|\psi'\rangle$ local on A. Since there is entanglement between A and C in $|\psi'\rangle$ if and only if $|S_{\psi'_A}| < 2^{k/2}$, our goal is to upper bound $\mathbb{E}[|S_{\psi'_A}|]$ with randomness over all the coarse-grained gates in the circuit. We will prove the following lemma.

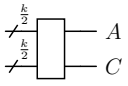
Lemma 3.3. *Under the setup described above, it holds that*

$$\mathbb{E}[|S_{\psi'_A}|] \leq \frac{2n}{k}. \tag{3.23}$$

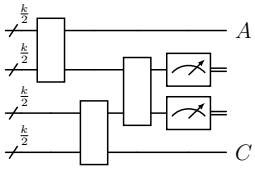
Proof. By delaying the gates and measurements, the circuit diagram in Figure 3.5 is equivalent to the following circuit diagram.



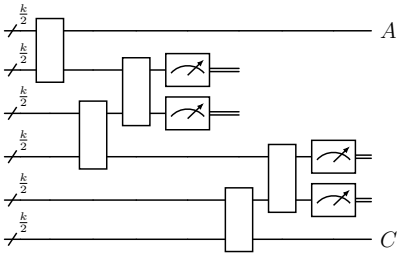
We can view the above circuit as starting with



and then iteratively build up the circuit by adding gates and measurements in a staircase pattern



and then



until eventually reaching the complete circuit shown in Figure 3.5.

Throughout this process, the set of qubits A remains fixed as the same top $k/2$ qubits, while the set of qubits C is redefined as the new bottom $k/2$ qubits in each iteration. We can upper bound $\mathbb{E}[|S_{\psi'_A}|]$ at the end by upper bounding the increment of this quantity in each iteration. Notice that the total number of iterations is equal to the number of second-layer gates, which is $n/k - 1$. For every $j = 0, \ldots, n/k - 1$, let $f(j)$ denote the value of $\mathbb{E}[|S_{\psi'_A}|]$ after the j-th iteration. By the second case of Lemma 3.1, we have $f(0) \leq 2$. For every

$j = 1, \ldots, n/k - 1$, by Lemma 3.2, we deduce that $f(j) \leq f(j-1) + 2$. Therefore, after the last iteration, we have

$$\mathbb{E}[|S_{\psi'_A}|] \leq 2 + 2\left(\frac{n}{k} - 1\right) = \frac{2n}{k}. \tag{3.24}$$

□

Proof of Theorem 3.1. For $k = c\log_2(n)$, by Lemma 3.3, we have $\mathbb{E}[|S_{\psi'_A}|] \leq \frac{2n}{c\log_2(n)}$. By Jensen's inequality, we get

$$\mathbb{E}[\dim(S_{\psi'_A})] = \mathbb{E}[\log_2(|S_{\psi'_A}|)] \leq \log_2(\mathbb{E}[|S_{\psi'_A}|]) \leq 1 + \log_2(n) - \log_2(c) - \log_2(\log_2(n)). \tag{3.25}$$

Therefore,

$$\mathbb{E}[E] = \frac{k}{2} - \mathbb{E}[\dim(S_{\psi'_A})] \geq \left(\frac{c}{2} - 1\right)\log_2(n) + \log_2(\log_2(n)) + \log_2(c) - 1. \tag{3.26}$$

□

3.3 Long-range MIE in 2D coarse-grained random Clifford circuits in annulus tripartitions

Recall that when we introduced the 2D coarse-grained architecture in Section 3.1, we left open the possibility of defining either layer of gates as the first layer. In this section, we make a choice so that whenever we refer to an n-qubit 2D coarse-grained random Clifford circuit U with $k = (\tau \times \tau)$-qubit gates, we mean a random Clifford circuit of the form

$$U = \underbrace{(W_{m^2} \otimes \cdots \otimes W_1)}_{W} \cdot \underbrace{(V_{(m-1)^2} \otimes \cdots \otimes V_1)}_{V} \tag{3.27}$$

acting on a $\sqrt{n} \times \sqrt{n}$ square grid of qubits where $m^2 = n/\tau^2$. Under this definition, in the example shown in Figure 3.4, $m = 4$, the first layer V contains 9 gates, and the second layer W contains 16 gates. The goal of this section is to prove the following theorem.

Theorem 3.2. *For sufficiently large n, the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $k = (\tau \times \tau)$-qubit gates with $\tau \geq \sqrt{1000\log_2(n)}$ satisfies Property 3.1.*

We conjecture that Property 3.1 already holds for 2D coarse-grained random Clifford circuits when τ is a sufficiently large constant. However, we are only able to prove it for $\tau \geq \sqrt{1000\log_2(n)}$. We will also obtain a corollary that 2D coarse-grained random Clifford circuits with $\tau \geq \sqrt{1000\log_2(n)}$ satisfy the anti-concentration property Definition 1.5.

To work towards proving Theorem 3.2, we will establish a series of lemmas that apply more generally to any tripartition $[n] = A \cup B \cup C$ with $|A| = 1$ and will only invoke the fact that $|C| \geq \sqrt{n}$ for an annulus tripartition at the very end. we first establish the condition that characterizes the existence of MIE for a tripartition $[n] = A \cup B \cup C$ such that $|A| = 1$. Let $|\psi\rangle$ be an n-qubit stabilizer state. Suppose we measure all qubits of $|\psi\rangle$ in B in the

computational basis and obtain a pure post-measurement state $|\psi'\rangle$ on $A \cup C$. Since $|A| = 1$, it is not hard to see that there is no entanglement between A and C in $|\psi'\rangle$ if and only if $|\psi'\rangle$ has a non-identity stabilizer local on A if and only if $|\psi\rangle$ has a stabilizer that is non-identity on A, Z-type on B, and the identity on C. The following lemma rephrases the above observation when $|\psi\rangle$ is the output state of a circuit.

Lemma 3.4. *Let $[n] = A \cup B \cup C$ be a tripartition such that $|A| = 1$. Let U be an n-qubit Clifford circuit and define $|\psi\rangle = U|0^n\rangle$. Suppose we measure all qubits of $|\psi\rangle$ in B in the computational basis and obtain a pure post-measurement state $|\psi'\rangle$ on $A \cup C$. For $P \in \{X, Y, Z\}$, define*

$$D_P = \{s \in \{0,1\}^n : UZ(s)U^\dagger = (-1)^\alpha P_A Z(r)_B \text{ for some } r \in \{0,1\}^{|B|} \text{ and } \alpha \in \{0,1\}\}$$

and define

$$D = D_X \cup D_Y \cup D_Z. \tag{3.28}$$

Then it holds that there is entanglement between A and C in $|\psi'\rangle$ if and only if $D = \emptyset$.

Proof. Let $S_\psi = \{UZ(s)U^\dagger : s \in \{0,1\}^n\}$ denote the stabilizer group of $|\psi\rangle = U|0^n\rangle$.

First suppose $D \neq \emptyset$. Then for some $P \in \{X, Y, Z\}$, $r \in \{0,1\}^{|B|}$, and $\alpha \in \{0,1\}$, $(-1)^\alpha P_A Z(r)_B \in S_\psi$. After measuring all qubits of $|\psi\rangle$ in B in the computational basis, the post-measurement state $|\psi'\rangle$ is stabilized by $\pm P_A$ for some choice of ± 1 sign determined by the measurement outcomes. In other words, the state $\rho = \mathrm{Tr}_C(|\psi'\rangle\langle\psi'|)$ is a pure 1-qubit stabilizer state, so there is no entanglement between A and C in $|\psi'\rangle$.

Next, suppose there is no entanglement between A and C in $|\psi'\rangle$, so $\rho = \mathrm{Tr}_C(|\psi'\rangle\langle\psi'|)$ is pure. Suppose $x \in \{0,1\}^{|B|}$ was the measurement outcome. The stabilizer group of the state $|\psi'\rangle_{A\cup C} \otimes |x\rangle_B$ is generated by

$$G = \{(-1)^{x_j} Z_j : j \in B\} \cup G' \tag{3.29}$$

where

$$G' = \{Q \in S_\psi : Q \text{ is } Z\text{-type on } B\}. \tag{3.30}$$

Since ρ is pure, it must have a stabilizer P for some $P \in \pm\{X, Y, Z\}$. This stabilizer can be expressed as a product of elements from the set of generators above. That is, for some $T \subseteq G'$ and $R \subseteq B$, we have

$$P_A = \prod_{Q \in T} Q \prod_{j \in R} (-1)^{x_j} Z_j. \tag{3.31}$$

Rearranging the above equation, we conclude that

$$(-1)^{\sum_{j\in R} x_j} P_A Z(r)_B = \prod_{Q \in T} Q = UZ(s)U^\dagger \tag{3.32}$$

for $r \in \{0,1\}^{|B|}$ such that $r_j = 1$ if and only if $j \in R$ and some binary string $s \in \{0,1\}^n$. Therefore, $s \in D$ and $D \neq \emptyset$. □

Given Lemma 3.4, our strategy is to upper bound $\mathbb{E}_U[|D|]$ with randomness over the choice of the 2D coarse-grained random Clifford circuit U. Given the definition of D, upper bounding $\mathbb{E}_U[|D|]$ reduces to upper bounding the probability that $|\psi\rangle = U|0^n\rangle$ has a stabilizer that is non-identity on A, Z-type on B, and the identity on C. In the Heisenberg picture, this reduces to upper bounding the probability that $U^\dagger$ conjugates a Pauli operator supported on $A \cup B$ to a Z-type Pauli operator.

We remark that a condition more restrictive than the one given in Lemma 3.4 is possible. With respect to an n-qubit Clifford circuit U, we say a subset of qubits $S \subseteq [n]$ is lightcone connected if for every bipartition $S = A \cup B$ with A and B being non-empty and disjoint, it holds that $\mathrm{blc}(A) \cap \mathrm{blc}(B) \neq \emptyset$ (recall the definition of backward lightcones in Section 1.3.4). We can define $\tilde{D} = \{s \in D : \mathrm{Supp}(UZ(s)U^\dagger) \text{ is lightcone connected}\}$. Recall that for every $P \in \pm\mathcal{P}^n$, $\mathrm{Supp}(P) \subseteq [n]$ denotes the subset of qubits on which P takes a value other than the identity. Clearly, if $\tilde{D} \neq \emptyset$, then $D \neq \emptyset$. It is also not hard to see that if $D \neq \emptyset$, then $\tilde{D} \neq \emptyset$, so it suffices to upper bound the smaller quantity $\mathbb{E}_U[|\tilde{D}|]$. That being said, we will work with the less refined set D defined in Lemma 3.4.

To proceed, we will need the following technical lemmas, which will be useful for analyzing the effect of each individual $k = \tau^2$-qubit gate in the first layer of a 2D coarse-grained random Clifford circuit. For every k-qubit Pauli operator $P \in \mathcal{P}^k$ written as $P = P_1 \otimes \cdots \otimes P_k$ and every subset $A \subseteq [k]$, we define the truncation of P to A to be the $|A|$-qubit Pauli operator $P_A = \bigotimes_{j \in A} P_j$.

Lemma 3.5. *Let $\mathcal{D}$ denote the uniform distribution over $\mathcal{P}^k \setminus \{I^{\otimes k}\}$. Let $A \subseteq [k]$ and define $B = [k] \setminus A$. For every $P \in \mathcal{P}^k$, let P_A and P_B denote the truncations of P to A and B respectively. Let $\emptyset \neq \mathcal{E} \subseteq \mathcal{P}^{|B|}$. Then it holds that*

$$\Pr_{P \sim \mathcal{D}}(P_A = I^{\otimes |A|} \mid P_B \in \mathcal{E}) \leq 4^{-|A|}. \tag{3.33}$$

Proof. By the definition of $\mathcal{D}$, for every $W \in \mathcal{P}^k \setminus \{I^{\otimes k}\}$, we have

$$\Pr_{P \sim \mathcal{D}}(P = W) = \frac{1}{4^k - 1}. \tag{3.34}$$

Since $\mathcal{E} \neq \emptyset$, either $I^{\otimes |B|} \in \mathcal{E}$ or $I^{\otimes |B|} \notin \mathcal{E}$. First consider the case where $I^{\otimes |B|} \notin \mathcal{E}$. In this case, we have

$$\Pr_{P \sim \mathcal{D}}(P_A = I^{\otimes |A|} \mid P_B \in \mathcal{E}) = \frac{\Pr_{P \sim \mathcal{D}}(P_A = I^{\otimes |A|} \text{ and } P_B \in \mathcal{E})}{\Pr_{P \sim \mathcal{D}}(P_B \in \mathcal{E})} \tag{3.35}$$

$$= \frac{|\mathcal{E}| \left(\frac{1}{4^k - 1}\right)}{|\mathcal{E}| 4^{|A|} \left(\frac{1}{4^k - 1}\right)} = 4^{-|A|}. \tag{3.36}$$

Secondly, for the case where $I^{\otimes|B|} \in \mathcal{E}$, we have

$$\Pr_{P\sim\mathcal{D}}(P_A = I^{\otimes|A|} \mid P_B \in \mathcal{E}) \tag{3.37}$$

$$= \frac{\Pr_{P\sim\mathcal{D}}(P_A = I^{\otimes|A|} \text{ and } P_B \in \mathcal{E})}{\Pr_{P\sim\mathcal{D}}(P_B \in \mathcal{E})} \tag{3.38}$$

$$= \frac{\Pr_{P\sim\mathcal{D}}(P_A = I^{\otimes|A|} \text{ and } P_B = I^{\otimes|B|}) + \Pr_{P\sim\mathcal{D}}(P_A = I^{\otimes|A|} \text{ and } P_B \in \mathcal{E}\setminus\{I^{\otimes|B|}\})}{\Pr_{P\sim\mathcal{D}}(P_B \in \mathcal{E})} \tag{3.39}$$

$$= \frac{(|\mathcal{E}|-1)(\frac{1}{4^k-1})}{(4^{|A|}|\mathcal{E}|-1)(\frac{1}{4^k-1})} \leq 4^{-|A|}. \tag{3.40}$$

□

Note that in the above lemma, taking $\mathcal{E} = \mathcal{P}^{|B|}$ corresponds to no conditioning.

Lemma 3.6. *Let $\mathcal{C}$ denote the uniform distribution over k-qubit Clifford unitaries, and let $\mathcal{D}$ denote the uniform distribution over $\mathcal{P}^k \setminus \{I^{\otimes k}\}$. Let $r \in \{1,2,3,4\}$. For every $i \in [r]$, for $P_i \sim \mathcal{D}$, we write $P_i = E_i \otimes F_i$ such that $E_i \in \mathcal{P}^{k/4}$ and $F_i \in \mathcal{P}^{3k/4}$. Define $P_A = \bigotimes_{i=1}^r E_i$, $P_B = \bigotimes_{i=1}^r F_i$, and $h = \frac{rk}{4}$. Let $\emptyset \neq \mathcal{E} \subseteq \mathcal{P}^{3h}$. Then it holds that*

1. *for $r = 4$,*
$$\Pr_{P_1,\ldots,P_4\sim\mathcal{D},C\sim\mathcal{C}}(C^\dagger P_A C \text{ is } Z\text{-type} \mid P_B \in \mathcal{E}) \leq 2^{-k}; \tag{3.41}$$
2. *for $r \in \{1,2,3\}$,*
$$\Pr_{P_1,\ldots,P_r\sim\mathcal{D},C\sim\mathcal{C}}(C^\dagger (P_A \otimes I^{\otimes k-h}) C \text{ is } Z\text{-type} \mid P_B \in \mathcal{E}) \leq 4^{-h} + 2^{-k}. \tag{3.42}$$

Proof. Recall that for every $W \in \mathcal{P}^k \setminus \{I^{\otimes k}\}$, we have (see Lemma 1.2)

$$\Pr_{C\sim\mathcal{C}}\left(C^\dagger W C \text{ is } Z\text{-type}\right) = \mathop{\mathbb{E}}_{C\sim\mathcal{C}}\left[\langle 0^k|C^\dagger W C|0^k\rangle^2\right] = \frac{1}{2^k+1}. \tag{3.43}$$

By the previous lemma and the independence among $F_1,\ldots,F_r$, we have

$$\Pr_{P_1,\ldots,P_r\sim\mathcal{D}}(P_A = I^{\otimes k} \mid P_B \in \mathcal{E}) \leq (4^{-\frac{k}{4}})^r = 4^{-h}. \tag{3.44}$$

Thus, by the law of total probability and then convexity, we get that

$$\Pr_{P_1,\ldots,P_r\sim\mathcal{D},C\sim\mathcal{C}}(C^\dagger (P_A \otimes I^{\otimes k-h}) C \text{ is } Z\text{-type} \mid P_B \in \mathcal{E}) \tag{3.45}$$

$$= \Pr_{P_1,\ldots,P_r\sim\mathcal{D}}(P_A = I^{\otimes k} \mid P_B \in \mathcal{E}) + \Pr_{P_1,\ldots,P_4\sim\mathcal{D}}(P_A \neq I^{\otimes k} \mid P_B \in \mathcal{E})\left(\frac{1}{2^k+1}\right) \tag{3.46}$$

$$\leq 4^{-h} + \left(1-4^{-h}\right)\left(\frac{1}{2^k+1}\right). \tag{3.47}$$

Notice that if $r = 4$, then $h = k$ and (3.47) $= 2^{-k}$. Otherwise, for $r \in \{1,2,3\}$, we get (3.47) $\leq 4^{-h} + 2^{-k}$. □

For any $P \in \mathcal{P}^n$ and n-qubit 2D coarse-grained random Clifford circuit U, it is relatively easy to see that the probability of $U^\dagger PU$ being Z-type is a function of the overlap of $\mathrm{Supp}(P)$ with the second layer of gates $W_1, \ldots, W_{m^2}$. To keep track of this overlap, for every $P \in \mathcal{P}^n$, we let $\mathcal{G}(P)$ denote the set of all second layer gates with non-trivial overlap with $\mathrm{Supp}(P)$, and we call $\mathcal{G}(P)$ a cluster, which can be specified by filling in the squares of an $m \times m$ grid. This grid representation leads naturally to the notion of the *size* and *perimeter* of a cluster $\mathcal{G}(P)$, which we define next.

Definition 3.1. Let $\mathcal{G}(P)$ be a cluster and identify it with a subset of the $m \times m$ grid in the natural way. Then we define the *size* of $\mathcal{G}(P)$, denoted $|\mathcal{G}(P)|$, to be the number of grid squares contained in the cluster. We define the *perimeter* of $\mathcal{G}(P)$, denoted $\mathrm{Per}(\mathcal{G}(P))$, to be the number of grid edges on the boundary of the cluster, excluding edges running along the outside of the grid. See Figure 3.6 for an example.

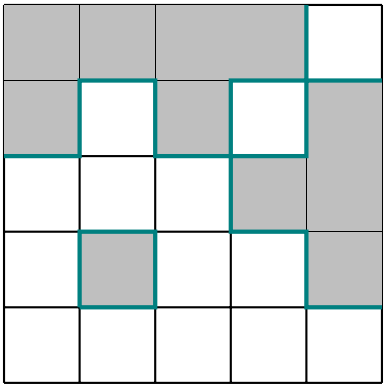


Figure 3.6: A cluster $\mathcal{G}(P)$ indicated in gray, with size $|\mathcal{G}(P)| = 11$ and perimeter $\mathrm{Per}(\mathcal{G}(P)) = 19$. Note that edges of the cluster running along the exterior of the grid do not contribute to the perimeter.

For a Pauli operator $P \in \mathcal{P}^n$, the following lemma bounds the probability that $U^\dagger PU$ is Z-type in terms of $|\mathcal{G}(P)|$ and $\mathrm{Per}(\mathcal{G}(P))$ with randomness over the choice of the n-qubit 2D coarse-grained random Clifford circuit U.

Lemma 3.7. *Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits with $k = (\tau \times \tau)$-qubit gates. Let $P \in \mathcal{P}^n$. Define $a = |\mathcal{G}(P)|$ and $l = \mathrm{Per}(\mathcal{G}(P))$. Then it holds that*

$$\Pr_{U \sim \mathcal{D}}(U^\dagger PU \text{ is } Z\text{-type}) \leq 2^{-a\tau^2 - \frac{1}{4}l(\tau^2/12 - 1)}. \tag{3.48}$$

Proof. Recall that we write $U = WV$ where V and W are the first and second layers of coarse-grained random Clifford gates respectively. We first consider the Pauli operator obtained by conjugating P with just the second layer of gates, which is $P' = W^\dagger PW$. Note that P' is a tensor product of uniformly random non-identity Pauli operators acting on every $\tau \times \tau$ patch of qubits in the support of $\mathcal{G}(P)$ and the identity everywhere else. There are a such patches by definition, so P' can only be non-identity on a subset $Q' \subseteq [n]$ of $a\tau^2$ qubits.

Now we consider the first layer of coarse-grained random Clifford gates $V = V_1 \otimes \cdots \otimes V_{(m-1)^2}$. For every $i \in [(m-1)^2]$, define $Q'_i = Q' \cap \mathrm{Supp}(V_i)$ and let P'_i be the truncation of

P' to the support of V_i. Finally, define $Q'_0 = Q' \setminus (\bigcup_{i=1}^{(m-1)^2} Q'_i)$ and P'_0 to be the truncation of P' to Q'_0. Also, as a notational convenience, let V_0 denote the identity gate acting on the qubits in Q'_0. This is to accommodate the qubits in Q' that lie on the boundary of the circuit and thus are not acted on by V. We want to upper bound the quantity

$$\Pr_{U\sim\mathcal{D}}(U^\dagger PU \text{ is } Z\text{-type}) \tag{3.49}$$

$$= \Pr_{U\sim\mathcal{D}}\left(V_i^\dagger P'_i V_i \text{ is } Z\text{-type for all } 0 \le i \le (m-1)^2\right) \tag{3.50}$$

$$= \prod_{i=0}^{(m-1)^2} \Pr_{U\sim\mathcal{D}}\left(V_i^\dagger P'_i V_i \text{ is } Z\text{-type} \mid V_j^\dagger P'_j V_j \text{ is } Z\text{-type for all } 0 \le j < i\right). \tag{3.51}$$

It is immediate from definitions that

$$\Pr_{U\sim\mathcal{D}}(V_0^\dagger P'_0 V_0 \text{ is } Z\text{-type}) = \Pr_{U\sim\mathcal{D}}(P'_0 \text{ is } Z\text{-type}) \le 2^{-|Q'_0|}. \tag{3.52}$$

Define $\mathcal{V}_{\text{int}}$ to be the set of all V_i with $0 < i \le (m-1)^2$ whose support is contained entirely in Q', and define $\mathcal{V}_{\text{perim}}$ to be the set of all V_i with $0 < i \le (m-1)^2$ whose support overlaps partially with Q'. By applying Lemma 3.6 (a), we see that for every $V_i \in \mathcal{V}_{\text{int}}$,

$$\Pr\left(V_i^\dagger P'_i V_i \text{ is } Z\text{-type} \mid V_j^\dagger P'_j V_j \text{ is } Z\text{-type for all } j < i\right) \le 2^{-\tau^2} = 2^{-|Q'_i|}. \tag{3.53}$$

Similarly, for every $V_i \in \mathcal{V}_{\text{perim}}$, Lemma 3.6 (b) gives that

$$\Pr\left(V_i^\dagger P'_i V_i \text{ is } Z\text{-type} \mid V_j^\dagger P'_j V_j \text{ is } Z\text{-type for all } j < i\right) \le 4^{-|Q'_i|} + 2^{-\tau^2}. \tag{3.54}$$

For every $V_i \in \mathcal{V}_{\text{perim}}$, we have that at least $1/4$ and at most $3/4$ of the qubits in the support of V_i overlap with Q', and hence $\tau^2/4 \le |Q'_i| \le 3\tau^2/4$. Then, for every $V_i \in \mathcal{V}_{\text{perim}}$, we can also write

$$\Pr\left(V_i^\dagger P'_i V_i \text{ is } Z\text{-type} \mid V_j^\dagger P'_j V_j \text{ is } Z\text{-type for all } j < i\right) \le 4^{-|Q'_i|} + 2^{-4|Q'_i|/3} \tag{3.55}$$

$$= 2^{-|Q'_i|}\left(2^{-|Q'_i|} + 2^{-|Q'_i|/3}\right) \tag{3.56}$$

$$\le 2^{-|Q'_i| - (\tau^2/12)+1}. \tag{3.57}$$

Combining (3.52), (3.53), and (3.57) with (3.51) gives

$$\Pr(U^\dagger PU \text{ is } Z\text{-type}) \le 2^{-|Q'_0|}\left(\prod_{V_i\in\mathcal{V}_{\text{int}}} 2^{-|Q'_i|}\right)\left(\prod_{V_j\in\mathcal{V}_{\text{perim}}} 2^{-|Q_j| - \tau^2/12+1}\right) \tag{3.58}$$

$$= 2^{-|\mathcal{V}_{\text{perim}}|(\tau^2/12-1) - \sum_{i=0}^{(m-1)^2}|Q'_i|} \tag{3.59}$$

$$= 2^{-|\mathcal{V}_{\text{perim}}|(\tau^2/12-1) - |Q'|}. \tag{3.60}$$

We observed near the beginning of the proof that $|Q'| = a\tau^2$. Every gate $V_i \in \mathcal{V}_{\text{perim}}$ intersects with four edges, and in the worst case, all four of the edges could be perimeter edges. At the same time, every perimeter edge intersects with at least one gate in $\mathcal{V}_{\text{perim}}$. Therefore, $|\mathcal{V}_{\text{perim}}| \ge \frac{1}{4}l$. Inserting those two values into the equation above completes the proof. □

Our strategy will be to partition D as

$$D = \bigcup_{l=0}^{\infty} \{s \in D : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}, \tag{3.61}$$

so next, we upper bound the expected size of each subset of D in the disjoint union.

Lemma 3.8. *Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits with $k = (\tau \times \tau)$-qubit gates. Let $[n] = A \cup B \cup C$ be a tripartition such that $|A| = 1$. Let $l \geq 0$ be an integer. Let D be the set defined in Lemma 3.4. Then with $m^2 = n/\tau^2$, it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}} \left[|\{s \in D : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}| \right] \leq \begin{cases} 3(2^{-|C|-1}) & \text{if } l = 0, \\ 6 \cdot 2^{l(2\log_2(m) - \tau^2/48 + 1/4)} & \text{if } l \geq 1. \end{cases} \tag{3.62}$$

Proof. For $U \sim \mathcal{D}$, by the definition of the set D, there is a one-to-one correspondence between every $s \in D$ and every pair (r, P) such that $r \in \{0,1\}^{|B|}$, $P \in \{X, Y, Z\}$, and $U^\dagger(P_A Z(r)_B)U$ is Z-type. Furthermore, for $U \sim \mathcal{D}$, there is a one-to-one correspondence between every $s \in D$ such that $\mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l$ and every pair (r, P) such that $r \in \{0,1\}^{|B|}$, $P \in \{X, Y, Z\}$, $U^\dagger(P_A Z(r)_B)U$ is Z-type, and $\mathrm{Per}(\mathcal{G}(P_A Z(r)_B)) = l$.

For a given cluster $\mathcal{G}'$, let $\Gamma(\mathcal{G}')$ denote the set of all pairs $r \in \{0,1\}^{|B|}$ and $P \in \{X, Y, Z\}$ such that $\mathcal{G}(P_A Z(r)_B) = \mathcal{G}'$. Then by Lemma 3.7, we get

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}} \left[|\{s \in D : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}| \right] \tag{3.63}$$

$$= \sum_{\mathcal{G}' : \mathrm{Per}(\mathcal{G}') = l} \sum_{(r,P) \in \Gamma(\mathcal{G}')} \Pr_{U\sim\mathcal{D}} \left(U^\dagger \left(P_A Z(r)_B \right) U \text{ is } Z\text{-type} \right) \tag{3.64}$$

$$\leq \sum_{\mathcal{G}' : \mathrm{Per}(\mathcal{G}') = l} |\Gamma(\mathcal{G}')| 2^{-|\mathcal{G}'|\tau^2 - \frac{1}{4} l(\tau^2/12 - 1)}. \tag{3.65}$$

First consider the case where $l = 0$. The only cluster $\mathcal{G}'$ with $\mathrm{Per}(\mathcal{G}') = 0$ is the entire $m \times m$ grid, which has $|\mathcal{G}'| = m^2$ and $|\Gamma(\mathcal{G}')| \leq 3(2^{|B|})$. Thus, we get the upper bound

$$(3.65) \leq 3(2^{|B|}) 2^{-m^2\tau^2} = 3(2^{n-1-|C|}) 2^{-n} = 3(2^{-|C|-1}). \tag{3.66}$$

For $l \geq 1$, we can bound the sum in (3.65) as

$$(3.65) \leq 3 \sum_{\mathcal{G}' : \mathrm{Per}(\mathcal{G}') = l} 2^{-\frac{1}{4} l(\tau^2/12 - 1)} \leq 6 \binom{2m(m-1)}{l} 2^{-\frac{1}{4} l(\tau^2/12 - 1)} \tag{3.67}$$

$$\leq 6 \cdot 2^{l(2\log_2(m) - \tau^2/48 + 1/4)}. \tag{3.68}$$

The first inequality holds since there are at most $3(2^{|\mathcal{G}'|\tau^2})$ pairs $r \in \{0,1\}^{|B|}$ and $P \in \{X, Y, Z\}$ such that $\mathcal{G}(P_A Z(r)_B) = \mathcal{G}'$. For the second inequality, we count the number of clusters $\mathcal{G}'$ with perimeter l. Any such cluster can be specified by first choosing the location of the l edges of the cluster and then specifying the interior and exterior of the cluster. A crude upper bound gives that there are at most $\binom{2m(m-1)}{l}$ ways to choose the l edges of the cluster, and then two possible options for how to specify the interior and exterior of the cluster once the edges are fixed. This argument gives $|\{\mathcal{G}' : \mathrm{Per}(\mathcal{G}') = l\}| \leq 2\binom{2m(m-1)}{l}$, which is the bound used above. □

Notice that we will require $k = \tau^2 \geq \Omega(\log n)$ in (3.68) because we will need τ^2 to be large enough to cancel the $\log_2(m)$ term coming from the binomial factor $\binom{2m(m-1)}{l}$. If we work with the lightcone connected set $\tilde{D}$ instead, then we can restrict the enumeration to clusters $\mathcal{G}'$ that are lightcone connected and act non-trivially on A. Clearly,

$$|\{\mathcal{G}' : \mathrm{Per}(\mathcal{G}') = l,\ \mathrm{Supp}(\mathcal{G}') \text{ is lightcone connected, and } \mathcal{G}' \text{ acts non-trivially on } A\}| \tag{3.69}$$

is only a function of l. If it is true that (3.69) $\leq 2^{cl}$ for any constant $c > 0$, then requiring τ to be a large enough constant would suffice to render (3.68) inverse exponential in l.

Finally, we are ready to upper bound the expected size of D.

Lemma 3.9. *Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits with $k = (\tau \times \tau)$-qubit gates. Let $[n] = A \cup B \cup C$ be a tripartition such that $|A| = 1$. Let D be the set defined in Lemma 3.4. Assume n is sufficiently large and $\tau \geq \sqrt{1000 \log_2(m)}$ where $m^2 = n/\tau^2$. Then it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|D|] \leq 3(2^{-|C|-1}) + 12(2^{-\tau^2/100}). \tag{3.70}$$

Proof. Since $\tau \geq \sqrt{1000 \log_2(m)}$, we get $2\log_2(m) - \tau^2/48 + 1/4 \leq -\tau^2/100$ for sufficiently large n. Thus, using Lemma 3.8, we get

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|D|] = \sum_{l=0}^{\infty} \mathop{\mathbb{E}}_{U\sim\mathcal{D}} \left[|\{s \in D : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}|\right] \tag{3.71}$$

$$\leq 3(2^{-|C|-1}) + \sum_{l=1}^{\infty} 6 \cdot 2^{l(2\log_2(m) - \tau^2/48 + 1/4)} \tag{3.72}$$

$$\leq 3(2^{-|C|-1}) + 6\sum_{l=1}^{\infty} 2^{l(-\tau^2/100)} \tag{3.73}$$

$$\leq 3(2^{-|C|-1}) + 12(2^{-\tau^2/100}). \tag{3.74}$$

□

Proof of Theorem 3.2. Under the assumptions of the theorem, Lemma 3.9 holds. Thus, we have $\Pr_{U\sim\mathcal{D}}(D \neq \emptyset) \leq \mathbb{E}_{U\sim\mathcal{D}}[|D|] \leq 1/\mathrm{poly}(n)$ since $|C| \geq \sqrt{n}$ in an annulus tripartition. Thus, $D = \emptyset$ with probability $1 - 1/\mathrm{poly}(n)$. Then the result follows from Lemma 3.4 and (3.1). □

By closely examining the proofs of Lemma 3.8 and Lemma 3.9, we realize that a slight modification of our arguments shows that 2D coarse-grained random Clifford circuits with $k = (\tau \times \tau)$-qubit gates for $\tau \geq \sqrt{1000 \log_2(n)}$ satisfy the anti-concentration property Definition 1.5.

Corollary 3.1. *Assume n is sufficiently large. Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $k = (\tau \times \tau)$-qubit gates with $\tau \geq$*

$\sqrt{1000\log_2(n)}$. *For $U \sim \mathcal{D}$, we use S_ψ^Z to denote the Z-type stabilizer subgroup of $|\psi\rangle = U|0^n\rangle$. Then it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|S_\psi^Z|] \leq 2 + 1/\mathrm{poly}(n), \tag{3.75}$$

which implies that $\mathcal{D}$ satisfies the anti-concentration property Definition 1.5.

Proof. We can see this by running the proofs of Lemma 3.8 and Lemma 3.9 on an "illegal tripartition" $[n] = A \cup B \cup C$ where $B = [n]$ and $A = C = \emptyset$.

For $U \sim \mathcal{D}$, define $D' = \{s \in \{0,1\}^n : UZ(s)U^\dagger \text{ is } Z\text{-type}\}$. We have

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|S_\psi^Z|] = \mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|D'|] = \sum_{l=0}^{\infty} \mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|\{s \in D' : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}|]. \tag{3.76}$$

For a given cluster $\mathcal{G}'$, let $\Gamma(\mathcal{G}')$ denote the set of all $r \in \{0,1\}^n$ such that $\mathcal{G}(Z(r)) = \mathcal{G}'$. Let $l \geq 0$. By Lemma 3.7, we get

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}\left[|\{s \in D' : \mathrm{Per}(\mathcal{G}(UZ(s)U^\dagger)) = l\}|\right] \tag{3.77}$$

$$= \sum_{\mathcal{G}':\mathrm{Per}(\mathcal{G}')=l} \sum_{r\in\Gamma(\mathcal{G}')} \Pr_{U\sim\mathcal{D}}(U^\dagger Z(r)U \text{ is } Z\text{-type}) \tag{3.78}$$

$$\leq \sum_{\mathcal{G}':\mathrm{Per}(\mathcal{G}')=l} |\Gamma(\mathcal{G}')| 2^{-|\mathcal{G}'|\tau^2 - \frac{1}{4}l(\tau^2/12-1)}. \tag{3.79}$$

For $l = 0$, $\mathcal{G}'$ is either empty or the entire $m \times m$ grid where $m^2 = n/\tau^2$. When $\mathcal{G}'$ is empty, $|\mathcal{G}'| = 0$ and $\Gamma(\mathcal{G}') = \{0^n\}$. When $\mathcal{G}'$ is the entire $m \times m$ grid, $|\mathcal{G}'| = m^2$ and $|\Gamma(\mathcal{G}')| \leq 2^n$. Thus, for $l = 0$, we have

$$(3.79) \leq 1 + 2^n 2^{-m^2\tau^2} = 2. \tag{3.80}$$

For $l \geq 1$, we can bound the sum in (3.79) as

$$(3.79) \leq \sum_{\mathcal{G}':\mathrm{Per}(\mathcal{G}')=l} 2^{-\frac{1}{4}l(\tau^2/12-1)} \leq 2\binom{2m(m-1)}{l} 2^{-\frac{1}{4}l(\tau^2/12-1)} \tag{3.81}$$

$$\leq 2 \cdot 2^{l(2\log_2(m) - \tau^2/48 + 1/4)}. \tag{3.82}$$

The first inequality holds since there are at most $2^{|\mathcal{G}'|\tau^2}$ strings $r \in \{0,1\}^n$ such that $\mathcal{G}(Z(r)) = \mathcal{G}'$. The second and third inequalities follow from the same bounds used in the proof of Lemma 3.8.

Now, we see that substituting (3.80) and (3.82) into (3.76) and summing the geometric series assuming $\tau \geq \sqrt{1000\log_2(n)}$ for sufficiently large n gives the claimed upper bound

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}}[|S_\psi^Z|] \leq 2 + 1/\mathrm{poly}(n). \tag{3.83}$$

□

3.4 Long-range tripartite MIE in 2D coarse-grained random Clifford circuits

In this section, we will take advantage of the fact that Lemma 3.9 from the last section applies to any tripartition $[n] = A \cup B \cup C$ with $|A| = 1$ to show that a certain notion of long-range tripartite MIE arises in the output states of 2D coarse-grained random Clifford circuits with gate size $k = \tau^2 \geq 1000 \log_2(n)$.

To set the stage, consider the 3-qubit GHZ state

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{3.84}$$

A 3-qubit stabilizer state is tripartite entangled if all three of its single-qubit reduced states are maximally mixed. The GHZ state is clearly tripartite entangled, but it is not the only 3-qubit stabilizer state that is tripartite entangled. We say two n-qubit stabilizer states $|\psi\rangle$ and $|\phi\rangle$ are locally Clifford equivalent if there exist single-qubit Clifford unitaries $C_1, \ldots, C_n$ such that

$$(C_1 \otimes \cdots \otimes C_n)|\psi\rangle = |\phi\rangle. \tag{3.85}$$

Locally Clifford equivalent stabilizer states have the same entanglement structure, and they define an equivalence relation on the set of all n-qubit stabilizer states. Clearly, all the 3-qubit stabilizer states that are locally Clifford equivalent to the GHZ state are also tripartite entangled, and we call them GHZ-type entangled. It turns out that every tripartite entangled 3-qubit stabilizer state is GHZ-type entangled [85]. One way to see this is through the graph state formalism, which we briefly introduce next.

For a graph $G = (V, E)$, the $|V|$-qubit graph state $|G\rangle$ is defined to be the stabilizer state whose stabilizer group is generated by

$$g_u = X_u \Bigg(\prod_{v:\{u,v\}\in E} Z_v \Bigg) \tag{3.86}$$

for every $u \in V$. It is known that every n-qubit stabilizer state $|\psi\rangle$ is locally Clifford equivalent to a graph state $|G(\psi)\rangle$ for some graph $G(\psi)$ with n vertices [86]. From this, we see that for any bipartition $[n] = A \cup B$, there is entanglement between A and B in $|\psi\rangle$ if and only if $|G(\psi)\rangle$ is entangled across the (A, B)-cut. It is not hard to see from the definition of $|G(\psi)\rangle$ that the later condition holds if and only if there is an edge crossing the cut (A, B) in $G(\psi)$.

The only connected graphs on three vertices are the path graph with two edges P_3 and the complete graph on three vertices K_3. For P_3, suppose the vertex with degree two corresponds to the second qubit; then we can verify that $(H \otimes I \otimes H)|P_3\rangle = |\mathrm{GHZ}\rangle$, where H is the Hadamard gate. For K_3, we can analogously verify that $(HS^\dagger \otimes HSH \otimes HS^\dagger)|K_3\rangle = |\mathrm{GHZ}\rangle$. Therefore, all tripartite entangled 3-qubit stabilizer states are GHZ-type entangled.

Now we are ready to define the long-range tripartite MIE property for random Clifford circuits.

Property 3.2 (Long-range tripartite MIE)**.** A distribution $\mathcal{D}$ of n-qubit random Clifford circuits satisfies the long-range tripartite MIE property if there exists a constant $c > 0$ such

that with probability at least $c > 0$ over the choice of $U \sim \mathcal{D}$ and a uniformly random triplet of qubits $\{h, i, j\} \subseteq [n]$, the 3-qubit pure post-measurement state $|\psi'\rangle$ obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in $[n] \setminus \{h, i, j\}$ in the computational basis is GHZ-type entangled.

Note that the architecture of $\mathcal{D}$ is not required to be geometrically local in 2D. Similar to Property 3.1, we also conjecture that Property 3.2 holds for 2D coarse-grained random Clifford circuits when τ is a sufficiently large constant and for 2D brickwork random Clifford circuits with circuit depth $d \geq 7$. The goal of this section is to establish the following weaker theorem.

Theorem 3.3. *For sufficiently large n, the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $k = (\tau \times \tau)$-qubit gates with $\tau \geq \sqrt{1000 \log_2(n)}$ satisfies Property 3.2 with constant $c > 0.005$.*

In addition to applying Lemma 3.9 multiple times, we will also need the following lemma, which shows that GHZ-type entanglement can always be produced by making computational basis measurements on a connected graph state with sufficiently many vertices.

Lemma 3.10. *Let $G = (V, E)$ be a connected graph with at least 3 vertices, and let $|G\rangle$ be the associated graph state. Then there exist vertices $\Omega = \{h, i, j\} \subseteq V$ such that for every $x \in \{0, 1\}^{|V|-3}$, the state $\langle x|_{V \setminus \Omega}|G\rangle$ is locally Clifford equivalent to the GHZ state after normalization.*

Proof. Let $u \in V$ and let $G - u$ denote the graph obtained from G by deleting the vertex u and all the edges incident to it. From the definition of $|G\rangle$, we see that the post-measurement state after measuring the qubit u of $|G\rangle$ in the computational basis and obtaining the outcome $y \in \{0, 1\}$ is

$$\left(\prod_{v:\{u,v\} \in E} Z_v^y \right) |G - u\rangle. \tag{3.87}$$

We choose $\Omega = \{h, i, j\} \subseteq V$ such that the subgraph G_Ω of G induced by Ω is a connected graph with three vertices. By applying successively the above fact about computational basis measurement, we see that after measuring all qubits of $|G\rangle$ in $V \setminus \Omega$, the post-measurement state is locally Clifford equivalent to the graph state $|G_\Omega\rangle$. Since G_Ω is either P_3 or K_3, we have already seen that $|G_\Omega\rangle$ is locally Clifford equivalent to the GHZ state. □

Proof of Theorem 3.3. Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $k = (\tau \times \tau)$-qubit gates with $\tau \geq \sqrt{1000 \log_2(n)}$. We say $\Omega = \{h, i, j\} \subseteq [n]$ is a good triplet if with probability at least 0.05 over the choice of $U \sim \mathcal{D}$, the 3-qubit pure post-measurement state $|\psi'\rangle$ obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in $[n] \setminus \Omega$ in the computational basis is GHZ-type entangled. Below, we prove that for every subset of qubits $V \subseteq [n]$ with $|V| = 5$, there exists a good triplet $\Omega \subseteq V$. The theorem then follows from this statement. To see this, consider the following procedure. First choose a uniformly random subset of qubits $V \subseteq [n]$ of size $|V| = 5$, and then choose a uniformly random triplet $\Omega \subseteq V$. By the above, the probability that we are lucky and choose a good triplet is at least $1/\binom{5}{3} = 0.1$. Moreover, clearly this procedure samples a triplet Ω from the uniform distribution over all $\binom{n}{3}$ triplets.

Let $V \subseteq [n]$ with $|V| = 5$. It remains to show that there exists a good triplet $\Omega = \{h, i, j\} \subseteq V$. For $U \sim \mathcal{D}$, let $|\psi'\rangle$ be the 5-qubit post-measurement state obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in $[n] \setminus V$ in the computational basis. For every qubit $v \in V$, by taking the tripartition $[n] = A \cup B \cup C$ where $A = \{v\}$ and $C = V \setminus \{v\}$, Lemma 3.4 and Lemma 3.9 imply that there is no entanglement between A and C in $|\psi'\rangle$ with probability at most

$$3(2^{-5}) + 12(2^{-\tau^2/100}) \leq 0.09375 + 1/\mathrm{poly}(n) < 0.1. \tag{3.88}$$

Then by the union bound, there exists a qubit $v \in V$ that is unentangled with $V \setminus \{v\}$ in $|\psi'\rangle$ with probability at most $|V| \cdot 0.1 < 0.5$. Thus, the event that there is entanglement between every qubit $v \in V$ and $V \setminus \{v\}$ in $|\psi'\rangle$ occurs with probability at least 0.5, and suppose this event occurs. This implies that $|\psi'\rangle$ is locally Clifford equivalent to a graph state $|G\rangle$ for some graph G with no isolated vertex. It follows that the graph G must have either a 3- or 5-vertex connected component. If G has a 3-vertex component, then the qubits of $|\psi'\rangle$ corresponding to this component are in a state locally Clifford equivalent to the GHZ state, and we are done. Otherwise, G has a 5-vertex connected component. In this case, by Lemma 3.10, there are vertices $\Omega = \{h, i, j\} \subseteq V$ such that for any $x \in \{0,1\}^2$, we have that

$$\langle x|_{V\setminus\Omega}|G\rangle \tag{3.89}$$

is locally Clifford equivalent to the GHZ state. Since $|\psi'\rangle$ is locally Clifford equivalent to $|G\rangle$, we have

$$|\psi'\rangle = \Big(\bigotimes_{v\in V} C_v\Big)|G\rangle \tag{3.90}$$

for some single-qubit Clifford unitaries C_v, $v \in V$. Furthermore, since the distribution $\mathcal{D}$ is invariant under local Clifford operations, we can assume the single-qubit Clifford unitaries C_v, $v \in V$ are chosen uniformly at random. Then the post-measurement state of the qubits in Ω when the two qubits of $|\psi'\rangle$ in $V \setminus \Omega$ are measured in the computational basis and outcome $x \in \{0,1\}^2$ is observed is given by

$$\langle x|_{V\setminus\Omega}\Big(\bigotimes_{v\in V} C_v\Big)|G\rangle. \tag{3.91}$$

Comparing with (3.89), we see that (3.91) is locally Clifford equivalent to the GHZ state provided the state

$$\Big(\bigotimes_{v\in V\setminus\Omega} C_v^\dagger\Big)|x\rangle_{V\setminus\Omega} \tag{3.92}$$

is a computational basis state. Since C_v, $v \in V$ are uniformly random single-qubit Clifford unitaries, this occurs with probability exactly $1/3^2 = 1/9$ since exactly 8 out of the 24 single-qubit Clifford unitaries map a computational basis state to a computational basis state. Putting everything together, we see that Ω is a good triplet since the state in (3.91) is locally Clifford to the GHZ state with probability at least $(1/9)(1/2) > 0.05$. □

3.5 Long-range MIE in 2D coarse-grained random Clifford circuits in the row-by-row tripartition

In this section, we try to generalize the approach of Section 3.2 from 1D to 2D.

First, we define the 2D coarse-grained architecture differently in this section. Compared to the definition adopted in Section 3.3, we swap the two layers of gates. Under the new convention, in the example shown in Figure 3.4, the first layer contains 16 gates, while the second layer contains 9 gates. This choice naturally generalizes the 1D coarse-grained circuit diagram Figure 3.3 to 2D.

Second, we define a 2D generalization of the tripartition considered in Section 3.2. For an n-qubit 2D coarse-grained architecture composed of $k = (\tau \times \tau)$-qubit gates, we consider the tripartition $[n] = A \cup B \cup C$, where A contains the top $\tau/2$ rows of qubits, C contains the bottom $\tau/2$ rows of qubits, and B contains all the other qubits. We call this tripartition the row-by-row tripartition, and it is a direct generalization of the tripartition considered in Section 3.2 for 1D. For this tripartition, we do not require the set of qubits $[n]$ to reside on a square grid, and we will generally consider a rectangular grid of $n = L_y \times L_x$ qubits for some $L_y > L_x$. Under this convention, we have $|A| = |C| = (\tau/2)(L_x)$.

The recent breakthrough work of [64] has rigorously proven the existence of long-range MIE between A and C in the row-by-row tripartition for 2D coarse-grained random quantum circuits composed of Haar random gates. Their proof applies to a slight variant of the 2D coarse-grained architecture where each coarse-grained gate acts on 4 qudits. They let the local qudit dimension q play the role of the gate size parameter τ in the qubit setting. Note that when each gate acts on 4 qudits, the sets A and C each contain a single row of qudits. More concretely, they have proved the following theorem.

Theorem 3.4 ([64])**.** *Let $[n]$ be n qudits with local qudit dimension q arranged on an $L_y \times L_x$ rectangular grid with $L_y = \mathrm{poly}(L_x)$. Let $\mathcal{D}$ denote the distribution of n-qudit 2D coarse-grained random quantum circuits composed of Haar random 4-qudit gates with local qudit dimension q. Let $[n] = A \cup B \cup C$ denote the row-by-row tripartition where A contains the top row of qudits, C contains the bottom row of qudits, and B contains all the other qudits. Then there exists some constant $\kappa > 0$ such that for every $q \geq 88$, it holds that*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D},x_B}[S(\rho)] \geq \kappa L_x \tag{3.93}$$

where for $U \sim \mathcal{D}$, $x_B \in \{0,\ldots,q-1\}^{|B|}$ is the outcome of measuring all qudits of $|\psi\rangle = U|0^n\rangle$ in B in the computational basis and obtaining pure post-measurement state $|\psi'\rangle$ on $A \cup C$, and $S(\rho)$ is the von Neumann entropy of $\rho = \mathrm{Tr}_C(|\psi'\rangle\langle\psi'|)$.

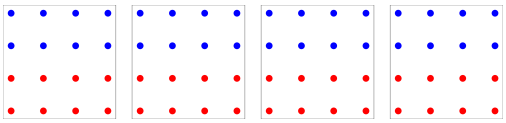
This theorem is the first to rigorously prove the existence of long-range MIE in a distribution of truly constant-depth 2D random quantum circuits with constant nominal circuit depth, gate size, and local qudit dimension. Remarkably, the theorem proves that MIE in the row-by-row tripartition is not only long-range as $L_y = \mathrm{poly}(L_x)$, but also extensive as the expected entanglement entropy is linear in $|A| = |C| = L_x$.

The goal of this section is to investigate the behaviour of MIE in 2D coarse-grained random Clifford circuits in the row-by-row tripartition, and we are able to provide numerical evidence for the following conjecture.

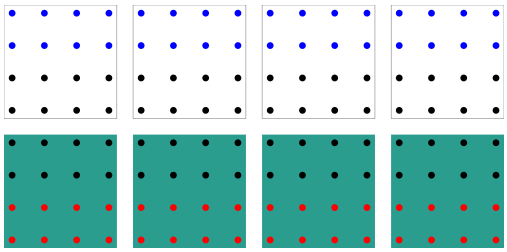
Conjecture 3.1. *Let $[n]$ be n qubits arranged on an $L_y \times L_x$ rectangular grid with $L_y = \mathrm{poly}(L_x)$. Let $\mathcal{D}$ denote the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $(\tau \times \tau)$-qubit gates for some even τ. Let $[n] = A \cup B \cup C$ denote the row-by-row tripartition where A contains the top $\tau/2$ rows of qubits, C contains the bottom $\tau/2$ rows of qubits, and B contains all the other qubits. Then there exists some constant $\kappa > 0$ such that for every $\tau \geq 4$, it holds that*

$$\mathop{\mathbb{E}}_{U \sim \mathcal{D}}[E] \geq \kappa L_x \tag{3.94}$$

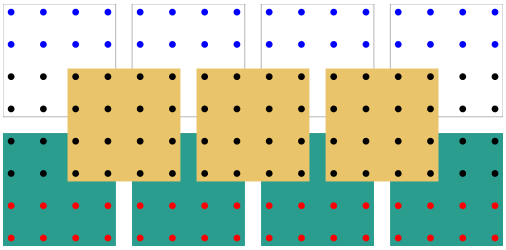
where for $U \sim \mathcal{D}$, E is the number of ebits of entanglement between A and C in the post-measurement state $|\psi'\rangle$ obtained by measuring all qubits of $|\psi\rangle = U|0^n\rangle$ in B in the computational basis.



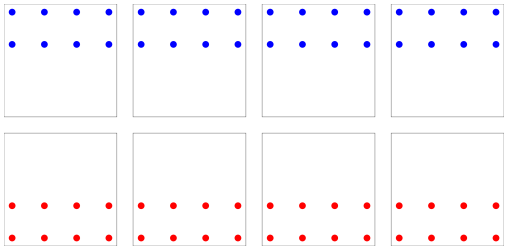
(a) Qubits from previous iteration.



(b) Adjoin new qubits and apply first-layer gates.



(c) Apply second-layer gates.

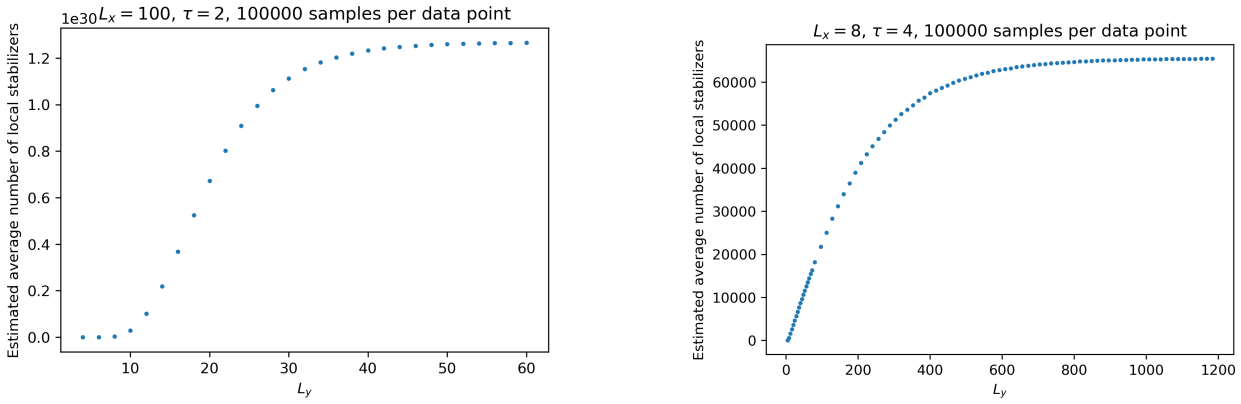


(d) Measure all qubits in B.

Figure 3.7: In each diagram, the blue (top) and red (bottom) qubits form the sets A and C respectively. The black qubits belong to the set B. The green and brown squares show the support of the gates in the first and second layer respectively.

Completely analogous to what we did in Section 3.2, we can build a 2D coarse-grained random Clifford circuit with measurements on B iteratively along the vertical direction. The four panels in Figure 3.7 from (a) to (d) show how to advance by one iteration for $L_x = 16$ and $\tau = 4$, and the process is the same for other values of L_x and τ. Just like in the 1D case, we can use a function $f(j)$, $j \geq 0$ to record the expected size of the stabilizer subgroup of $|\psi'\rangle$ local on A after the j-th iteration. As a sanity check, we see that $f(j)$ should be monotonically increasing in j, and for fixed constants L_x and τ, $f(j)$ should approach $2^{|A|} = 2^{L_x\tau/2}$ as $j \to \infty$ since the entanglement should decay to zero as $L_y \to \infty$.

Our goal is to find the maximum increment $\mathrm{MaxInc}(L_x, \tau) = \max_{j \geq 0}(f(j+1) - f(j))$, which is a function of L_x and τ. With this, we obtain the upper bound $f(j) \leq f(0) + j \cdot \mathrm{MaxInc}(L_x, \tau)$ for every $j \geq 1$, and we can conclude that extensive entanglement survives after j iterations for some $j \geq 1$ if $f(j) \leq f(0) + j \cdot \mathrm{MaxInc}(L_x, \tau) << 2^{L_x\tau/2}$.



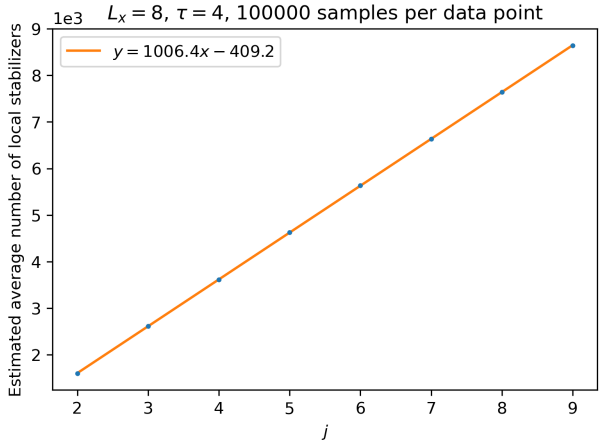
(a) (b)

Figure 3.8: Numerical estimation of the expected number of stabilizers of $|\psi'\rangle$ local on A as a function of L_y for (a) $L_x = 100$ and $\tau = 2$ and (b) $L_x = 8$ and $\tau = 4$.

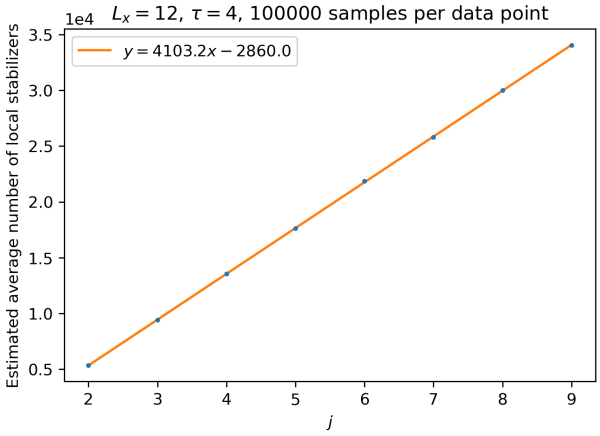
Figure 3.8 shows the results of numerically estimating the value of $f(j)$ plotted as a function of $L_y = (j+1)\tau$ for $(L_x, \tau) = (100, 2)$ and $(L_x, \tau) = (8, 4)$. While it is more visually evident for $(L_x, \tau) = (100, 2)$, we observe logistic growth in both cases. For $(L_x, \tau) = (100, 2)$, $f(j)$ plateaus towards the limiting value $2^{100} \approx 1.268 \times 10^{30}$ very quickly at values of $L_y < L_x = 100$, so the entanglement appears to be short lived. In sharp contrast, for $(L_x, \tau) = (8, 4)$, the numerics predict that it requires $L_y \gtrsim 2^{L_x}$ for $f(j)$ to reach half of the maximum value $2^{16} = 65536$, so the entanglement survives for significantly longer.

As the maximum slope of a logistic curve occurs during the linear segment, we can extract, for example, an estimate for $\mathrm{MaxInc}(8, 4)$ from the slope of a best linear fit to the linear region of Figure 3.8(b). By repeating this for multiple values of L_x with $\tau = 4$ fixed, we can hope to extract an expression for $\mathrm{MaxInc}(L_x, 4)$. The plots (a) through (f) in Figure 3.9 show the results of performing the fitting for $L_x \in \{8, 12, 16, 20, 24, 28\}$. We obtain the estimates $\mathrm{MaxInc}(8, 4) \approx 1006.4$, $\mathrm{MaxInc}(12, 4) \approx 4103.2$, $\mathrm{MaxInc}(16, 4) \approx 16474.7$, $\mathrm{MaxInc}(20, 4) \approx 65971.8$, $\mathrm{MaxInc}(24, 4) \approx 265776.3$, and $\mathrm{MaxInc}(28, 4) \approx 1056667.6$. A clear pattern emerges that for each L_x, $\mathrm{MaxInc}(L_x + 4, 4) \approx 4 \cdot \mathrm{MaxInc}(L_x, 4)$. This suggests that $\mathrm{MaxInc}(L_x, 4)$ scales exponentially in L_x. We can fit an exponential model $a2^{bL_x}$ with parameters a and b to $\mathrm{MaxInc}(L_x, 4)$ by fitting a linear model $bL_x + \log_2(a)$ to $\log_2(\mathrm{MaxInc}(L_x, 4))$. Figure 3.10(b) shows the result of performing the fitting on a log-linear plot, and the model predicted by the data is $\mathrm{MaxInc}(L_x, 4) \propto 2^{L_x/2}$. By repeating the whole procedure for $\tau \in \{2, 6, 8\}$, the observed data shown in Figure 3.10 collectively predict that $\mathrm{MaxInc}(L_x, \tau) \propto 2^{2L_x/\tau}$.

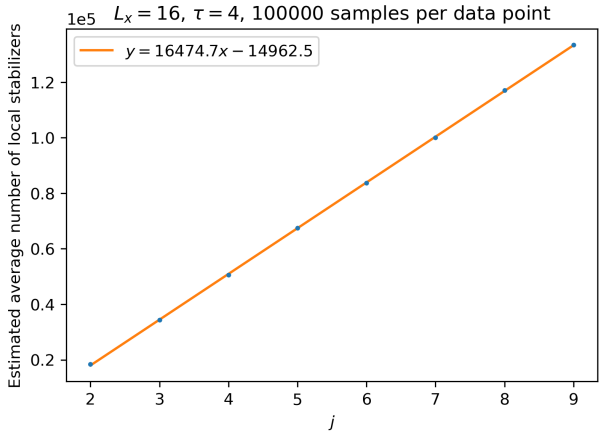
Overall, our numerical experiments suggest that for $\tau = 2$, $\mathrm{MaxInc}(L_x, 4)$ scales like $2^{2L_x/\tau} = 2^{L_x}$ while no entanglement survives when $f(j)$ reaches $2^{L_x\tau/2} = 2^{L_x}$. Thus, MIE is short-range for $\tau = 2$. In contrast, for $\tau = 4$, $\mathrm{MaxInc}(L_x, 4) \propto 2^{2L_x/\tau} = 2^{L_x/2}$; therefore, it takes a number of iterations that is exponential in L_x for $f(j)$ to reach $2^{L_x\tau/2} = 2^{2L_x}$, so MIE is extensive and long-range. The same conclusion can be drawn for all $\tau \geq 4$, and these constitute our numerical evidence for Conjecture 3.1.



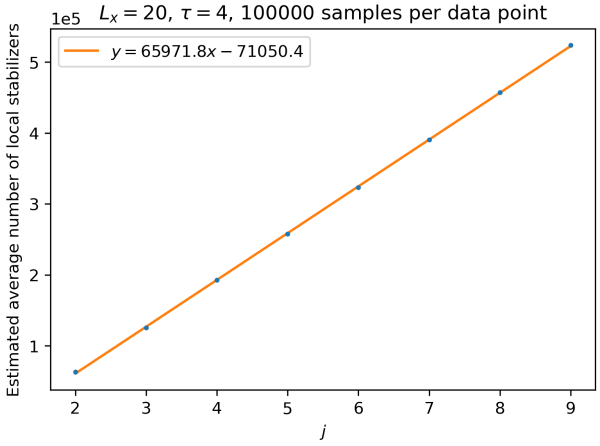
(a)



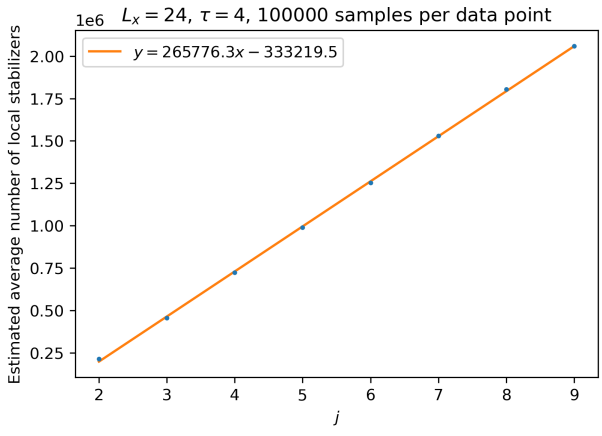
(b)



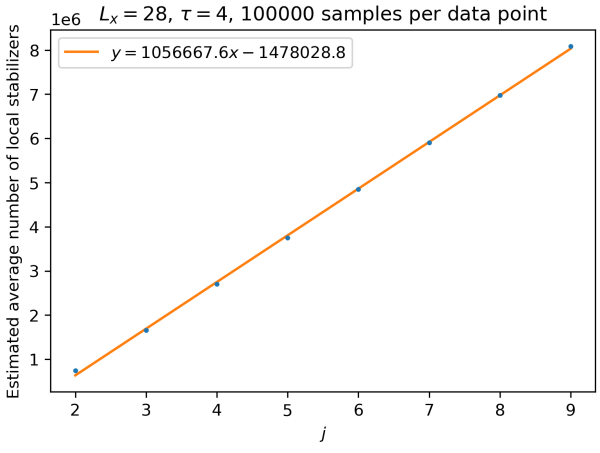
(c)



(d)



(e)



(f)

Figure 3.9: Numerical estimation of $\mathrm{MaxInc}(L_x, 4)$ for $L_x \in \{8, 12, 16, 20, 24, 28\}$ by estimating the slope of the linear segment of $f(j)$ as a function of j.

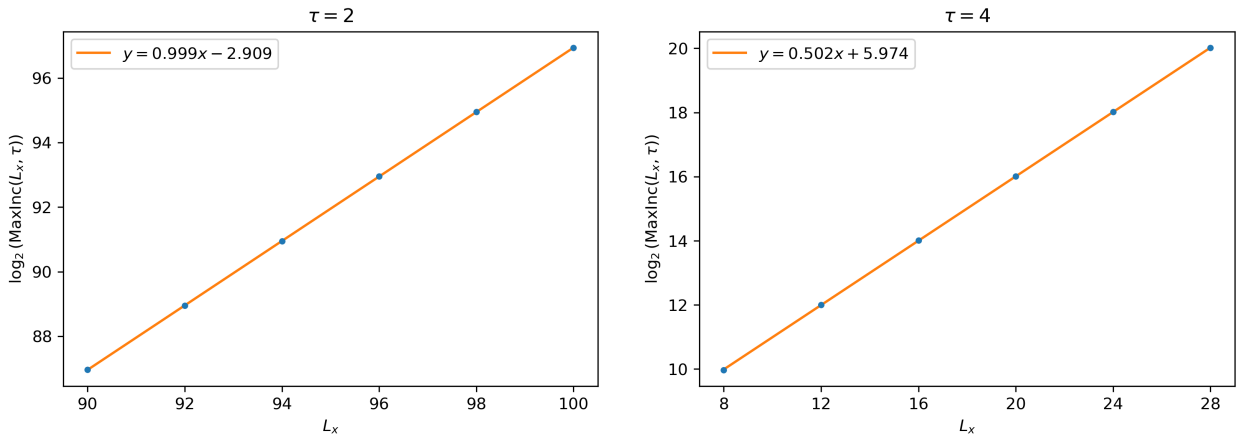


Figure 3.10: Fitting an exponential model $a2^{bL_x}$ with parameters a and b to $\text{MaxInc}(L_x, \tau)$ as a function of L_x for $\tau \in \{2, 4, 6, 8\}$ by fitting a linear model $bL_x + \log_2(a)$ to $\log_2(\text{MaxInc}(L_x, \tau))$. The data predicts that $b = 2/\tau$.

3.6 Quantum advantage from tripartite MIE

The main motivation for us to investigate the emergence of long-range MIE in the output states of shallow 2D random quantum circuits is the identification by the work [5] that it may be the great watershed separating quantum advantage from efficient classical simulation for the task of sampling from the output distributions of shallow 2D random quantum circuits. Establishing this quantum advantage for shallow random quantum circuits composed of Haar random gates has remained elusive, as it would imply a formal separation between efficient quantum and classical computation more broadly.

A series of recent works pioneered by [19] have provided unconditional results addressing a scaled-down notion of quantum advantage. In the work [19], a computational problem is introduced that can be solved by constant-depth 2D Clifford circuits over a fixed architecture, but not by any classical probabilistic circuit with bounded fan-in and whose depth grows sub-logarithmically with the input size. The work also provides distributions over instances such that shallow classical circuits cannot solve an instance in the average case. Existing results along these lines make use of specially tailored families of shallow quantum circuits, and it is a family of 2D Clifford circuits in the case of [19]. This raises the question of whether shallow 2D random Clifford circuits can be used to demonstrate a similar quantum advantage.

By closely examining the proof given in [19], we realize that the hard n-qubit circuit instances U constructed have the key property that there exist three specially planted qubits $\Omega = \{h, i, j\}$ such that when all qubits of the output state $|\psi\rangle = U|0^n\rangle$ not in Ω are measured in the computational basis, the pure post-measurement state $|\psi'\rangle$ on Ω is guaranteed to be GHZ-type entangled. This is almost exactly our long-range tripartite MIE property (Property 3.2), except with 100% success probability and that the triplet Ω is not selected randomly. While the proof given in [19] makes use of further structures in their construction, we will show that Property 3.2 is all that is needed for a quantum advantage.

Every computational quantum advantage with classical input and classical output can be phrased as the inability to classically simulate, for some suitable definition of classical simulation. When specifying a classical algorithm for simulating a quantum circuit, we need to specify what the classical description of the circuit is that is fed as the input to the algorithm. For algorithms that run in at least polynomial time, like the ones described in Chapter 2, the detailed format of the classical description does not matter much since polynomial time is enough for the algorithm to parse and reorganize any sensible classical description. However, for the quantum advantage that we will prove in this section, the allowed classical simulators will be limited to shallow classical circuits with bounded fan-in, which have restricted data processing flexibility. Therefore, we need to be explicit about the input format and ensure that it entails a fair comparison between quantum and classical.

We will consider classical probabilistic circuits defined as follows. First, we need to fix a random circuit distribution. Let $\mathcal{D}$ be any distribution of n-qubit depth-d random Clifford circuits composed of k-qubit gates. By fixing $\mathcal{D}$, we have fixed a specific architecture, which may or may not be geometrically local. Let R be the total number of k-qubit Clifford gates in the architecture associated with $\mathcal{D}$, and we label the gates $U_1, \ldots, U_R$. Let c_k denote the exact number of bits needed to specify a k-qubit Clifford gate. We know that $c_k = \Theta(k^2)$ since the total number of k-qubit Clifford gates is $2^{\Theta(k^2)}$ [87]. A classical probabilistic circuit

F_n with fan-in K and depth D for simulating a Clifford circuit $U \sim \mathcal{D}$ is a classical circuit of the following form. Every gate in F_n has K input bits and is allowed to have an unlimited number of output bits. The Boolean function computed by each gate in F_n can be arbitrary, and F_n is allowed to have an arbitrary number of gates per layer. If for some $i, j \in [D]$, an output bit of a gate in the i-th layer is fed as input to a gate in the j-th layer, it is required that $j > i$. Such a circuit computes a function with the following input-output behaviour, and we will use F_n to denote both the function and the circuit. F_n takes as input $c_k R$ bits, which are used to specify the gates $U_1, \ldots, U_R$. F_n also takes as input a random bit string $r \in \{0,1\}^{\ell(n)}$ drawn from some probability distribution q_n. Here $\ell(n)$ can be any function of n, and q_n can be any probability distribution over $\ell(n)$-bit strings. The function outputs a binary string

$$x = F_n(U_1, \ldots, U_R, r) \in \{0,1\}^n. \tag{3.95}$$

For ease of notation, we will sometimes write $x = F_n(U, r)$. Note that U is used to refer to both a quantum circuit $U \sim \mathcal{D}$ and its classical description $U = U_1 \cdots U_R \in \{0,1\}^{c_k R}$. This is well-defined as $U = U_1 \cdots U_R \in \{0,1\}^{c_k R}$ unambiguously specifies a circuit $U \sim \mathcal{D}$ for F_n since the information about the circuit architecture associated with $\mathcal{D}$ is assumed to be hardwired into F_n.

If we draw $r \sim q_n$ and compute $x = F_n(U, r)$, then $x \in \{0,1\}^n$ samples from a distribution that we denote by $P_U(x)$, $x \in \{0,1\}^n$. Let us also write $Q_U(x) = |\langle x|U|0^n\rangle|^2$, $x \in \{0,1\}^n$ for the true output distribution of a quantum circuit $U \sim \mathcal{D}$. For ease of notation, we will call a classical probabilistic circuit F_n of the kind we defined above with fan-in K and depth D an (n, K, D)-classical circuit.

Given the circuit structure of an (n, K, D)-classical circuit F_n, notice that each output bit z_j of $F_n(U, r)$ can depend on at most K^D input bits. In particular, this implies that each output bit z_j can depend on at most K^D of the gates in the input Clifford circuit. Recall from Lemma 1.1 that the size of the backward lightcone of a qubit in a quantum circuit $U \sim \mathcal{D}$ is also constrained to be at most k^d, and the reduced state of each qubit of $U|0^n\rangle$ depends on at most $\sum_{j=0}^{d-1} k^j = (k^d - 1)/(k-1) < k^d$ gates of U. Since each k-qubit Clifford gate is specified by $c_k = \Theta(k^2)$ bits, we need to allow at least $K \geq \text{poly}(k) \geq k^2$ for a fair comparison.

The notion of classical simulation we consider can be stated as follows.

Definition 3.2 (Classical simulation by (n, K, D)-classical circuits)**.** We say that a distribution $\mathcal{D}$ of n-qubit depth-d random Clifford circuits composed of k-qubit gates can be classically simulated by (n, K, D)-classical circuits if there exists an (n, K, D)-classical circuit F_n such that for some $\delta = 1/\text{poly}(n)$ and $\varepsilon = 1/\text{poly}(n)$, it holds that

$$\Pr_{U \sim \mathcal{D}}(\|P_U - Q_U\|_1 \leq \varepsilon) \geq 1 - \delta. \tag{3.96}$$

Therefore, to show that a distribution $\mathcal{D}$ exhibits a quantum advantage, we need to prove a statement implying the logical negation of the statement in Definition 3.2. The larger the values of K and/or D for which we can prove the logical negation, the wider the gap between quantum and classical becomes.

The main result of this section is the following theorem.

Theorem 3.5. *Let $\mathcal{D}$ be any distribution of n-qubit depth-d random Clifford circuits composed of k-qubit gates that satisfies Property 3.2. Assume n is sufficiently large, and in particular, $n \geq d^4$. Assume $k \leq n^{1/8}$. Then there exists a positive constant $\alpha > 0$ such that if an (n, K, D)-classical circuit with output distribution P_U on input circuit $U \sim \mathcal{D}$ satisfies*

$$\mathop{\mathbb{E}}_{U\sim\mathcal{D}} [\|P_U - Q_U\|_1] < \alpha \tag{3.97}$$

where $Q_U(x) = |\langle x|U|0^n\rangle|^2$ for every $x \in \{0,1\}^n$, then it holds that $D > \frac{\ln(n)}{4\ln(K)}$.

Theorem 3.5 implies that a distribution of random Clifford circuits with $k = O(1)$-qubit gates and depth $d = O(1)$ that can generate long-range tripartite MIE cannot be simulated by (n, K, D)-classical circuits with $K = O(1)$ and $D = O(1)$. Since it is conjectured that 2D brickwork random Clifford circuits with depth $d \geq 7$ satisfy Property 3.2, Theorem 3.5 provides evidence that such shallow random Clifford circuits achieve a quantum advantage over shallow classical circuits with bounded fan-in. This scaled-down quantum advantage can also be taken as evidence supporting the conjectured hardness of classically simulating 2D brickwork random quantum circuits composed of Haar random gates with depth $d > d^*$.

By directly composing Theorem 3.3 with Theorem 3.5, we obtain an end-to-end, unconditional quantum advantage using a distribution of n-qubit 2D coarse-grained random Clifford circuits with gate size $k = O(\log n)$ and depth $d = 2$ against all (n, K, D)-classical circuits with fan-in $K = \mathrm{polylog}(n)$ and depth $D \leq \ln(n)/4\ln(K)$.

Theorem 3.6. *Assume n is sufficiently large. It holds that the distribution of n-qubit 2D coarse-grained random Clifford circuits composed of $k = (\tau \times \tau)$-qubit gates with $\tau \geq \sqrt{1000\log_2(n)}$ cannot be simulated by any (n, K, D)-classical circuit with depth $D \leq \frac{\ln(n)}{4\ln(K)}$ and fan-in $K = \mathrm{polylog}(n)$.*

The remainder of this section is devoted to proving Theorem 3.5. We begin by considering what kinds of classical circuits are capable of reproducing the measurement statistics of GHZ-type entangled quantum states.

Suppose U is an n-qubit Clifford circuit and $|\psi\rangle = U|0^n\rangle$. Recall that we say that three qubits $\Omega = \{h, i, j\} \subseteq [n]$ exhibit GHZ-type MIE if and only if the state $|\psi'\rangle$ obtained after measuring all qubits of $|\psi\rangle$ in $[n] \setminus \Omega$ in the computational basis is locally Clifford equivalent to the GHZ state. We can rephrase this definition in terms of the stabilizer group of $|\psi\rangle$ as follows. Three qubits $\Omega = \{h, i, j\}$ of $|\psi\rangle$ exhibit GHZ-type MIE if and only if there exist single-qubit Clifford unitaries $C^{(h)}$, $C^{(i)}$, and $C^{(j)}$ such that the stabilizer group of $C_h^{(h)} C_i^{(i)} C_j^{(j)} |\psi\rangle$ contains elements

$$\begin{aligned} g_1 &= X_h X_i X_j Z(s) \\ g_2 &= Z_h Z_i I_j Z(t) \\ g_3 &= I_h Z_i Z_j Z(u) \end{aligned} \tag{3.98}$$

for some binary strings $s, t, u \in \{0,1\}^n$ such that $s_r = t_r = u_r = 0$ for every $r \in \Omega$.

In the following, we consider a function $m : \{0,1,2\}^3 \to \{0,1\}^n$ which takes as input $b = b_h b_i b_j \in \{0,1,2\}^3$ describing the measurement bases X, Y, Z for the qubits h, i, j respectively

and produces an output $m(b) \in \{0,1\}^n$. Let $Q(0), Q(1), Q(2)$ denote the single-qubit Clifford unitaries that change basis from the Z basis to the X, Y, Z bases respectively (i.e., $Q(0) = (X+Z)/\sqrt{2}$, $Q(1) = (Y+Z)/\sqrt{2}$, and $Q(2) = I$). For every $b = b_h b_i b_j \in \{0,1,2\}^3$, let $W(b) = Q(b_h)_h Q(b_i)_i Q(b_j)_j$ be the n-qubit unitary that changes the measurement bases for the qubits h, i, j to the bases specified by b. We consider a restricted notion of classical simulation where we say m simulates $|\psi\rangle$ if $m(b)$ is always a possible outcome of measuring all qubits of $|\psi\rangle$ in Ω in the given bases b and measuring all other qubits in the computational basis. Using the notation defined above, the requirement for the restricted notion of classical simulation that we consider can be stated more formally as

$$\langle m(b)|W(b)|\psi\rangle \neq 0 \text{ for every } b \in \{0,1,2\}^3. \tag{3.99}$$

We show a condition under which (3.99) is not possible.

Lemma 3.11. *Suppose the three qubits $\{h,i,j\} \subseteq [n]$ exhibit GHZ-type MIE in the state $|\psi\rangle = U|0^n\rangle$. Let $m : \{0,1,2\}^3 \to \{0,1\}^n$ and write $m = m(b)$ with $b = b_h b_i b_j \in \{0,1,2\}^3$. Suppose that each output bit m_r depends on at most one of the input variables b_h, b_i, b_j, and that m_h is independent of b_i and b_j, m_i is independent of b_h and b_j, and m_j is independent of b_h and b_i. Then it holds that*

$$\Pr_{b \sim \mathrm{unif}(\{0,1,2\}^3)} \left(\langle m(b)|W(b)|\psi\rangle = 0\right) \geq \frac{1}{27} \tag{3.100}$$

where $\mathrm{unif}(\{0,1,2\}^3)$ *is the uniform distribution over* $\{0,1,2\}^3$.

Proof. Without loss of generality, let us fix a local basis so that $C^{(h)}$, $C^{(i)}$, and $C^{(j)}$ are all the identity, so the stabilizer group of $|\psi\rangle$ contains the stabilizers g_1, g_2, g_3 in (3.98). By taking products of g_1, g_2, g_3, we see that the stabilizer group of $|\psi\rangle$ contains the following elements:

$$\begin{aligned}
g_1 &= X_h X_i X_j Z(s) \\
g_2 &= Z_h Z_i I_j Z(t) \\
g_3 &= I_h Z_i Z_j Z(u) \\
g_4 &= -Y_h Y_i X_j Z(s \oplus t) \\
g_5 &= -X_h Y_i Y_j Z(s \oplus u) \\
g_6 &= -Y_h X_i Y_j Z(s \oplus t \oplus u) \\
g_7 &= Z_h I_i Z_j Z(u \oplus t).
\end{aligned} \tag{3.101}$$

Let us assume, to reach a contradiction, that m satisfies (3.99). Write

$$m_s = \prod_{v \in [n]} m_v^{s_v}, \quad m_t = \prod_{v \in [n]} m_v^{t_v}, \quad \text{and} \quad m_u = \prod_{v \in [n]} m_v^{u_v}. \tag{3.102}$$

Note that our function m must only output binary strings that are consistent with stabilizers that are diagonal in the measurement basis b. In this way, each of the stabilizers (3.101) places a constraint on the possible outputs of m.

For example, from the stabilizer g_2, we infer that

$$m_h \cdot m_i \cdot m_t = 1 \quad \text{if} \quad b_h = b_i = 2. \tag{3.103}$$

Since each output bit depends on only one input bit and m_h and m_i are independent of b_j, we can infer from the above that m_t is independent of b_j. We now explain this in more detail. Let $L_h, L_i, L_j \subseteq [n]$ be the set of output bits m_v that depend on inputs b_h, b_i, b_j respectively, so that

$$m_t = (-1)^a \prod_{v \in L_h} m_v(b_h) \prod_{v \in L_i} m_v(b_i) \prod_{v \in L_j} m_v(b_j) \tag{3.104}$$

for some $a \in \{0,1\}$. Then for $b_j \in \{0,1,2\}$, (3.103) implies

$$\prod_{v \in L_j} m_v(b_j) = (-1)^a m_h(2) \cdot m_i(2) \cdot \prod_{v \in L_h} m_v(2) \prod_{v \in L_i} m_v(2) \tag{3.105}$$

where $a \in \{0,1\}$ is fixed. Plugging (3.105) into (3.104) gives

$$m_t = \prod_{v \in L_h} m_v(b_h) \prod_{v \in L_i} m_v(b_i) \left[m_h(2) \cdot m_i(2) \cdot \prod_{v \in L_h} m_v(2) \prod_{v \in L_i} m_v(2) \right], \tag{3.106}$$

which shows that m_t is independent of b_j.

In exactly the same way, from the stabilizers g_3 and g_7, we infer that m_u is independent of b_h, and $m_u \cdot m_t$ is independent of b_i.

Next, consider the subset of inputs $b_h, b_i, b_j \in \{0,1\}$ which correspond to measuring the qubits in $\{h,i,j\}$ in the X or Y basis. Since m_t is independent of b_j and m_u is independent of b_h, there exist affine functions $e, f, g, h : \{0,1\} \to \{0,1\}$ such that for $b_h, b_i, b_j \in \{0,1\}$,

$$m_t = (-1)^{e(b_h)+f(b_i)} \quad \text{and} \quad m_u = (-1)^{g(b_i)+h(b_j)}. \tag{3.107}$$

Since $m_t \cdot m_u$ is independent of b_i, we may set

$$g(b_i) = f(b_i). \tag{3.108}$$

From the stabilizers g_1, g_4, g_5, g_6, we infer that for $b_h, b_i, b_j \in \{0,1\}$,

$$i^{b_h+b_i+b_j} m_h m_i m_j m_s m_t^{b_h} m_u^{b_j} = 1 \quad \text{if} \quad b_h \oplus b_i \oplus b_j = 0. \tag{3.109}$$

Plugging (3.107),(3.108) into (3.109) gives, for $b_h, b_i, b_j \in \{0,1\}$,

$$i^{b_h+b_i+b_j} m_h m_i m_j m_s (-1)^{b_h(e(b_h)+f(b_i))} (-1)^{b_j(f(b_i)+h(b_j))} = 1 \quad \text{if} \quad b_h \oplus b_i \oplus b_j = 0. \tag{3.110}$$

Now, since each output bit m_v depends on at most one of the variables b_h, b_i, b_j, we infer from the above that for $b_h, b_i, b_j \in \{0,1\}$, we have

$$i^{b_h+b_i+b_j} (-1)^{\alpha_0+\alpha_1 b_h+\alpha_2 b_i+\alpha_3 b_j} (-1)^{\beta b_h b_i} (-1)^{\beta b_i b_j} = 1 \quad \text{if} \quad b_h \oplus b_i \oplus b_j = 0 \tag{3.111}$$

for some coefficients

$$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta \in \{0,1\}. \tag{3.112}$$

The claim below shows that we have reached a contradiction, and we conclude that (3.99) is not satisfied by any function m of the form described in the theorem statement. (3.100) then follows directly.

Claim 1. (3.111) *is not satisfied by any choice of the coefficients in* (3.112).

Proof. Assume by way of contradiction that (3.111) holds. We will plug in values for b_h, b_i, b_j and infer a set of linear equations that must be satisfied by the coefficients $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta$.

By plugging in $b_h = b_i = b_j = 0$, we infer $\alpha_0 = 0$.

Next, setting $(b_h, b_i, b_j) \in \{(1,1,0), (1,0,1), (0,1,1)\}$, we get

$$\begin{aligned} \alpha_1 + \alpha_2 + \beta &= 1 \mod 2 \\ \alpha_1 + \alpha_3 &= 1 \mod 2 \\ \alpha_2 + \alpha_3 + \beta &= 1 \mod 2. \end{aligned} \tag{3.113}$$

Summing these three equations leads to a contradiction ($0 = 1 \mod 2$). □

□

Now we are ready to establish the key lemma for proving Theorem 3.5.

Lemma 3.12. *Let $\mathcal{D}$ be a distribution of n-qubit depth-d random Clifford circuits composed of k-qubit gates that satisfies Property 3.2. Assume n is sufficiently large, and in particular, $n \geq d^4$. Suppose each $U \sim \mathcal{D}$ contains R gates, and we label the gates $U_1, \ldots, U_R$. Let F_n be an (n, K, D)-classical circuit such that every output bit of F_n depends on at most $n^{1/4}$ of the input Clifford gates $U_1, \ldots, U_R$. Then for any $r \in \{0,1\}^{\ell(n)}$, it holds that*

$$\Pr_{U \sim \mathcal{D}} \left(\langle F_n(U, r) | U | 0^n \rangle = 0 \right) \geq \frac{c}{27} - O\left(\frac{k}{n^{1/4}} \right) \tag{3.114}$$

where $c > 0$ is the absolute constant from Property 3.2.

Proof. In the following, we fix $r \in \{0,1\}^{\ell(n)}$ and write $F_n(U) = F_n(U, r)$.

Let $U \sim \mathcal{D}$, and let $U_1, \ldots, U_R$ be the k-qubit gates in U. For each qubit $t \in [n]$, there is exactly one k-qubit gate that is the last one to act on qubit t, and we call this gate V_t. Note that we may have $V_t = V_l$ for $t \neq l$.

For each gate $w \in [R]$, let $\mathcal{L}^{\rightarrow}(U_w) \subseteq [n]$ be the union of the set of all output bits of F_n that depend on the gate U_w and $\mathrm{Supp}(U_w)$. Recall that $\mathrm{Supp}(U_w) \subseteq [n]$ is the set of k qubits acted on by the gate U_w. Define a bipartite graph G with a vertex for each gate U_w with $w \in [R]$, a vertex for each output bit $v \in [n]$ of F_n, and an edge $\{U_w, v\}$ whenever $v \in \mathcal{L}^{\rightarrow}(U_w)$. Since each output bit of F_n depends on at most $n^{1/4}$ input gates and each qubit can be in the support of at most d gates in U, the maximum degree of any $v \in [n]$ in G is at most $d + n^{1/4} \leq 2n^{1/4}$ since we assumed $n \geq d^4$. Therefore, the total number of edges in G satisfies

$$\sum_{w=1}^{R} |\mathcal{L}^{\rightarrow}(U_w)| \leq 2n^{5/4}. \tag{3.115}$$

Define

$$\mathrm{Bad} = \{t \in [n] : |\mathcal{L}^{\rightarrow}(V_t)| \geq k n^{1/2}\} \quad \text{and} \quad \mathrm{Good} = [n] \setminus \mathrm{Bad}. \tag{3.116}$$

Using (3.115), observe that

$$kn^{1/2}|\text{Bad}| \leq \sum_{t\in\text{Bad}} |\mathcal{L}^{\rightarrow}(V_t)| \leq k\sum_{w=1}^{R} |\mathcal{L}^{\rightarrow}(U_w)| \leq 2kn^{5/4}. \tag{3.117}$$

Therefore, we have

$$|\text{Bad}| \leq 2n^{3/4} \quad \text{and} \quad |\text{Good}| \geq n(1-2n^{-1/4}). \tag{3.118}$$

Define another graph G' with vertex set Good, and there is an edge $\{s,t\}$ if and only if

$$\mathcal{L}^{\rightarrow}(V_s) \cap \mathcal{L}^{\rightarrow}(V_t) \neq \emptyset. \tag{3.119}$$

The maximum degree Δ of G' satisfies

$$\Delta \leq \max_{t\in\text{Good}} |\mathcal{L}^{\rightarrow}(V_t)| \cdot (d+n^{1/4}) \leq kn^{1/2}\cdot 2n^{1/4} = 2kn^{3/4}. \tag{3.120}$$

Now suppose we choose a triplet of qubits $\Omega = \{h,i,j\} \subseteq [n]$ uniformly at random. From (3.118), we see that the probability that $\Omega \subseteq \text{Good}$ is $1-O(1/n^{1/4})$. Conditioned on this event, the probability that Ω forms an independent set in G' can be lower bounded as $1-O(k/n^{1/4})$. To see this, note that since the maximum degree of G' is at most $O(kn^{3/4})$, two randomly chosen vertices in Good have an edge between them with probability at most $O(k/n^{1/4})$. Then apply this to the three pairs $\{h,i\}$, $\{i,j\}$, and $\{h,j\}$ and use the union bound. Let $\mathcal{E}$ be the event that Ω forms an independent set in G'. Then we have shown that $\Pr(\mathcal{E}) \geq 1-O(k/n^{1/4})$. Moreover, if $\mathcal{E}$ occurs, then

$$\begin{aligned}
\mathcal{L}^{\rightarrow}(V_h) \cap \mathcal{L}^{\rightarrow}(V_i) &= \emptyset \\
\mathcal{L}^{\rightarrow}(V_i) \cap \mathcal{L}^{\rightarrow}(V_j) &= \emptyset \\
\mathcal{L}^{\rightarrow}(V_h) \cap \mathcal{L}^{\rightarrow}(V_j) &= \emptyset.
\end{aligned} \tag{3.121}$$

Let $|\psi\rangle = U|0^n\rangle$ be the output state of the circuit U. Let $\mathcal{E}'$ denote the event that Ω exhibits GHZ-type MIE in $|\psi\rangle$. The event $\mathcal{E}'$ occurs if and only if the state $|\psi'\rangle$ obtained by measuring all qubits of $|\psi\rangle$ in $[n]\setminus\Omega$ in the computational basis is GHZ-type entangled. By the definition of Property 3.2, the event $\mathcal{E}'$ occurs with probability at least $c>0$.

By the union bound, the probability that both events $\mathcal{E}$ and $\mathcal{E}'$ occur is at least

$$\Pr(\mathcal{E} \text{ and } \mathcal{E}') \geq c - O(k/n^{1/4}). \tag{3.122}$$

Now let us suppose that events $\mathcal{E}$ and $\mathcal{E}'$ both occur. Starting with our circuit U, we define a set of 27 quantum circuits as follows. For each of the qubits h,i,j, there are three choices for a single-qubit Pauli basis X,Y,Z. Let us index these choices by a string $b = b_h b_i b_j \in \{0,1,2\}^3$ where $0,1,2$ correspond to X,Y,Z respectively. Recall that we write $Q(0),Q(1),Q(2)$ for the single-qubit Clifford unitaries that change basis from the Z basis to the X,Y,Z bases respectively (i.e., $Q(0) = (X+Z)/\sqrt{2}$, $Q(1) = (Y+Z)/\sqrt{2}$, and

$Q(2) = I$). Without loss of generality, assume h, i, j are the first qubit in the supports of V_h, V_i, V_j respectively. For every $b = b_h b_i b_j \in \{0,1,2\}^3$, define

$$\begin{aligned} V_h(b_h) &= (Q(b_h) \otimes I^{\otimes k-1}) V_h \\ V_i(b_i) &= (Q(b_i) \otimes I^{\otimes k-1}) V_i \\ V_j(b_j) &= (Q(b_j) \otimes I^{\otimes k-1}) V_j, \end{aligned} \tag{3.123}$$

and let $U(b)$ denote the circuit obtained from U by making the replacements $V_h \leftarrow V_h(b_h)$, $V_i \leftarrow V_i(b_i)$, and $V_j \leftarrow V_j(b_j)$.

Define a function $m : \{0,1,2\}^3 \to \{0,1\}^n$ by $m(b) = F_n(U(b))$. From (3.121), we see that each output bit m_j depends on at most one of the bits b_h, b_i, b_j, and that m_h is independent of b_i and b_j, m_i is independent of b_h and b_j, and m_j is independent of b_h and b_i. Therefore, by Lemma 3.11, we conclude that

$$\langle F_n(U(b)) | U(b) | 0^n \rangle = 0 \text{ for some } b \in \{0,1,2\}^3. \tag{3.124}$$

Note that each of the circuits $U(b)$ with $b \in \{0,1,2\}^3$ occurs in our random circuit distribution $\mathcal{D}$ with the same probability as U. Therefore,

$$\Pr_{U \sim \mathcal{D}} \left(\langle F_n(U) | U | 0^n \rangle = 0 \right) \geq \frac{\Pr(\mathcal{E} \text{ and } \mathcal{E}')}{27} \geq \frac{c}{27} - O\left(\frac{k}{n^{1/4}} \right). \tag{3.125}$$

□

Proof of Theorem 3.5. Let F_n be an (n, K, D)-classical circuit. Assume $D \leq \ln(n)/4\ln(K)$. For $U \sim \mathcal{D}$, recall that F_n takes as input the gates $U_1, \ldots, U_R$ in the Clifford circuit U as well as a random bit string $r \in \{0,1\}^{\ell(n)}$ drawn from some distribution q_n, and outputs a binary string $x = F_n(U_1, \ldots, U_R, r) = F_n(U, r)$.

Since $D \leq \ln(n)/4\ln(K)$, every output bit of F_n depends on at most $K^D \leq n^{1/4}$ input bits. In particular, each output bit of F_n depends on at most $n^{1/4}$ of the input gates $U_1, \ldots, U_R$. Therefore, by Lemma 3.12, we have

$$\mathop{\mathbb{E}}_{U \sim \mathcal{D}} \left[\sum_{x \in \{0,1\}^n} |P_U(x) - Q_U(x)| \right] \tag{3.126}$$

$$\geq \mathop{\mathbb{E}}_{U \sim \mathcal{D}} \left[\sum_{x \in \{0,1\}^n \,:\, Q_U(x) = 0} P_U(x) \right] \tag{3.127}$$

$$= \mathop{\mathbb{E}}_{U \sim \mathcal{D}} \left[\Pr_{r \sim q_n} \left(\langle F_n(U, r) | U | 0^n \rangle = 0 \right) \right] \tag{3.128}$$

$$= \mathop{\mathbb{E}}_{r \sim q_n} \left[\Pr_{U \sim \mathcal{D}} \left(\langle F_n(U, r) | U | 0^n \rangle = 0 \right) \right] \tag{3.129}$$

$$\geq \mathop{\mathbb{E}}_{r \sim q_n} \left[\frac{c - O(k/n^{1/4})}{27} \right] \tag{3.130}$$

$$= \frac{c - O(k/n^{1/4})}{27}. \tag{3.131}$$

We assume n is large enough so that $O(k/n^{1/4})$ is at most $c/2$ and then set $\alpha = c/54$. □

Chapter 4

The stabilizer Scrooge ensemble and shallow 2D random Clifford circuits

In Chapter 3, we showed the emergence of long-range measurement-induced entanglement (MIE) in the output states of low-depth random Clifford circuits. In this chapter, we focus on developing a more detailed understanding of the computational-basis output distributions of shallow 2D random Clifford circuits. This chapter is based on the paper [3].

The output distribution of a Clifford circuit is fully determined by the Z-type stabilizer subgroup of its output state. More concretely, let U be an n-qubit Clifford circuit with output state $|\psi\rangle = U|0^n\rangle$, and let S_ψ^Z denote the Z-type stabilizer subgroup of $|\psi\rangle$. In the Clifford setting, the output distribution of U is always the uniform distribution over an affine subspace of $\{0,1\}^n$ given by

$$\{x \in \{0,1\}^n : P|x\rangle = |x\rangle \text{ for every } P \in S_\psi^Z\}. \tag{4.1}$$

Each $P \in S_\psi^Z$ imposes a linear constraint on a subset of the bits that a possible outcome string must satisfy. Equivalently, each $P \in S_\psi^Z$ (of weight at least 2) enforces some correlation among the bits of an outcome string. As a result, our goal is made equivalent to understanding the statistical properties of just the Z-type stabilizers of the output states of shallow 2D random Clifford circuits.

Like what we did for MIE, we first probe our problem using numerical simulations.

4.1 Numerical investigation of the output distributions of shallow 2D random Clifford circuits

In this section, we investigate numerically the key question of this chapter: what does the output distribution of a typical constant-depth 2D brickwork random Clifford circuit look like? We consider the canonical 2D brickwork architecture described in Section 1.3.3 and Figure 1.2. In this chapter, all numerical simulations of Clifford circuits are performed using Stim [35].

We will focus our attention on depth-7 2D brickwork random Clifford circuits because it has been observed in previous numerical investigations that with respect to this specific architecture, depth-7 is right after the conjectured complexity phase transition has taken place; see Figure 3.2.

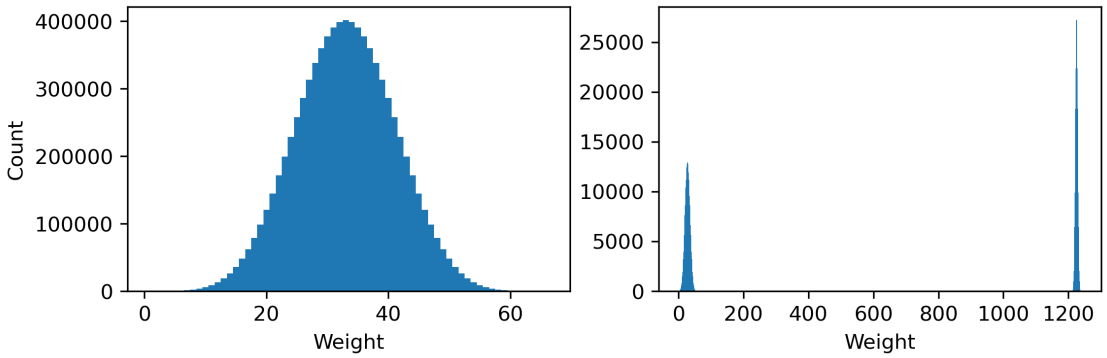


Figure 4.1: Examples of typical weight distributions of the Z-type stabilizers of the output states of depth-7 2D brickwork random Clifford circuits acting on a 50×50 grid of qubits. The example on the left panel shows a unimodal histogram over a total of $2^{23}-1$ non-identity Z-type stabilizers centred around a weight of 33. The right panel shows a sharply different bimodal histogram over a total of $2^{19}-1$ non-identity Z-type stabilizers, with a low-weight peak centred around 28 and a far-apart high-weight peak centred around 1225. In both cases, the total number of non-identity Z-type stabilizers, $2^{23}-1$ and $2^{19}-1$ respectively, far exceed the average value of roughly 1 expected in the anti-concentration regime.

We first investigate the output distributions of typical 2D brickwork random Clifford circuits when all the qubits are measured in the computational basis. As we already explained, in the Clifford setting, this amounts to looking at the Z-type stabilizer subgroup of the output state.

What do the Z-type stabilizers of the output state of a typical depth-7 2D brickwork random Clifford circuit look like? We know that on average, an output state should have at least $\exp(\Omega(n))$ Z-type stabilizers in total [57] where the constant in $\Omega()$ scales inverse exponentially with the circuit depth. In contrast, when the circuit depth reaches the anti-concentration regime, which is believed to occur at $\Omega(\log n)$ depth for our specific architecture, the expected total number of non-identity Z-type stabilizers converges to approximately 1 [57]. Recall that on average, an n-qubit uniformly random stabilizer state has roughly one non-identity Z-type stabilizer (see (1.9)), whose weight is concentrated around $\frac{n}{2}$. The weight of a Pauli operator P is $|\mathrm{Supp}(P)|$. We numerically simulated depth-7 2D brickwork random Clifford circuits acting on a 50×50 grid of qubits and solved for the Z-type stabilizer subgroups of the output states. We found that the most common Z-type stabilizer weight distributions/histograms follow the shapes exemplified in Figure 4.1.

In the case depicted in the right panel of Figure 4.1, it is possible to find a set of independent Z-type stabilizer generators for the Z-type stabilizer subgroup consisting of 18 low-weight generators, whose supports are confined to small-sized local regions, and a single high-weight generator, whose support spans the entire grid. In this particular case, due to the clear gap between the two peaks in the weight histogram, we can be certain that the sole high-weight generator cannot be obtained as a product involving only low-weight Z-type stabilizers. We say that this high-weight generator is indecomposable, in the sense that it cannot be decomposed into a product of a large number of Z-type stabilizers, each supported on a local region. An indecomposable Z-type stabilizer embeds an indecomposable

long-range correlation into the output distribution, serving as a signal of global randomness in the output state. Note that a Z-type stabilizer being high-weight does not mean that it must be indecomposable; for example, $Z^{\otimes n}$ is a maximal weight Z-type stabilizer of the state $|0^n\rangle$, but $Z^{\otimes n}$ is clearly decomposable as $Z^{\otimes n} = \prod_{j=1}^{n} Z_j$. A high-weight Z-type stabilizer being present in a particular generating set also does not mean that it is indecomposable, as we could extend $\{Z^{\otimes n}\}$ to a generating set for the stabilizer group of $|0^n\rangle$, which will include the high-weight but decomposable $Z^{\otimes n}$ as a generator by construction.

Motivated by our findings so far, we move on to investigate the distribution of the number of indecomposable Z-type stabilizers of the output state $|\psi\rangle$ of a constant-depth 2D brickwork random Clifford circuit. To make this question well-defined, we consider a tripartition of the $\sqrt{n}\times\sqrt{n}$ grid where we take A to be the top row of qubits, C to be the bottom row of qubits, B to be all the other qubits, and consider a rather stringent notion of decomposability. Let S_ψ^Z denote the Z-type stabilizer subgroup of $|\psi\rangle$, and for every $D \subseteq [n]$, let $S_{\psi_D}^Z$ denote the subgroup of S_ψ^Z local on D. We call $P \in S_\psi^Z$ indecomposable if $P \neq QR$ for every $Q \in S_{\psi_{AB}}^Z$ and $R \in S_{\psi_{BC}}^Z$ and decomposable otherwise. This operational notion of decomposability is captured precisely by a quotient group structure: every indecomposable $P \in S_\psi^Z$ belongs to a non-identity element of the quotient group $S_\psi^Z/(S_{\psi_{AB}}^Z S_{\psi_{BC}}^Z)$. As a result, the number of indecomposable Z-type stabilizer generators is given by

$$\log_2\left(\frac{|S_\psi^Z|}{|S_{\psi_{AB}}^Z S_{\psi_{BC}}^Z|}\right) \tag{4.2}$$

$$= \log_2(|S_\psi^Z|) - \log_2(|S_{\psi_{AB}}^Z|) - \log_2(|S_{\psi_{BC}}^Z|) + \log_2(|S_{\psi_B}^Z|) \tag{4.3}$$

$$= (|AB| - \log_2(|S_{\psi_{AB}}^Z|)) + (|BC| - \log_2(|S_{\psi_{BC}}^Z|)) - (n - \log_2(|S_\psi^Z|)) - (|B| - \log_2(|S_{\psi_B}^Z|)) \tag{4.4}$$

$$= H(AB) + H(BC) - H(ABC) - H(B) \tag{4.5}$$

$$= H(A|B) + H(C|B) - H(AC|B) \tag{4.6}$$

$$= I(A;C|B) \tag{4.7}$$

where H denotes the Shannon entropy, and $I(A;C|B)$ is the conditional mutual information between A and C given B in the computational-basis output distribution of $|\psi\rangle$.

In the following, we numerically simulate depth-7 2D brickwork random Clifford circuits acting on a 100×100 grid of qubits and compute the value of $I(A;C|B)$ for each circuit instance. Since

$$I(A;C|B) = \log_2(|S_\psi^Z|) - \log_2(|S_{\psi_{AB}}^Z|) - \log_2(|S_{\psi_{BC}}^Z|) + \log_2(|S_{\psi_B}^Z|) \tag{4.8}$$

$$= \log_2(|S_\psi^Z/S_{\psi_B}^Z|) - \log_2(|S_{\psi_{AB}}^Z/S_{\psi_B}^Z|) - \log_2(|S_{\psi_{BC}}^Z/S_{\psi_B}^Z|), \tag{4.9}$$

we can compute $I(A;C|B)$ using the following procedure. For each circuit instance U, we first compute its output state $|\psi\rangle = U|0^n\rangle$ and then measure all the qubits of $|\psi\rangle$ in B in the computational basis to get a pure post-measurement state $|\psi'\rangle$ on $A \cup C$. Next, we solve for a set of independent generators G^Z for the Z-type stabilizer subgroup of $|\psi'\rangle$ and sets of independent generators G_A^Z and G_C^Z for the Z-type stabilizer subgroups of $|\psi'\rangle$ local on A and C, respectively. Then we have $I(A;C|B) = |G^Z| - |G_A^Z| - |G_C^Z|$, which can also be interpreted as the number of independent non-local Z-type stabilizer generators of $|\psi'\rangle$. Table 4.1 records

r	$\Pr(I(A;C\|B)=r)$ 2D Clifford depth-5	$\Pr(I(A;C\|B)=r)$ 2D Clifford depth-6	$\Pr(I(A;C\|B)=r)$ 2D Clifford depth-7	$\Pr(I(A;C\|B)=r)$ random stab. state as $n\to\infty$
0	0.99999	0.46207	0.41815	0.41942
1	0.00001	0.41336	0.42077	0.41942
2	0	0.11267	0.13952	0.13981
3	0	0.01141	0.02022	0.01997
4	0	0.00049	0.00131	0.00133
5	0	0	0.00003	0.00004

Table 4.1: Numerical estimation of the distribution of the integer $r = I(A;C|B)$ for depth-$d \in \{5,6,7\}$ 2D brickwork random Clifford circuits acting on a 100×100 grid of qubits using 10^5 random circuit samples.

the data obtained from numerically estimating the distribution of the integer $r = I(A;C|B)$ using 10^5 random circuit samples for depth-d 2D brickwork random Clifford circuits with $d \in \{5,6,7\}$. The distribution catalogued in the last column for uniformly random stabilizer states in the large-n limit will be worked out in Section 4.2.

At depth-5, the value $I(A;C|B) = 0$ is observed in all but one of the 10^5 runs. In contrast, we see that long-range conditional mutual information is present at $d = 7$, which is believed to be above the critical depth for the 2D brickwork architecture; this is to be expected as a consequence of long-range MIE. However, we also observe an unexpected signature of global randomness—the probability distribution that governs the conditional mutual information appears to be indistinguishable from that of a uniformly random stabilizer state (as $n \to \infty$). Our findings at $d = 6$ are inconclusive. It is possible that there are strong finite-size effects near criticality, and the distribution of $I(A;C|B)$ for $d = 6$ will eventually converge to either the distribution for $d = 5$ or $d = 7$ as $n \to \infty$. However, given the data, it is also possible that as $n \to \infty$, the distributions of $I(A;C|B)$ are qualitatively distinct for $d = 6$ and $d = 7$ even though extensive long-range MIE is already present at $d = 6$. The results in Table 4.1 predict that the unimodal and bimodal weight histograms shown in Figure 4.1 each occur with probability approximately 0.419, corresponding to the cases $r = 0$ and $r = 1$. Our numerical findings strongly suggest that some form of global randomness resides within the output distributions of depth-7 2D brickwork random Clifford circuits.

To better understand the characteristics of the global randomness in the output distributions of depth-7 2D brickwork random Clifford circuits predicted by our previous numerics, we next look at the marginal output distributions of those circuits. Recall that for an n-qubit uniformly random stabilizer state $|\phi\rangle$, the computational-basis marginal output distribution of $|\phi\rangle$ on a subset A of $m = n - \Omega(\log n)$ qubits will be the uniform distribution over $\{0,1\}^m$ with high probability. We can again understand this intuitively through the Z-type stabilizers S_ϕ^Z of $|\phi\rangle$. The marginal output distribution is dictated by the Z-type stabilizer subgroup of the reduced state $\mathrm{Tr}_{[n]\setminus A}(|\phi\rangle\langle\phi|)$, which is in bijection with $S_{\phi_A}^Z$, the subgroup of S_ϕ^Z local on A. Recall that $|\phi\rangle$ typically has at most a few high-weight Z-type stabilizers, so that with high probability, all have non-trivial support on the $\Omega(\log n)$-sized subset $[n] \setminus A$.

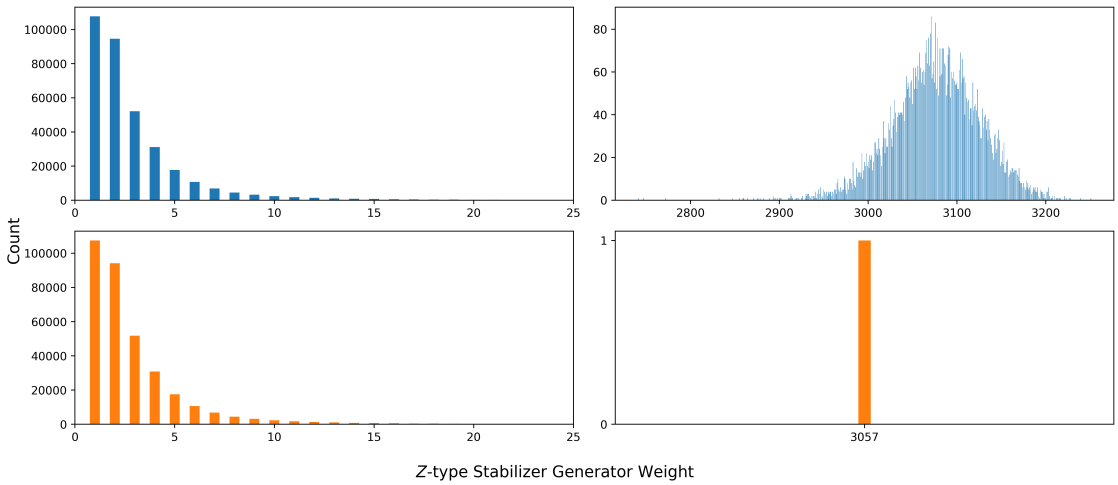


Figure 4.2: Weight histograms of the Z-type stabilizer generators of the full output states (top left and top right panels) versus reduced output states (bottom left and bottom right panels) of (80×80)-qubit depth-7 2D brickwork random Clifford circuits aggregated over 10^4 circuit instances. The top left and right panels show that a full output state tends to admit a Z-type stabilizer subgroup generating set containing many low-weight (at most 166) generators and perhaps a few high-weight (around 3000) generators. There are 1603 generators with weights between 26 and 166 not shown in the top left panel. In our experiment, after tracing out 15 random qubits, the reduced output state admits a Z-type stabilizer subgroup generating set with only low-weight (at most 166) generators in all but one out of the 10^4 circuit instances for which the solved generating set contained an element of weight 3057. There are 1485 generators with weights between 26 and 166 not shown in the bottom left panel.

Equivalently, with high probability, $S^Z_{\phi_A} = \emptyset$, so the marginal output distribution will be uniform.

We do not expect the marginal output distribution of a depth-7 2D brickwork random Clifford circuit to be uniform on a large region A of size $n - \Omega(\log n)$ because the output state will have many low-weight Z-type stabilizers with support contained in A. However, what could happen is that the reduced state $\mathrm{Tr}_{[n]\setminus A}(|\psi\rangle\langle\psi|)$ on A, with $|\psi\rangle$ being the output state of the circuit, may no longer have any high-weight indecomposable Z-type stabilizers, for the same reason that all the indecomposable Z-type stabilizers are inevitably supported non-trivially on $[n]\setminus A$. Note that this can be certified by the existence of a set of independent generators for $S^Z_{\psi_A}$ for which every generator has low-weight and is locally supported.

In this experiment, we numerically simulate (80×80)-qubit depth-7 2D brickwork random Clifford circuits with 10^4 circuit samples. For each circuit instance U with output state $|\psi\rangle = U|0^n\rangle$, we first solve for a set of independent generators for S^Z_ψ, then choose a uniformly random subset B of 15 qubits, and lastly solve for a set of independent generators for $S^Z_{\psi_A}$ with $A = [n] \setminus B$. Figure 4.2 plots the aggregated weight histograms of all the generators encountered throughout the 10^4 circuit samples. As we have explained earlier, while the presence of a high-weight Z-type stabilizer in a generating set does not imply that it is necessarily indecomposable, we do suspect that all the high-weight generators encountered in

this experiment are actually indecomposable due to the significant weight difference between the low-weight and high-weight generators. In contrast, in 9999 out of the 10^4 instances, after tracing out $|B| = 15$ random qubits, $S^Z_{\psi_A}$ admits an independent generating set consisting of only low-weight generators, and this fact does imply that every Z-type stabilizer of the reduced state is decomposable. Similar to uniformly random stabilizer states, our numerics suggest that any global correlation present in the output distribution of $|\psi\rangle$ ceases to exist in its marginal output distributions even after tracing out just a small subset of qubits.

In summary, the following intuitive picture emerges from our numerical investigation of the output distribution, or equivalently the Z-type stabilizer subgroup, of depth-7 2D brickwork random Clifford circuits. From Figure 4.1, we infer that the Z-type stabilizer subgroup S^Z_ψ of the output state of a typical circuit instance $|\psi\rangle = U|0^n\rangle$ admits an independent generating set consisting of a large number of low-weight local Z-type generators and perhaps a few high-weight non-local generators. From Table 4.1, we infer that the distribution of the number of indecomposable Z-type stabilizer generators closely matches the distribution of the number of Z-stabilizer generators of a uniformly random stabilizer state. From Figure 4.2, we infer that just like for uniformly random stabilizer states, any global correlation present in the output distribution of $|\psi\rangle$ is fragile against depolarizing noise; the Z-type stabilizer subgroup of the reduced state of $|\psi\rangle$ admits a generating set composed solely of low-weight local Z-type stabilizers after tracing out only a small number of qubits.

4.2 The stabilizer Scrooge ensemble

A Scrooge ensemble or distribution is a distribution of pure states that generalizes the distribution of Haar random states [88]. A Scrooge ensemble is defined by a mixed background state ρ. Given a ρ, there are many equivalent ways to define its Scrooge ensemble. We give one definition below.

Definition 4.1 (Scrooge ensemble)**.** Let ρ be an n-qubit mixed state. Let $|\Psi\rangle$ be a purification of ρ over registers A and B with $|A| = n$ and $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = \rho$. The Scrooge ensemble defined by ρ is the distribution of pure post-measurement states $|\psi\rangle$ obtained by running the following quantum circuit. On initial state $|\Psi\rangle$, we first apply a $|B|$-qubit Haar random unitary U to the B register and then measure all qubits in B in the computational basis to obtain an outcome $x_B \in \{0,1\}^{|B|}$ and an n-qubit pure post-measurement state $|\psi\rangle$ on A.

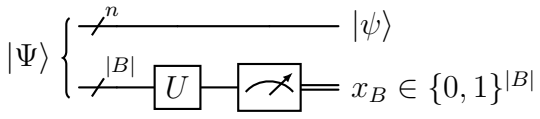


Figure 4.3: Circuit diagram for drawing from a Scrooge ensemble. The $|B|$-qubit unitary U is Haar random.

Figure 4.3 shows a circuit diagram. Note that since the $|B|$-qubit unitary U is Haar random, by its unitary invariance property, all purifications lead to the same Scrooge ensemble. We see that when ρ is the n-qubit maximally mixed state, we can take $|\Psi\rangle$ to be

the $2n$-qubit maximally entangled state. In this case, by the transpose trick, we see that the Scrooge ensemble reduces to the distribution of n-qubit Haar random states. On the other extreme, when $\rho = |\psi\rangle\langle\psi|$ is pure, the ensemble becomes degenerate and produces the pure state $|\psi\rangle$ with probability one.

We can also define a Scrooge ensemble in the Clifford/stabilizer setting that generalizes the uniform distribution over stabilizer states in an analogous way.

Definition 4.2 (Stabilizer Scrooge ensemble)**.** Let ρ be an n-qubit mixed stabilizer state. Let $|\Psi\rangle$ be a stabilizer purification of ρ over registers A and B with $|A| = n$ and $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = \rho$. The stabilizer Scrooge ensemble $\mathcal{S}_\rho$ defined by ρ is the distribution of pure post-measurement states $|\psi\rangle$ obtained by running the following quantum circuit. On initial state $|\Psi\rangle$, we first apply a $|B|$-qubit uniformly random Clifford unitary U to the B register and then measure all qubits in B in the computational basis to obtain an outcome $x_B \in \{0,1\}^{|B|}$ and an n-qubit pure post-measurement state $|\psi\rangle$ on A.

For stabilizer Scrooge, the circuit diagram is the same as Figure 4.3, except that the unitary U is Clifford. Similar to the Haar random case, the exact purification does not matter, and the stabilizer Scrooge ensemble defined by the n-qubit maximally mixed state reduces to the uniform distribution over n-qubit stabilizer states.

Why are we talking about Scrooge ensembles when the goal of this chapter is to understand the output distributions of shallow 2D random Clifford circuits? In fact, a clean example of Scrooge randomness has already appeared in the analysis of long-range MIE in 1D coarse-grained random Clifford circuits in Section 3.2. By pattern matching each circuit diagram shown in the proof of Lemma 3.3 with Figure 4.3, we realize that the 1D coarse-grained random Clifford circuit with measurements shown in Figure 3.5 can be interpreted as iteratively drawing from Scrooge ensembles, where the background state for one iteration is sampled randomly based on all previous iterations. Something similar can be said about 2D coarse-grained random Clifford circuits, so long as all the qubits acted on by a second-layer gate are measured in the computational basis.

While Scrooge ensembles reside naturally in coarse-grained random circuits in a local sense, in the paper [3], my collaborators and I conjecture that Scrooge randomness dictates the structure of conditional dependence in the output states of shallow 2D random quantum circuits composed of either Haar random or random Clifford gates in a global sense. The conjecture pertains to depth $d > d^*$ circuits where d^* is the critical depth at which the complexity phase transition occurs. To concretely state the conjecture, we will focus on the Clifford setting.

Let $[n] = A \cup B \cup C$ be a tripartition of n qubits living on a $\sqrt{n} \times \sqrt{n}$ grid. Let $\mathcal{U}$ denote the distribution of depth-d 2D brickwork random Clifford circuits with $d > d^*$. For a circuit $U \sim \mathcal{U}$, we can always write U as $U = U_C U_{AB}$, where U_{AB} contains all the gates of U in the backward lightcone of $A \cup B$, and U_C contains all the other gates. Note that U_{AB} acts on some qubits $C' \subseteq C$ at the boundary of C, where $|C'|$ depends on the circuit depth d. Then, $\mathcal{U}$ can be analogously separated into two independent constituents $\mathcal{U}_{AB}$ and $\mathcal{U}_C$ so that $U_{AB} \sim \mathcal{U}_{AB}$ and $U_C \sim \mathcal{U}_C$. Consider a distribution $\mathcal{B}$ of $|A|$-qubit mixed stabilizer states defined by the following protocol. We draw $U_{AB} \sim \mathcal{U}_{AB}$ and measure all qubits of $U_{AB}|0^n\rangle$ in B in the computational basis to obtain an outcome $x_B \in \{0,1\}^{|B|}$ and a pure post-measurement state $|\varphi\rangle_{AC} = |\psi\rangle_{AC'} \otimes |0^{|C\setminus C'|}\rangle$. Then $\mathcal{B}$ is the distribution of the state

$\rho = \mathrm{Tr}_C(|\varphi\rangle\langle\varphi|)$, with randomness over $U_{AB} \sim \mathcal{U}_{AB}$ and the measurement outcome x_B. Note that in the Clifford setting, for fixed $U_{AB} \sim \mathcal{U}_{AB}$, the measurement outcome x_B only affects the $\pm$ signs of the stabilizers of ρ. Conditioned on a fixed $|\varphi\rangle_{AC}$ and ρ, we will compare the Scrooge ensemble $\mathcal{S}_\rho$ defined by ρ to another distribution of $|A|$-qubit pure stabilizer states $\mathcal{D}_\rho$ obtained by drawing $U_C \sim \mathcal{U}_C$ and measuring all qubits of $U_C|\varphi\rangle_{AC}$ in C in the computational basis. Note that for $|\psi'\rangle \sim \mathcal{D}_\rho$, the measurement outcome on C only affects the $\pm$ signs of the stabilizers of $|\psi'\rangle$ and that strictly speaking, $\mathcal{D}_\rho$ depends on not just ρ but also the specific purification $|\psi\rangle_{AC'}$ of ρ. Notice that if U_C is drawn from the uniform distribution over $|C|$-qubit Clifford unitaries, then by Definition 4.2, $\mathcal{D}_\rho = \mathcal{S}_\rho$. Our Scrooge hypothesis postulates that $\mathcal{D}_\rho$ is well-approximated by $\mathcal{S}_\rho$ under suitable conditions in the following specific sense.

Conjecture 4.1 (Scrooge hypothesis)**.** *Let $[n] = A \cup B \cup C$ be a tripartition of a $\sqrt{n} \times \sqrt{n}$ grid of qubits. Let the distributions $\mathcal{B}$, $\mathcal{S}_\rho$, and $\mathcal{D}_\rho$ be defined as above. Fix $t = O(1)$ and a $2^{|A|t}$-by-$2^{|A|t}$ Hermitian positive semidefinite matrix O. Assume C contains an $l_C \times l_C$ square of qubits. Then for*

$$\varepsilon = \mathrm{poly}(n) e^{-\Omega(\min(l_C, \mathrm{dist}(A,C)))}, \tag{4.10}$$

with probability $1 - \varepsilon$ over the randomness of drawing $\rho \sim \mathcal{B}$, it holds that

$$(1-\varepsilon) \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} [\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}] \leq \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}_\rho} [\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}] \leq (1+\varepsilon) \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} [\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}], \tag{4.11}$$

and the background state ρ has min-entropy $S_\infty(\rho) = \Omega(|\partial A|)$. In the above, $|\partial A|$ denotes the number of qubits on the boundary of A and $\mathrm{dist}(A, C) = \min_{x\in A, y\in C} \mathrm{dist}(x, y)$ denotes the minimum L1 distance between $x \in A$ and $y \in C$ on the 2D grid. The $\Omega()$ notation hides factors depending on t and the circuit depth d.

Note that in the Scrooge hypothesis, the observable O must be fixed beforehand and is not allowed to depend on the random background state ρ. For a $|A|$-qubit stabilizer state ρ, its Renyi entropy of order α is the same for every $\alpha \in [0, \infty]$, so if ρ has 2^k stabilizers, then $S_\infty(\rho) = |A| - k$. Notice that $S_\infty(\rho)$ is precisely the number of ebits of entanglement between A and C in the purification $|\varphi\rangle_{AC}$, so the Scrooge hypothesis includes in itself the existence of extensive long-range MIE.

In the remainder of this section, we develop a better understanding of the Z-type stabilizers of random states drawn from a stabilizer Scrooge ensemble, with a particular focus on the case where the background state is a product state $\rho = \rho_A \otimes \rho_B$. In Section 4.3, we show that the Scrooge hypothesis implies detailed structural properties of the output distributions of shallow 2D random Clifford circuits that are very suggestive of what we observed numerically in Section 4.1. In Section 4.4, we provide numerical evidence for the Scrooge hypothesis in the Clifford setting. The paper [3] contains further analytical evidence for the Scrooge hypothesis in the Haar random setting that is not included in this thesis.

We first provide a useful structural characterization of stabilizer Scrooge distributions.

Lemma 4.1. *Let ρ be an n-qubit mixed stabilizer state with independent stabilizer generators $\{g_1, \ldots, g_k\}$. Let U be any n-qubit Clifford unitary such that $\{UZ_jU^\dagger : j \in [k]\} = \{g_1, \ldots, g_k\}$. Then the stabilizer Scrooge distribution defined by ρ is the uniform distribution over the set $\{U(|0^k\rangle \otimes |\phi\rangle) : |\phi\rangle$ is an $(n-k)$-qubit stabilizer state$\}$.*

Proof. By the definition of U, we have

$$\rho = \frac{1}{2^{n-k}} \prod_{j=1}^{k} \frac{I^{\otimes n} + g_j}{2} = \frac{1}{2^{n-k}} \prod_{j=1}^{k} \frac{I^{\otimes n} + UZ_jU^\dagger}{2} = \frac{1}{2^{n-k}} U(|0^k\rangle\langle 0^k| \otimes I^{\otimes n-k})U^\dagger. \tag{4.12}$$

We define A_1 to be the first k qubits of ρ and A_2 to be the last $n-k$ qubits of ρ. Notice that $|\Psi\rangle = |0^k\rangle_{A_1} \otimes |\Phi\rangle_{A_2B}$ is a purification of $\frac{1}{2^{n-k}}|0^k\rangle\langle 0^k| \otimes I^{\otimes n-k}$ where $|\Phi\rangle$ is the $2(n-k)$-qubit maximally entangled state and the ancilla register B contains $n-k$ qubits. Then $U_{A_1A_2}|\Psi\rangle$ is a purification of ρ. Thus, by Definition 4.2 and the transpose trick applied with $|\Phi\rangle$, the claim follows. □

Given ρ with stabilizer generators $\{g_1, \ldots, g_k\}$, one way to find a U that satisfies $\{UZ_jU^\dagger : j \in [k]\} = \{g_1, \ldots, g_k\}$ is to extend $\{g_1, \ldots, g_k\}$ to a complete set $\{g_1, \ldots, g_k, g_{k+1}, \ldots, g_n\}$ of independent stabilizer generators for some n-qubit pure stabilizer state $|\psi\rangle$ and take a U that satisfies $|\psi\rangle = U|0^n\rangle$.

When ρ has 2^k stabilizers for some $k \in \{0, 1, \ldots, n\}$, we can think of ρ as representing a stabilizer code that encodes $n-k$ logical qubits. Lemma 4.1 shows that under this interpretation, the stabilizer Scrooge distribution is the uniform distribution over the set of all codewords of the stabilizer code represented by ρ. Before we proceed, we describe another way to sample from a stabilizer Scrooge ensemble.

Definition 4.3 (Alternative stabilizer Scrooge sampler)**.** Let ρ be an n-qubit mixed stabilizer state.

1. Draw a uniformly random n-qubit stabilizer state $|\varphi\rangle$.
2. If $\rho|\varphi\rangle = 0$, go back to step 1.
3. Output

$$\frac{\rho|\varphi\rangle}{\|\rho|\varphi\rangle\|}. \tag{4.13}$$

We will implement the algorithm in Definition 4.3 to sample from stabilizer Scrooge ensembles in our numerical experiments in Section 4.4. The next lemma shows that the alternative protocol indeed samples from the correct distribution.

Lemma 4.2. *On input ρ, the algorithm given in Definition 4.3 samples from the stabilizer Scrooge distribution defined by ρ.*

Proof. It suffices to show that under the conditions of Lemma 4.1, the algorithm given in Definition 4.3 also samples from the uniform distribution over the set $\{U(|0^k\rangle \otimes |\phi\rangle) : |\phi\rangle$ is an $(n-k)$-qubit stabilizer state$\}$.

First consider the case where $\rho = \frac{1}{2^{n-k}}(|0^k\rangle\langle 0^k| \otimes I)$. In this case, clearly, the algorithm only ever outputs states of the form $|0^k\rangle \otimes |\phi\rangle$ for some $(n-k)$-qubit stabilizer state $|\phi\rangle$. Moreover, the distribution is invariant under unitary transformations of the form $I \otimes V$ for every $(n-k)$-qubit Clifford unitary V. Therefore, the procedure samples from the uniform distribution over

$$\{|0^k\rangle \otimes |\phi\rangle : |\phi\rangle \text{ is an } (n-k)\text{-qubit stabilizer state}\}. \tag{4.14}$$

By the definition of ρ, we have $\{g_1, \ldots, g_k\} = \{Z_1, \ldots, Z_k\}$, so $U(|0^k\rangle\langle 0^k| \otimes I)U^\dagger = |0^k\rangle\langle 0^k| \otimes I$. We show that in this case, there exists an $(n-k)$-qubit Clifford unitary V such that for every $(n-k)$-qubit stabilizer state $|\phi\rangle$, $U(|0^k\rangle \otimes |\phi\rangle) = |0^k\rangle \otimes (V|\phi\rangle)$. Consider the $(n-k)$-qubit operator V defined by

$$V|\phi'\rangle = (\langle 0^k| \otimes I)U(|0^k\rangle \otimes |\phi'\rangle) \tag{4.15}$$

for every $(n-k)$-qubit state $|\phi'\rangle$. First observe that indeed

$$|0^k\rangle \otimes (V|\phi'\rangle) = (|0^k\rangle\langle 0^k| \otimes I)U(|0^k\rangle \otimes |\phi'\rangle) = U(|0^k\rangle \otimes |\phi'\rangle). \tag{4.16}$$

Then observe that

$$\||V|\phi'\rangle\| = \||0^k\rangle \otimes (V|\phi'\rangle)\| = \|U(|0^k\rangle \otimes |\phi'\rangle)\| = \||0^k\rangle \otimes |\phi'\rangle\| = 1, \tag{4.17}$$

so V is a unitary. For every $(n-k)$-qubit stabilizer state $|\phi\rangle$, $V|\phi\rangle$ is again a stabilizer state, so V is a Clifford unitary. The claim then follows since the sampled distribution is invariant under $|0^k\rangle \otimes |\phi\rangle \to U(|0^k\rangle \otimes |\phi\rangle)$.

Next consider an arbitrary n-qubit stabilizer state ρ with stabilizer group $\langle g_1, g_2 \ldots, g_k\rangle$. Recall that for a Clifford unitary U satisfying $\{UZ_jU^\dagger : j \in [k]\} = \{g_1, \ldots, g_k\}$, we can write

$$\rho = \frac{1}{2^{n-k}} U(|0^k\rangle\langle 0^k| \otimes I)U^\dagger. \tag{4.18}$$

Therefore, we can write

$$\rho|\varphi\rangle = \frac{1}{2^{n-k}} U(|0^k\rangle\langle 0^k| \otimes I)U^\dagger|\varphi\rangle. \tag{4.19}$$

By the unitary invariance property of $|\varphi\rangle \to U^\dagger|\varphi\rangle$, the algorithm in Definition 4.3 is equivalent to the following procedure.

1. Draw a uniformly random n-qubit stabilizer state $|\varphi\rangle$.
2. If $(|0^k\rangle\langle 0^k| \otimes I)|\varphi\rangle = 0$, go back to step 1.
3. Define
$$|\alpha\rangle = \frac{(|0^k\rangle\langle 0^k| \otimes I)|\varphi\rangle}{\|(|0^k\rangle\langle 0^k| \otimes I)|\varphi\rangle\|}. \tag{4.20}$$
4. Output
$$U|\alpha\rangle. \tag{4.21}$$

Notice that by the first case, $|\alpha\rangle$ is a uniformly random state drawn from the set

$$\{|0^k\rangle \otimes |\phi\rangle : |\phi\rangle \text{ is an } (n-k)\text{-qubit stabilizer state}\}. \tag{4.22}$$

Therefore, $U|\alpha\rangle$ is a uniformly random state drawn from the set

$$\{U(|0^k\rangle \otimes |\phi\rangle) : |\phi\rangle \text{ is an } (n-k)\text{-qubit stabilizer state}\}. \tag{4.23}$$

□

Next, we establish some properties about an illuminating special case of stabilizer Scrooge distributions where $\rho = \rho_A \otimes \rho_B$ is a product state. In particular, we will focus on understanding the distribution of the number of local and non-local Z-type stabilizers in a Scrooge random state defined by product-state ρ. We first look at local Z-type stabilizers. In the following, let $[n] = A \cup B$ be a bipartition of the n qubits with $|A| = n_A$ and $|B| = n_B$ so that $n_A + n_B = n$. Let ρ_A be an n_A-qubit mixed stabilizer state with 2^{k_A} stabilizers, and let ρ_B be an n_B-qubit mixed stabilizer state with 2^{k_B} stabilizers. Define $\rho = \rho_A \otimes \rho_B$ and let $\mathcal{S}_\rho$ denote the stabilizer Scrooge distribution defined by ρ. Let $S^Z_{\rho_A}$ denote the Z-type stabilizer subgroup of ρ local on A, which is in bijection with the Z-type stabilizer subgroup of ρ_A. We use $|\psi\rangle$ to denote a Scrooge random state drawn from $\mathcal{S}_\rho$, and we use $S^Z_{\psi_A}$ to denote the Z-type stabilizer subgroup of $|\psi\rangle$ local on A. In particular, this consists of all Z-type stabilizers of $|\psi\rangle$ that act nontrivially only on A and as the identity on B. We first give a lower bound on the probability that $S^Z_{\psi_A}$ remains unchanged from $S^Z_{\rho_A}$.

Recall that for every integer $m \geq 1$, we have defined $\mathcal{P}^m = \{I, X, Y, Z\}^{\otimes m}$, and for every set S consisting of Pauli operators, we have defined the notation $\pm S = S \cup \{-P : P \in S\}$.

Lemma 4.3. *It holds that*

$$\Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) \geq 1 - \frac{1}{2^{n_B - k_B}}. \tag{4.24}$$

Proof. By applying Lemma 4.1 to $\rho_A \otimes \rho_B$, we can choose n_A- and n_B-qubit Clifford unitaries U and V such that $\mathcal{S}_\rho$ is the uniform distribution over the set

$$\{(U \otimes V)(|0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle) : |\phi\rangle \text{ is an } (n - k_A - k_B)\text{-qubit stabilizer state}\}. \tag{4.25}$$

Define

$$R^* = \pm U^\dagger \{I, Z\}^{\otimes n_A} U, \tag{4.26}$$

i.e., R^* is the set of Pauli operators P on A such that $UPU^\dagger$ is Z-type. Define

$$R' = \{Q \in \mathcal{P}^{n_A - k_A} : Z(s) \otimes Q \in R^* \text{ for some } s \in \{0,1\}^{k_A}\}. \tag{4.27}$$

Notice that R' is a commuting set of Pauli operators supported on $n_A - k_A$ qubits with $+$ sign. Therefore, by maximality, $|R'| \leq 2^{n_A - k_A}$. Note that $I^{\otimes n_A - k_A} \in R'$, and R' is determined entirely by U, which is independent of any randomness involving $|\phi\rangle$.

Recall that we have $|\psi\rangle \sim \mathcal{S}_\rho$, so $|\psi\rangle = (U \otimes V)(|0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle)$ where $|\phi\rangle$ is a uniformly random $(n - k_A - k_B)$-qubit stabilizer state. Let S_ϕ denote the stabilizer group of $|\phi\rangle$. Define

$$C = \{Q \in R' : \pm Q \otimes I^{\otimes n_B - k_B} \in S_\phi\}. \tag{4.28}$$

For every $Q \in C$, fix a representative $s_Q \in \{0,1\}^{k_A}$ that satisfies $Z(s_Q) \otimes Q \in R^*$, and let $c_Q \in \{-1, 1\}$ be the sign that makes $c_Q \cdot Q \otimes I^{\otimes n_B - k_B} \in S_\phi$. Define

$$C' = \{c_Q \cdot U(Z(s_Q) \otimes Q)U^\dagger : Q \in C\}. \tag{4.29}$$

Clearly, $|C'| = |C|$. Notice that by construction, $C' \subseteq S^Z_{\psi_A}$.

We next show that $S^Z_{\psi_A} \subseteq S^Z_{\rho_A} \times C'$, which implies $|S^Z_{\psi_A}| \leq |S^Z_{\rho_A}| \cdot |C|$. Let $P \in S^Z_{\psi_A}$. Then by the definition of $S^Z_{\psi_A}$, $P = \pm U(Z(s) \otimes Q)U^\dagger$ for some $s \in \{0,1\}^{k_A}$ and $Q \in \mathcal{P}^{n_A - k_A}$

satisfying $Z(s) \otimes Q \in R^*$ and $\pm Q \otimes I^{\otimes n_B - k_B} \in S_\phi$. Observe that the condition on Q implies $Q \in C$, so $P' = c_Q \cdot U(Z(s_Q) \otimes Q)U^\dagger \in C'$. Define $Z(r) = Z(s)Z(s_Q)$. Then $P \in S^Z_{\psi_A}$ and $P' \in S^Z_{\psi_A}$ imply

$$PP' = U(Z(r) \otimes I^{\otimes n_A - k_A})U^\dagger \in S^Z_{\psi_A}. \tag{4.30}$$

Notice that by the definition of $S^Z_{\rho_A}$, $PP' \in S^Z_{\rho_A}$. Therefore, $P = (PP')P'$ with $PP' \in S^Z_{\rho_A}$ and $P' \in C'$.

Observe that

$$\mathbb{E}[|C|] = \sum_{Q \in R'} \Pr(\pm Q \otimes I^{\otimes n_B - k_B} \in S_\phi) \leq 1 + \frac{2^{n_A - k_A} - 1}{2^{n - k_A - k_B} + 1} \leq 1 + \frac{1}{2^{n_B - k_B}}. \tag{4.31}$$

Therefore,

$$\mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|} \right] \leq \mathbb{E}[|C|] \leq 1 + \frac{1}{2^{n_B - k_B}}. \tag{4.32}$$

By the law of total expectation,

$$\mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|} \right] = 1 \cdot \Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) + \mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|} \middle| S^Z_{\psi_A} \neq S^Z_{\rho_A} \right] \left(1 - \Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) \right) \tag{4.33}$$

$$\geq \Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) + 2 \left(1 - \Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) \right). \tag{4.34}$$

Therefore,

$$\Pr_{|\psi\rangle \sim \mathcal{S}_\rho}(S^Z_{\psi_A} = S^Z_{\rho_A}) \geq 2 - \mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|} \right] \geq 1 - \frac{1}{2^{n_B - k_B}}. \tag{4.35}$$

□

Lemma 4.3 implies that as long as $n_B - k_B \geq \Omega(\log n)$, then with probability at least $1 - 1/\mathrm{poly}(n)$, the Z-type stabilizer subgroup of $|\psi\rangle$ local on A will be the same as that of ρ; in other words, applying Scrooge to $\rho_A \otimes \rho_B$ does not introduce to $|\psi\rangle$ new Z-type stabilizers local on A with high probability. Since Z-type stabilizers dictate the output distribution of $|\psi\rangle$ when all qubits are measured in the computational basis, $S^Z_{\psi_A} = S^Z_{\rho_A}$ also implies that the marginal output distribution of $|\psi\rangle$ on A will be the same as the output distribution of ρ_A. Also note that since ρ_B is a mixed stabilizer state, $n_B - k_B$ equals the order-α Renyi entropy of ρ_B for every α.

In the following, we extend Lemma 4.3 to the more general case where $|\psi\rangle$ is allowed to be drawn from not necessarily the exact stabilizer Scrooge distribution defined by $\rho_A \otimes \rho_B$ but instead from a relative error ε-approximate Scrooge 2-design. We first define relative error ε-approximate Scrooge 2-designs.

Definition 4.4. Let $\varepsilon \geq 0$. Let ρ be an n-qubit mixed stabilizer state, and let $\mathcal{S}_\rho$ denote the stabilizer Scrooge distribution defined by ρ. We say a distribution $\mathcal{D}'_\rho$ of n-qubit pure states forms a relative error ε-approximate Scrooge 2-design with respect to $\mathcal{S}_\rho$ if for every 2^{2n}-by-2^{2n} Hermitian positive-semidefinite matrix O, it holds that

$$(1 - \varepsilon) \mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} [\langle\psi|^{\otimes 2} O |\psi\rangle^{\otimes 2}] \leq \mathop{\mathbb{E}}_{|\psi'\rangle \sim \mathcal{D}'_\rho} [\langle\psi'|^{\otimes 2} O |\psi'\rangle^{\otimes 2}] \leq (1 + \varepsilon) \mathop{\mathbb{E}}_{|\psi\rangle \sim \mathcal{S}_\rho} [\langle\psi|^{\otimes 2} O |\psi\rangle^{\otimes 2}]. \tag{4.36}$$

Contrasting Definition 4.4 with the statement in Conjecture 4.1, notice that in Definition 4.4, the approximation guarantee (4.36) needs to hold simultaneously for all Hermitian positive semidefinite observables O, even those chosen adversarially as a function of ρ. The next corollary assumes a condition weaker than the one stated in Definition 4.4; it only requires $|\psi\rangle$ to be drawn from a distribution $\mathcal{D}'_\rho$ that satisfies (4.36) for one specific observable.

Corollary 4.1. *Let $\mathcal{D}'_\rho$ be any distribution of n-qubit pure states that satisfies (4.36) with the specific observable*

$$O = 2^{n_A} \sum_{x\in\{0,1\}^{n_A}} |x\rangle\langle x| \otimes I^{\otimes n_B} \otimes |x\rangle\langle x| \otimes I^{\otimes n_B}, \tag{4.37}$$

and where $\rho = \rho_A \otimes \rho_B$ is a product state. For a state $|\psi'\rangle \sim \mathcal{D}'_\rho$, let $S^Z_{\psi'_A}$ denote the Z-type stabilizer subgroup of $|\psi'\rangle$ local on A. Then it holds that

$$\Pr_{|\psi'\rangle\sim\mathcal{D}'_\rho}(S^Z_{\psi'_A} = S^Z_{\rho_A}) \geq 1 - \varepsilon - \frac{1+\varepsilon}{2^{n_B-k_B}}. \tag{4.38}$$

Proof. In the proof of Lemma 4.3, we have established that for $|\psi\rangle$ drawn from the stabilizer Scrooge distribution defined by $\rho_A \otimes \rho_B$,

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|}\right] \leq 1 + \frac{1}{2^{n_B-k_B}}. \tag{4.39}$$

Through standard manipulations, we can write

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} \left[|S^Z_{\psi_A}|\right] = \mathrm{Tr}\left(O \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} \left[|\psi\rangle\langle\psi|^{\otimes 2}\right]\right), \tag{4.40}$$

where

$$O = 2^{n_A} \sum_{x\in\{0,1\}^{n_A}} |x\rangle\langle x| \otimes I^{\otimes n_B} \otimes |x\rangle\langle x| \otimes I^{\otimes n_B}. \tag{4.41}$$

Using (4.36) with the observable O gives

$$\mathop{\mathbb{E}}_{|\psi'\rangle\sim\mathcal{D}'_\rho} \left[\frac{|S^Z_{\psi'_A}|}{|S^Z_{\rho_A}|}\right] \leq (1+\varepsilon) \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} \left[\frac{|S^Z_{\psi_A}|}{|S^Z_{\rho_A}|}\right] \leq (1+\varepsilon)\left(1 + \frac{1}{2^{n_B-k_B}}\right). \tag{4.42}$$

Therefore,

$$\Pr_{|\psi'\rangle\sim\mathcal{D}'_\rho}(S^Z_{\psi'_A} = S^Z_{\rho_A}) \geq 2 - \mathop{\mathbb{E}}_{|\psi'\rangle\sim\mathcal{D}'_\rho} \left[\frac{|S^Z_{\psi'_A}|}{|S^Z_{\rho_A}|}\right] \geq 1 - \varepsilon - \frac{1+\varepsilon}{2^{n_B-k_B}}. \tag{4.43}$$

□

Note that just like in Lemma 4.3, if we have $n_B - k_B \geq \Omega(\log n)$ and $\varepsilon \leq 1/\mathrm{poly}(n)$ in Corollary 4.1, then we obtain $S^Z_{\psi'_A} = S^Z_{\rho_A}$ with probability $1 - 1/\mathrm{poly}(n)$.

Lastly, we work out the distribution of the mutual information $I(A;B)$ for the output distribution of $|\psi\rangle \sim \mathcal{S}_\rho$ where $\rho = \rho_A \otimes \rho_B$. Note that $I(A;B)$ is a random variable with

randomness induced by $|\psi\rangle \sim \mathcal{S}_\rho$. Analogous to $S^Z_{\psi_A}$, let $S^Z_{\psi_B}$ denote the Z-type stabilizer subgroup of $|\psi\rangle$ local on B. We also let S^Z_ψ denote the whole Z-type stabilizer subgroup of $|\psi\rangle$. We say $P \in S^Z_\psi$ is non-local if P cannot be decomposed as a product $P = P_A P_B$ for some $P_A \in S^Z_{\psi_A}$ and $P_B \in S^Z_{\psi_B}$. To find a non-local Z-type stabilizer of $|\psi\rangle$, we can first find a set of independent generators G^Z_A and G^Z_B for $S^Z_{\psi_A}$ and $S^Z_{\psi_B}$ respectively, and then try to extend $G^Z_A \cup G^Z_B$ to a full set of independent generators G^Z for S^Z_ψ. With G^Z constructed this way, $|\psi\rangle$ has a non-local Z-type stabilizer if and only if $G^Z \neq G^Z_A \cup G^Z_B$ and every $P \in G^Z \setminus (G^Z_A \cup G^Z_B)$ is a non-local Z-type stabilizer of $|\psi\rangle$ (recall that G^Z is a set of generators so $I \notin G^Z$). More generally, for a non-local Z-type stabilizer P of $|\psi\rangle$, the product PQ remains non-local for every $Q \in S^Z_{\psi_A} \cup S^Z_{\psi_B}$, i.e., non-local times local is non-local. We wish to characterize the distribution of the mutual information

$$I(A;B) = H(A) + H(B) - H(AB) = (n_A - |G^Z_A|) + (n_B - |G^Z_B|) - (n - |G^Z|) \tag{4.44}$$
$$= |G^Z| - |G^Z_A| - |G^Z_B| \tag{4.45}$$

which can also be interpreted as the number of independent non-local Z-type stabilizer generators of $|\psi\rangle$. Recall that we have $\rho = \rho_A \otimes \rho_B$ where ρ_A and ρ_B have 2^{k_A} and 2^{k_B} stabilizers respectively. Let $\mathcal{D}$ denote the uniform distribution over the set of all $(n-k_A-k_B)$-qubit stabilizer states. We will view each $|\phi\rangle \sim \mathcal{D}$ as a bipartite state between the first $n_A - k_A$ and last $n_B - k_B$ qubits and consider the mutual information of the computational-basis output distribution of $|\phi\rangle$.

Theorem 4.1. *It holds that for every integer $r \geq 0$,*

$$\Pr_{|\psi\rangle\sim\mathcal{S}_\rho}[I(A;B) = r] = \Pr_{|\phi\rangle\sim\mathcal{D}}[I(A;B) = r]\,. \tag{4.46}$$

Proof. By applying Lemma 4.1 to $\rho_A \otimes \rho_B$, we can choose n_A- and n_B-qubit Clifford unitaries U and V such that $\mathcal{S}_\rho$ is the uniform distribution over the set

$$\{(U \otimes V)(|0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle) : |\phi\rangle \text{ is an } (n - k_A - k_B)\text{-qubit stabilizer state}\}. \tag{4.47}$$

Let $|\psi\rangle = (U \otimes V)(|0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle) \sim \mathcal{S}_\rho$. Let A_1 denote the first k_A qubits of A, and let A_2 denote the last $n_A - k_A$ qubits of A. Similarly, let B_1 denote the first k_B qubits of B, and let B_2 denote the last $n_B - k_B$ qubits of B. Note that under this further subdivision, the support of $|\phi\rangle$ is precisely $A_2 \cup B_2$.

Define

$$R = \pm(U^\dagger \otimes V^\dagger)\{I, Z\}^{\otimes n}(U \otimes V) = \{\pm(U^\dagger \otimes V^\dagger)Z(s)(U \otimes V) : s \in \{0,1\}^n\} \tag{4.48}$$
$$= \pm(U^\dagger\{I, Z\}^{\otimes n_A}U) \otimes (V^\dagger\{I, Z\}^{\otimes n_B}V). \tag{4.49}$$

By construction, $(U \otimes V)R(U^\dagger \otimes V^\dagger)$ is the set of all n-qubit Z-type Pauli operators with $\pm$ signs. Note that R is a maximal set of n-qubit commuting Pauli operators with $\pm$ signs, and for every Pauli operator $P \in R$, we call P an R-type Pauli operator.

Let us also define

$$T = \{P \in \pm\mathcal{P}^{n-k_A-k_B} : Z(w) \otimes P \in R \text{ for some } w \in \{0,1\}^{k_A+k_B}\}. \tag{4.50}$$

It is also the case that T is a maximal set of $(n-k_A-k_B)$-qubit commuting Pauli operators with $\pm$ signs, i.e.,

$$|T| = 2 \cdot 2^{n-k_A-k_B}. \tag{4.51}$$

To establish (4.51), consider starting from the n-qubit state $(U^\dagger \otimes V^\dagger)|0^n\rangle$, which has a stabilizer group $S' \subseteq R$, and then measuring all qubits in $A_1 \cup B_1$ in the computational basis. The resulting state (conditioned on a particular measurement outcome) is an $(n-k_A-k_B)$-qubit pure state with a stabilizer group S'' such that $\pm S'' = T$. Since the post-measurement state is pure, we have $|S''| = 2^{n-k_A-k_B} = |T|/2$. Note that by (4.49), the set T factorizes across the (A_2, B_2)-bipartition.

We call every $P \in T$ a T-type Pauli operator. Note that the set T is a fixed set fully determined by $U \otimes V$.

Next, we show that the number of independent non-local R-type stabilizer generators of $|0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle$ equals the number of independent non-local T-type stabilizer generators of $|\phi\rangle$. More precisely, let

$$|\phi\rangle = |0^{k_A}\rangle \otimes |0^{k_B}\rangle \otimes |\phi\rangle \tag{4.52}$$

and define $S(\Phi)$ to be the stabilizer group of Φ. Let $S_A(\Phi), S_B(\Phi)$ be the subgroups of $S(\Phi)$ that are local on subsystems A, B respectively. Finally, let

$$S^R(\Phi) = S(\Phi) \cap R, \qquad S^R_A(\Phi) = S_A(\Phi) \cap R, \qquad S^R_B(\Phi) = S_B(\Phi) \cap R \tag{4.53}$$

be the corresponding R-type subgroups.

Similarly, define the group of stabilizers $S(\phi)$ of $|\phi\rangle$, its local subgroups $S_{A_2}(\phi), S_{B_2}(\phi)$, and their T-type subgroups:

$$S^T(\phi) = S(\phi) \cap T, \qquad S^T_{A_2}(\phi) = S_{A_2}(\phi) \cap T, \qquad S^T_{B_2}(\phi) = S_{B_2}(\phi) \cap T. \tag{4.54}$$

We are interested in the quotient groups

$$H(\Phi) \equiv S^R(\Phi)/(S^R_A(\Phi)S^R_B(\Phi)), \qquad H(\phi) \equiv S^T(\phi)/(S^T_{A_2}(\phi)S^T_{B_2}(\phi)). \tag{4.55}$$

The mutual information $I(A;B)$ of the output distribution of $|\psi\rangle$ is equal to $\log_2(|H(\Phi)|)$. Since $|\phi\rangle$ is a uniformly random stabilizer state, the distribution of $H(\phi)$ is invariant if we replace $|\phi\rangle \to W|\phi\rangle$ where $W = W_{A_2} \otimes W_{B_2}$ is an $(n-k_A-k_B)$-qubit Clifford unitary that maps the set of T-type Pauli operators to the set of Z-type Pauli operators. Therefore, the distribution of $\log_2(|H(\phi)|)$ is the same as the distribution of the mutual information $I(A_2;B_2)$ of the output distribution of $|\phi\rangle$. The following claim then completes the proof.

Claim 2. $|H(\Phi)| = |H(\phi)|$

Proof. Let us define a function $f : H(\Phi) \to H(\phi)$ as follows. Let $P \in \pm\mathcal{P}^n$ such that the equivalence class

$$[P] \in H(\Phi). \tag{4.56}$$

Then we can write

$$P = Z(s) \otimes Q \tag{4.57}$$

for some $s \in \{0,1\}^{k_A+k_B}$ and $Q \in \pm\mathcal{P}^{n-k_A-k_B}$. Note that since P is non-local, Q is non-local. We define

$$f([P]) = [Q] \in H(\phi). \tag{4.58}$$

Note that this function is well-defined; if we choose another representative P' such that $[P'] = [P] \in H(\Phi)$, then $P' = PL$ where $L \in S_A^R(\Phi)S_B^R(\Phi)$ is a local R-type Pauli operator, i.e.,

$$P' = (Z(s) \otimes Q)(Z(t) \otimes K)(Z(u) \otimes J) \tag{4.59}$$

for some bit strings $t, u \in \{0,1\}^{k_A+k_B}$ and Pauli operators $K, J \in \pm\mathcal{P}^{n-k_A-k_B}$ such that $\mathrm{Supp}(K) \subseteq A_2$ and $\mathrm{Supp}(K) \subseteq B_2$. From the above, we see that $K \in S_{A_2}^T(\phi)$ and $J \in S_{B_2}^T(\phi)$ and therefore $f([P]) = [Q] = [Q \cdot K \cdot J] = f([P'])$, establishing that f is well defined.

To complete the proof of the claim, we show that f is invertible and therefore a bijection. Suppose that $[D] \in H(\phi)$ for some $(n-k_A-k_B)$-qubit Pauli operator D. Then by the definition of T, we have

$$Z(r) \otimes D \in R \tag{4.60}$$

for some $r \in \{0,1\}^{k_A+k_B}$. Then the inverse function is

$$f^{-1}([D]) = [Z(r) \otimes D] \tag{4.61}$$

We can confirm that this function is well-defined by checking that (a) if we start with some other representative D' such that $[D'] = [D]$ in $H(\phi)$, then $f^{-1}([D]) = f^{-1}([D'])$, and (b) if there is another string $r' \in \{0,1\}^{k_A+k_B}$ such that $Z(r') \otimes D \in R$, then $[Z(r') \otimes D] = [Z(r) \otimes D]$ in $H(\Phi)$. Finally, we can confirm it is a valid inverse, i.e., $f^{-1}(f([P])) = [P]$ for all $[P] \in H(\Phi)$ and $f(f^{-1}([D])) = [D]$ for all $[D] \in H(\phi)$. □

□

Notice that for a uniformly random $(n-k_A-k_B)$-qubit stabilizer state $|\phi\rangle \sim \mathcal{D}$, when viewed as a bipartite state between the first $n_A - k_A$ and the last $n_B - k_B$ qubits, the probabilities for $|\phi\rangle$ to have a non-identity Z-type stabilizer local on A or B can be upper bounded by $1/2^{n_B-k_B}$ or $1/2^{n_A-k_A}$ respectively. Thus, when both $n_A - k_A$ and $n_B - k_B$ are moderately large, it becomes overwhelmingly unlikely to draw a $|\phi\rangle$ that has a local Z-type stabilizer. Therefore, when both $n_A - k_A$ and $n_B - k_B$ are large, $I(A;B)$ effectively becomes the number of independent Z-type stabilizer generators of $|\phi\rangle$. Note that the same is true for $I(A;C|B)$ for a tripartition $[m] = A \cup B \cup C$ where $|A| = n_A - k_A$, $|C| = n_B - k_B$, and $|B| = m - |A| - |C|$ since measuring all qubits in B of an m-qubit uniformly random stabilizer state in the computational basis produces a post-measurement state that is a uniformly random $(n-k_A-k_B)$-qubit stabilizer state on $A \cup C$, i.e., a state $|\phi\rangle \sim \mathcal{D}$.

A closed-form formula for the number of independent Z-type stabilizer generators of a uniformly random stabilizer state is known.

Lemma 4.4 ([89])**.** *For every integer $r \in \{0,1,\ldots,n\}$, the probability that a uniformly random n-qubit stabilizer state has r independent Z-type stabilizer generators is*

$$\frac{\eta(r)}{\sum_{k=0}^{n}\eta(k)} \tag{4.62}$$

where $\eta(0) = 1$ *and*

$$\eta(r) = 2^{-r(r+1)/2} \prod_{a=1}^{r} \frac{1 - 2^{r-n-a}}{1 - 2^{-a}}. \tag{4.63}$$

As n increases, the probabilities in Lemma 4.4 quickly converge to the following values:

$$P(0) = P(1) = \prod_{k=0}^{\infty} \left(1 - \frac{1}{2^{2k+1}}\right) \approx 0.41942244, \tag{4.64}$$

$$P(2) = \frac{1}{3} P(1) \approx 0.13980748, \tag{4.65}$$

$$P(3) = \frac{1}{7} P(2) \approx 0.01997250, \tag{4.66}$$

$$P(4) = \frac{1}{15} P(3) \approx 0.00133150, \tag{4.67}$$

$$P(5) = \frac{1}{31} P(4) \approx 0.00004295. \tag{4.68}$$

Theorem 4.1 together with Lemma 4.4 predict that as long as $n_A - k_A$ and $n_B - k_B$ are both reasonably large, $\Pr(I(A;B) = 0) \approx 0.42$, $\Pr(I(A;B) = 1) \approx 0.42$, $\Pr(I(A;B) = 2) \approx 0.14$, and so on for $|\psi\rangle \sim \mathcal{S}_\rho$ with $\rho = \rho_A \otimes \rho_B$.

4.3 Structure of the output distributions of shallow 2D random Clifford circuits

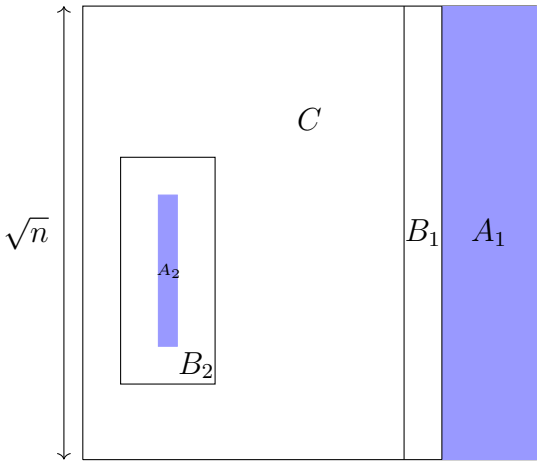


Figure 4.4: An example of an ABC tripartition in the set $\mathcal{T}$. In this example, the shaded A_1 and A_2 regions come from $A_1 \in \mathcal{K}_R$ and $A_2 \in \mathcal{H}_L$. We have $|B_1| = \sqrt{n} \times b \log_2 n$, $|A_2| = a \log_2 n \times 1$, and $|B_2| = O(\log n) \times O(\log n)$ for some constants a and b. Notice that A_1 and A_2 are always lightcone separated.

In this section, we show how we can use our Scrooge hypothesis Conjecture 4.1 to deduce detailed structural properties of the output distributions of shallow 2D random Clifford

circuits that are fully consistent with what we observed numerically in Section 4.1. As such, the Scrooge hypothesis provides an explanation for why we observed what we did in Section 4.1. The results in this section, together with Section 4.1, could be viewed as evidence for the Scrooge hypothesis in the stabilizer setting.

We consider depth-d 2D brickwork random Clifford circuits acting on a $\sqrt{n} \times \sqrt{n}$ grid of qubits with $d > d^*$, and we assume the Scrooge hypothesis holds. We will label the columns of the grid 1 through $\sqrt{n}$ from left to right. In the following, whenever we refer to a rectangular region of $y \times x$ qubits, y denotes the vertical side length, and x denotes the horizontal side length. Let L be the set of qubits in columns 1 through $\sqrt{n}/2$, and let R be the set of qubits in columns $\sqrt{n}/2+1$ through $\sqrt{n}$, so L and R partition the grid into left and right halves. Let $a \geq 1$ be a constant to be chosen later. Let $\mathcal{H}_L$ and $\mathcal{H}_R$ denote the set of all rectangular regions of $a \log_2 n \times 1$ qubits in L and R respectively, so each region is a thin rectangular strip of height $a \log_2 n$ and width 1. Note that $|\mathcal{H}_L| = |\mathcal{H}_R| \leq n/2$. Then, for every $j = 1, \ldots, \sqrt{n}/2$, let K_j denote the set of all qubits in columns 1 through j, and define $\mathcal{K}_L = \{K_j : j \in [\sqrt{n}/2]\}$. Similarly, for every $j = \sqrt{n}/2+1, \ldots, \sqrt{n}$, let K_j denote the set of all qubits in columns j through $\sqrt{n}$, and define $\mathcal{K}_R = \{K_j : j \in \{\sqrt{n}/2+1, \ldots, \sqrt{n}\}\}$. Note that $|\mathcal{K}_L| = |\mathcal{K}_R| = \sqrt{n}/2$. Next, we define a set of ABC tripartitions of the $\sqrt{n} \times \sqrt{n}$ grid, with each tripartition being specified by an element in the set

$$\mathcal{T} = \mathcal{T}_L \cup \mathcal{T}_R \tag{4.69}$$

where

$$\mathcal{T}_L = \{(A_1, A_2) \in \mathcal{K}_L \times \mathcal{H}_R : \text{dist}(A_1, A_2) \geq \sqrt{n}/3\} \tag{4.70}$$

and

$$\mathcal{T}_R = \{(A_1, A_2) \in \mathcal{K}_R \times \mathcal{H}_L : \text{dist}(A_1, A_2) \geq \sqrt{n}/3\}. \tag{4.71}$$

Let $b \geq 1$ be another constant to be chosen later. Each $(A_1, A_2) \in \mathcal{T}$ specifies an ABC tripartition where $A = A_1 \cup A_2$, and $B = B_1 \cup B_2$ is the disjoint union of two shielding regions separating A from C, the rest of the qubits, so that $\text{dist}(A, C) = b \log_2 n + 1$; see Figure 4.4 for an example with $(A_1, A_2) \in \mathcal{K}_R \times \mathcal{H}_L$. Note that for every ABC tripartition in $\mathcal{T}$, the C region contains a square region of qubits of size at least $\sqrt{n}/4 \times \sqrt{n}/4$, and that $|\mathcal{T}| \leq n^{1.5}$.

Let $\mathcal{U}$ denote the distribution of depth-d 2D brickwork random Clifford circuits with $d > d^*$. Next, we invoke the Scrooge hypothesis with $t = 2$ for a fixed $U \sim \mathcal{U}$ on every ABC tripartition in $\mathcal{T}$ twice in the following way. Let $c \geq 2$ be a constant. Let ABC with $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ be a tripartition specified by an element in $\mathcal{T}$. We choose the constant b to be large enough so that the Scrooge hypothesis holds for the ABC tripartition with $\varepsilon \leq 1/n^c$. If we were to apply the circuit, measure all qubits in B, and trace out all qubits in C, we would obtain a background state $\rho = \rho_{A_1} \otimes \rho_{A_2}$ since A_1 and A_2 are lightcone separated. Depending on the measurement outcome x_B, ρ belongs to an equivalence class $[\rho]$ with ρ being a representative where all $\rho' \in [\rho]$ have the same stabilizers as ρ modulo the $\pm$ signs. Next, we invoke Corollary 4.1 on $[\rho_{A_1} \otimes \rho_{A_2}]$ with the single observable, so the event in (4.38) occurs with probability $1 - \varepsilon - (1+\varepsilon)/2^{n_{A_2} - k_{A_2}}$ for all $\rho' \in [\rho_{A_1} \otimes \rho_{A_2}]$. To lower bound $n_{A_2} - k_{A_2}$, we choose the constant a to be large enough so that the Scrooge hypothesis holds for another tripartition $\bar{A} = A_2$, $\bar{B} = B_2$, and $\bar{C} = A_1 \cup B_1 \cup C$ with $\varepsilon \leq 1/n^c$ and

$S_\infty(\rho'_{A_2}) = n_{A_2} - k_{A_2} \geq c \log_2 n$ for all $\rho'_{A_2} \in [\rho_{A_2}]$. With this, we get

$$1 - \varepsilon - \frac{1+\varepsilon}{2^{n_{A_2}-k_{A_2}}} \geq 1 - \frac{3}{n^c}. \tag{4.72}$$

By a union bound, all of the above events occur simultaneously with probability at least $1 - 5n^{1.5}/n^c$, which is $1 - 1/\text{poly}(n)$. We call a $U \sim \mathcal{U}$ Good if all the above events occur simultaneously. We have established that we draw a Good circuit with probability at least $1 - 1/\text{poly(n)}$.

In the following, let $|\psi\rangle$ be the output state of a Good constant-depth 2D random Clifford circuit. Let P be a non-identity Z-type stabilizer of $|\psi\rangle$. Let j and k be the indices of the leftmost and rightmost $\sqrt{n} \times 1$ columns of qubits on which P has non-trivial support. We define the column width of P to be $cw(P) = k - j + 1$. We say P is column-wise decomposable if P can be written as a product

$$P = \prod_{j=1}^{r} Q_j \tag{4.73}$$

for some $r \geq 1$ and every Q_j is a Z-type stabilizer of $|\psi\rangle$ that has column width at most $b \log_2 n + 1$. Recall that for every $P = P_1 \otimes \cdots \otimes P_n$ and $R \subseteq [n]$, we defined the truncation of P to R to be the $|R|$-qubit Pauli operator $P_R = \bigotimes_{j \in R} P_j$. We say P has a hole $H \subseteq [n]$ if H has dimension $a \log_2(n) \times 1$ and P is all identity on H, or equivalently, $P_H = I^{\otimes |H|}$.

Theorem 4.2. *Let $|\psi\rangle$ be the output state of a Good constant-depth 2D random Clifford circuit U. For every non-identity Z-type stabilizer P of $|\psi\rangle$, if P has a hole, then P is column-wise decomposable. That is, P can be written as a product*

$$P = \prod_{j=1}^{r} Q_j \tag{4.74}$$

where $r \leq cw(P)$ and every Q_j is a Z-type stabilizer of $|\psi\rangle$ with column width at most $b \log_2 n + 1$.

Proof. Let S^Z be the set of all non-identity Z-type stabilizers of $|\psi\rangle$. We proceed by induction on the column width. The base case includes all $P \in S^Z$ that have $cw(P) \leq b \log_2 n + 1$ for which the claim holds trivially. For the inductive step, let $P \in S^Z$ such that P has a hole and $cw(P) > b \log_2 n + 1$. Let j and k be the indices of the leftmost and rightmost columns of qubits that P has non-trivial support on. Let H_k and H_j be any holes of P that have the largest column-wise distance to j and k respectively. Without loss of generality, assume H_j and k are the more distant pair. Note that the column-wise distance between H_j and k is necessarily at least $\sqrt{n}/2 - 2$. Define A_1 to be the union of the columns k through $\sqrt{n}$. Define $A_2 = H_j$. Let R denote the k-th column of qubits. First, notice that since P is a Z-type stabilizer of $|\psi\rangle$ and $P_{A_2} = I^{\otimes |A_2|}$, if we were to measure all qubits of $|\psi\rangle$ in $[n] \setminus (A_1 \cup A_2)$ in the computational basis to get a pure post-measurement state $|\psi'\rangle$, P_R or $-P_R$ will be a Z-type stabilizer of $|\psi'\rangle$ that is local on A_1. Our choice of (A_1, A_2) corresponds to an ABC tripartition contained in $\mathcal{T}$ with a corresponding region $B = B_1 \cup B_2$. Let W denote the backward lightcone circuit of $A_1 \cup B_1$, and let $[\rho_{A_1} \otimes \rho_{A_2}]$ denote the equivalence class of

background states sampled when invoking the Scrooge hypothesis on this ABC tripartition. Since U is Good, P_R or $-P_R$ being a Z-type stabilizer of $|\psi'\rangle$ local on A_1 implies that P_R or $-P_R$ is a Z-type stabilizer of ρ_{A_1}. By lightcone cancellation, this implies that $(W \otimes I)|0^n\rangle$ has a Z-type stabilizer Q supported on $R \cup B_1$ such that $Q_R = P_R$, and Q is also a Z-type stabilizer of $|\psi\rangle$. Define $Q' = PQ$. Then we have $P = QQ'$ satisfying $cw(Q) \leq b \log_2 n + 1$, $Q' \in S^Z$, Q' having a hole, and $cw(Q') \leq cw(P) - 1$. Notice that $cw(P) > b \log_2 n + 1$ implies $P \neq Q$, which implies $Q' \neq I^{\otimes n}$. Thus, the claim follows by applying the inductive hypothesis to Q'. □

As a sanity check, note that the converse of Theorem 4.2 is not true. Clearly, P could still be column-wise decomposable even if P does not have a hole. The contrapositive of Theorem 4.2 implies that every Z-type stabilizer of $|\psi\rangle$ that is not column-wise decomposable cannot have any holes, so the weight of a column-wise indecomposable Z-type stabilizer needs to be at least $\frac{n}{a \log_2(n)}$.

Corollary 4.2. *Let $|\psi\rangle$ be the output state of a Good constant-depth 2D random Clifford circuit. The Z-type stabilizer subgroup of $|\psi\rangle$ admits a set of independent generators G such that for every $P \in G$, either $cw(P) \leq b \log_2 n + 1$ or P is not column-wise decomposable and has no hole.*

Proof. Let S^Z be the set of all non-identity Z-type stabilizers of $|\psi\rangle$. Consider constructing a set of independent generators for S^Z using the following procedure. We sort every element in S^Z in increasing order of column width with arbitrary tie breaking, so that $S^Z = \{P_1, P_2, \ldots, P_{|S^Z|}\}$ with $cw(P_1) \leq cw(P_2) \leq \cdots \leq cw(P_{|S^Z|})$. Define $G_0 = \emptyset$. We then scan through every element in S^Z. For every $j \in [|S^Z|]$, for the j-th iteration, we set $G_j = G_{j-1}$ if $P_j \in \langle G_{j-1} \rangle$ and $G_j = G_{j-1} \cup \{P_j\}$ otherwise. By construction, $G_{|S^Z|}$ is a set of independent generators for S^Z. Let k be the cutoff such that $cw(P_j) \leq b \log_2 n + 1$ if and only if $j \leq k$. Suppose for some $j > k$, P_j is added to G_{j-1} in the j-th iteration. This implies $P_j \notin \langle G_{j-1} \rangle$, which implies P_j is not column-wise decomposable. If P_j has a hole, by Theorem 4.2, P_j would be column-wise decomposable. Therefore, P_j cannot have a hole. □

At this point, we would like to apply the Scrooge hypothesis three more times with A being the leftmost column, the rightmost column, and the union of the leftmost and rightmost columns together with Theorem 4.1 to argue that the distribution of the number of column-wise indecomposable Z-type stabilizer generators converges to the distribution characterized in Lemma 4.4 as $n \to \infty$. Unfortunately, we do not know how to prove a generalization of Theorem 4.1 for ε-approximate Scrooge t-designs, and we do not get exact Scrooge from the Scrooge hypothesis. Intuitively speaking, since the distribution in Lemma 4.4 effectively has a sample space of size 6, and every discrete probability distribution over a sample space of size t is completely characterized by its first $t - 1$ moments, a generalization requiring only the guarantees of an ε-approximate Scrooge t-design seems plausible.

If we trace out from $|\psi\rangle$ any subset of qubits $H \subseteq [n]$ that contains a hole of dimension $a \log_2(n) \times 1$, then every Z-type stabilizer of $|\psi\rangle$ that does not have any hole is deleted, which includes every Z-type stabilizer of $|\psi\rangle$ that is not column-wise decomposable. Therefore, every Z-type stabilizer of $\mathrm{Tr}_H(|\psi\rangle\langle\psi|)$ is column-wise decomposable in terms of the Z-type stabilizers of $|\psi\rangle$. However, it may not be generically true that a Z-type stabilizer of

$\mathrm{Tr}_H(|\psi\rangle\langle\psi|)$ is column-wise decomposable using only the Z-type stabilizers of $\mathrm{Tr}_H(|\psi\rangle\langle\psi|)$. To illustrate this subtlety, consider a 3-qubit example where $S_\psi^Z = \langle ZZI, IZZ\rangle$. We have $ZIZ \in S_\psi^Z$ and ZIZ is decomposable as $ZIZ = (ZZI)(IZZ)$. However, if we trace out the middle qubit, ZIZ is the sole non-identity element in S_ψ^Z that is not deleted, so it becomes indecomposable. Nevertheless, we show below that in our case, all the Z-type stabilizers of the reduced state are column-wise decomposable in terms of the Z-type stabilizers of the reduced state in a slightly relaxed sense.

Corollary 4.3. *Let $|\psi\rangle$ be the output state of a Good constant-depth 2D random Clifford circuit. Let $H \subseteq [n]$ be a hole of dimension $a\log_2 n \times 1$ and define $\rho = \mathrm{Tr}_H(|\psi\rangle\langle\psi|)$. It holds that every non-identity Z-type stabilizer P' of ρ can be decomposed as*

$$P' = \prod_{j=1}^{r'} Q'_j \tag{4.75}$$

where $r' \le cw(P')$ and each Q'_j is a Z-type stabilizer of ρ. Moreover, every $j \in [r']$ has $cw(Q'_j) \le b\log_2(n) + 1$, except for at most a single $j^* \in [r']$ such that H is supported on a column between the leftmost and rightmost columns that Q'_{j^*} is supported on and $cw(Q'_{j^*}) \le 2b\log_2(n) + 1$.

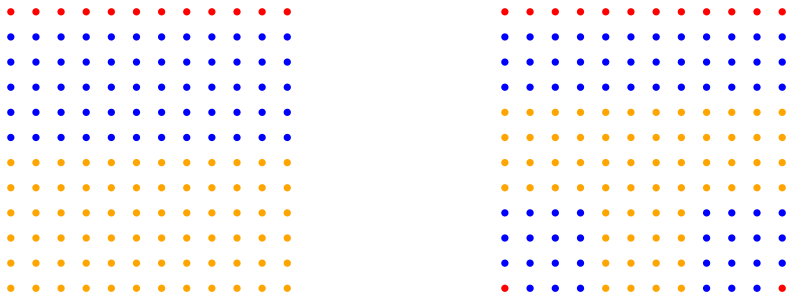
Proof. Let P' be a non-identity Z-type stabilizer of ρ. Then $P = P' \otimes I^{\otimes|H|}$ is a Z-type stabilizer of $|\psi\rangle$ with a hole. By Theorem 4.2, let $P = \prod_{j=1}^{r} Q_j$ be a column-wise decomposition of P. Note that there could be $j \in [r]$ such that Q_j is supported non-trivially on H. Define $B = \{j \in [r] : (Q_j)_H \neq I^{\otimes|H|}\}$ and consider $S = \prod_{j\in B} Q_j$ and $T = \prod_{j\in[r]\setminus B} Q_j$. Since $P = ST$ and $P_H = T_H = I^{\otimes|H|}$, we have $S_H = I^{\otimes|H|}$. Since $cw(Q_j) \le b\log_2 n + 1$ for every $j \in B$, we deduce that $cw(S) \le 2b\log_2 n + 1$. Therefore,

$$P' = \mathrm{Tr}_H(S) \prod_{j\in[r]\setminus B} \mathrm{Tr}_H(Q_j) \tag{4.76}$$

is a decomposition of P' satisfying all the requirements. □

4.4 Numerical evidence for the Scrooge hypothesis

In this section, we numerically investigate the Scrooge hypothesis in the stabilizer setting for diagonal observables like (4.37). All the following numerical experiments share a common setup. We take an ABC tripartition of an $L_y \times L_x$ grid of qubits, and we define $\mathrm{width}(B) = \mathrm{dist}(A, C) - 1$ where $\mathrm{dist}(A, C) = \min_{x\in A, y\in C} \mathrm{dist}(x, y)$ denotes the minimum L1 distance between $x \in A$ and $y \in C$ on the 2D grid. The two types of ABC tripartitions we consider are depicted in Figure 4.5, where in both figures, each dot represents a qubit, and the red, blue, and orange qubits form the regions A, B, and C respectively. We write a depth-7 2D brickwork circuit U as $U = U_C U_{AB}$ where U_{AB} is a backward lightcone circuit acting non-trivially on $A \cup B$ and some boundary qubits $C' \subseteq C$, and U_C is a forward lightcone circuit acting non-trivially on C. To generate a random ρ sample, we first sample random Clifford gates for U_{AB}, then apply U_{AB} to $|0^n\rangle$, and finally measure all qubits in B in the



(a) Row-by-row

(b) Row-by-row with two corners

Figure 4.5: 12×12 examples of the ABC tripartitions considered in our numerical experiments. For both cases, the red, blue, and orange grid points depict the qubits in the regions A, B, and C respectively. For (a) the row-by-row tripartition, we always have $|A| = L_x$, and for (b) with the two corners, we always have $|A| = L_x + 2$. In the example (a), $\text{width}(B) = 5$ while in the example (b), $\text{width}(B) = 3$. We will always simulate a square grid ($L_y = L_x$) of qubits for the row-by-row with two corners tripartition, but we will consider the row-by-row tripartition also on rectangular grids with $L_y \neq L_x$.

computational basis to get a pure post-measurement state $|\varphi\rangle_{AC} = |\psi\rangle_{AC'} \otimes |0^{|C\setminus C'|}\rangle$ on $A \cup C$. With this, we obtain $\rho = \text{Tr}_C(|\varphi\rangle\langle\varphi|)$. Conditioned on each $|\varphi\rangle_{AC}$ and ρ generated in this way, we consider two distributions of pure stabilizer states supported on A. The first is $\mathcal{S}_\rho$, the stabilizer Scrooge distribution defined by ρ. The second, which we call $\mathcal{D}_\rho$, is the distribution sampled by the following procedure: first draw independent fresh random gates for U_C, then apply U_C to $|\varphi\rangle_{AC}$, and finally measure all qubits in C in the computational basis to get a pure post-measurement state on A. Note that strictly speaking, $\mathcal{D}_\rho$ depends on not just ρ but also the specific purification $|\psi\rangle_{AC'}$ of ρ.

For constant $t \geq 1$ and diagonal Hermitian positive semidefinite t-th moment observable O, the Scrooge hypothesis concerns itself with the scaling of the relative error

$$\left| \frac{\mathbb{E}_{|\psi\rangle\sim\mathcal{D}_\rho}\left[\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}\right] - \mathbb{E}_{|\psi\rangle\sim\mathcal{S}_\rho}\left[\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}\right]}{\mathbb{E}_{|\psi\rangle\sim\mathcal{S}_\rho}\left[\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}\right]} \right| \tag{4.77}$$

as a function of $\text{width}(B)$ and $|\partial A|$ when the ABC tripartition satisfies some regularity conditions.

Let $\mathcal{M}$ denote the completely dephasing channel, or equivalently the computational-basis measurement channel, which sets all off diagonal entries of an input density matrix to zero. We call the expression

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}_\rho}\left[\mathcal{M}\left(|\psi\rangle\langle\psi|^{\otimes t}\right)\right] \tag{4.78}$$

the t-th diagonal moment operator of $\mathcal{D}_\rho$. Notice that

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}_\rho}\left[\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}\right] = \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho}\left[\langle\psi|^{\otimes t} O |\psi\rangle^{\otimes t}\right] \tag{4.79}$$

for every diagonal t-th moment observable O if and only if

$$\mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}_\rho}\left[\mathcal{M}\left(|\psi\rangle\langle\psi|^{\otimes t}\right)\right] = \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho}\left[\mathcal{M}\left(|\psi\rangle\langle\psi|^{\otimes t}\right)\right]. \tag{4.80}$$

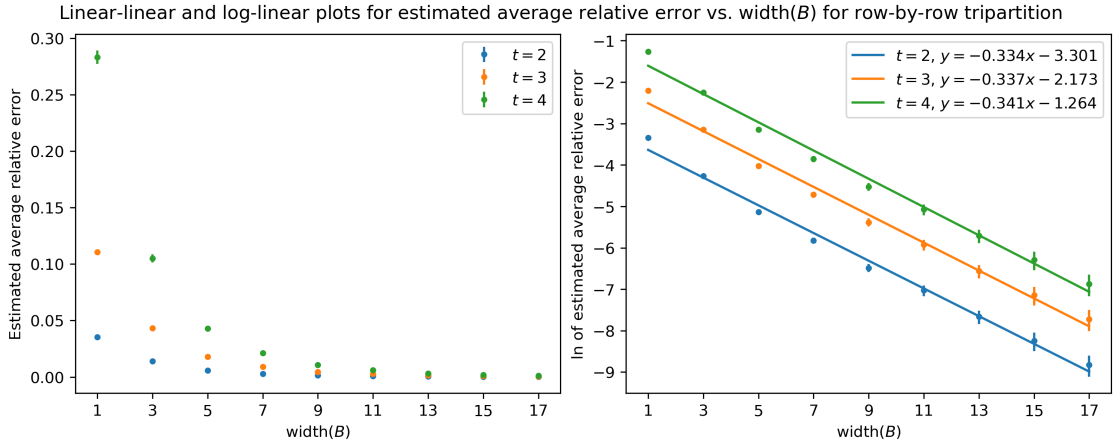


Figure 4.6: In this experiment, we investigate the dependence of the average relative error on width(B) for the row-by-row tripartition. The other architectural parameters are held constant at $L_x = 100$, $|A| = 100$, and $|C| = 600$. Each data point is estimated using 5000 independent ρ samples, and for each ρ, the relative error is computed from t-th diagonal frame potentials of $\mathcal{S}_\rho$ and $\mathcal{D}_\rho$ estimated using 10^4 independent state samples. From the right panel, we see that as width(B) increases from 1 to 17, the data points show good agreement with the linear least squares best fits, and that the y-axis spans a good range from, for example, e^{-3} to e^{-9} for $t = 2$. Therefore, the data predict that the average relative error scales inverse exponentially with respect to width(B). The error bars show $2\times$ standard error.

By mimicking the derivation of the state frame potential [90], we can quantify how well $\mathcal{D}_\rho$ approximates $\mathcal{S}_\rho$ on all diagonal observables by

$$\left\| \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{D}_\rho} \left[\mathcal{M}(|\psi\rangle\langle\psi|^{\otimes t})\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mathcal{S}_\rho} \left[\mathcal{M}(|\psi\rangle\langle\psi|^{\otimes t})\right] \right\|_F^2 \geq 0. \tag{4.81}$$

Expanding the LHS of the above expression, the first term we get is

$$\mathrm{Tr}\left(\mathop{\mathbb{E}}_{|\psi\rangle,|\phi\rangle\sim\mathcal{D}_\rho} \left[\mathcal{M}(|\psi\rangle\langle\psi|^{\otimes t})\mathcal{M}(|\phi\rangle\langle\phi|^{\otimes t})\right] \right) = \mathop{\mathbb{E}}_{|\psi\rangle,|\phi\rangle\sim\mathcal{D}_\rho} \left[\left(\sum_{x\in\{0,1\}^{|A|}} |\langle x|\psi\rangle|^2 |\langle x|\phi\rangle|^2 \right)^t \right]. \tag{4.82}$$

Analogous to the derivation of the state frame potential, we call (4.82) the t-th diagonal frame potential of $\mathcal{D}_\rho$. Note that if (4.82) is zero, then (4.80) holds. We adopt the relative error between the t-th diagonal frame potential of $\mathcal{D}_\rho$ and $\mathcal{S}_\rho$ as a proxy for quantifying the convergence of the t-th diagonal moment operator of $\mathcal{D}_\rho$ to that of $\mathcal{S}_\rho$.

After having sampled a list of $|A|$-qubit states $|\psi_1\rangle, \ldots, |\psi_T\rangle$ from $\mathcal{S}_\rho$ or $\mathcal{D}_\rho$, we can estimate the corresponding t-th diagonal frame potential using the empirical average

$$\frac{1}{\binom{T}{2}} \sum_{i=1}^{T} \sum_{j=i+1}^{T} \left(\sum_{x\in\{0,1\}^{|A|}} |\langle x|\psi_i\rangle|^2 |\langle x|\psi_j\rangle|^2 \right)^t. \tag{4.83}$$

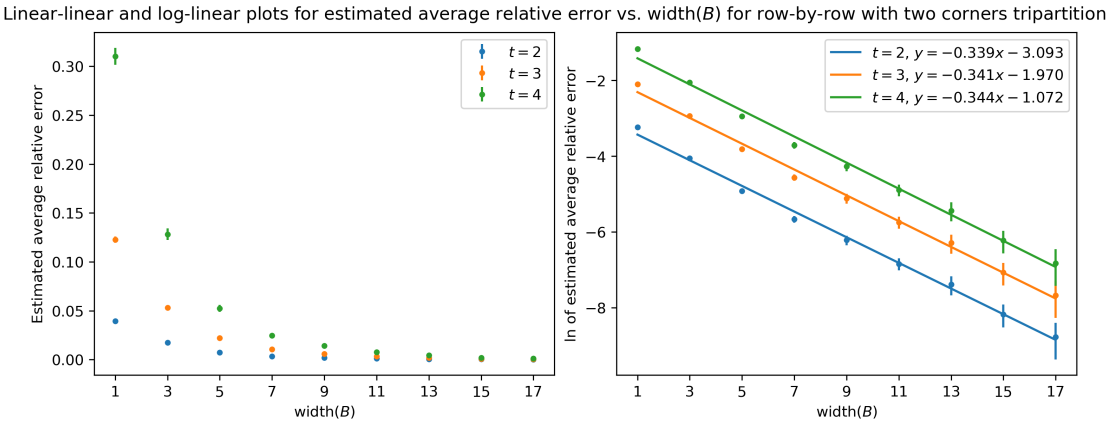


Figure 4.7: This experiment investigates the dependence of the average relative error on width(B) for the row-by-row with two corners tripartition. we fix $L_x = L_y = 60$ so $|A| = 62$. Note that as $|B|$ increases, $|C|$ decreases accordingly as the total number of qubits is always fixed to be 60×60. All the other experimental parameters are set identically to those used for Figure 4.6, and the same overall qualitative scaling behaviour is observed.

For a pair of stabilizer states $|\psi_i\rangle$ and $|\psi_j\rangle$, the inner-most sum satisfies

$$\sum_{x \in \{0,1\}^{|A|}} |\langle x|\psi_i\rangle|^2 \cdot |\langle x|\psi_j\rangle|^2 = \frac{|\mathrm{Supp}(|\psi_i\rangle) \cap \mathrm{Supp}(|\psi_j\rangle)|}{|\mathrm{Supp}(|\psi_i\rangle)| \cdot |\mathrm{Supp}(|\psi_j\rangle)|} \tag{4.84}$$

where the RHS can be computed using linear algebra in $O(|A|^3)$ time.

In our first experiment, we investigate the scaling of the average relative error between the t-th diagonal frame potentials of $\mathcal{D}_\rho$ and $\mathcal{S}_\rho$ as a function of width(B) for the two types of ABC tripartitions depicted in Figure 4.5. For the row-by-row tripartition, we fix $L_x = 100$ and A to be the top row of qubits, so $|A| = L_x = 100$. We let width(B) vary from 1 to 17 and fix C to contain 6 rows of qubits below B, so $|C| = 600$. To estimate each data point, we generate 5000 independent ρ (strictly speaking $|\varphi\rangle_{AC}$) samples using the procedure described earlier. Conditioned on each ρ, we draw 10^4 independent state samples from $\mathcal{S}_\rho$ and $\mathcal{D}_\rho$ to estimate the t-th diagonal frame potentials of $\mathcal{S}_\rho$ and $\mathcal{D}_\rho$ using (4.83) and (4.84). In Figure 4.6, we plot the estimated average relative error against width(B) for $t \in \{2, 3, 4\}$, as well as in log-linear scale with the best linear fits. The results shown in Figure 4.6 predict an inverse exponential relationship between the average relative error and width(B). Also note that we were not bottlenecked by the dimension of C. We repeat the same experiment for the row-by-row with two corners tripartition on a 60×60 grid, and the results are shown in Figure 4.7. Again, the data are consistent with a scaling that is inverse exponential in width(B).

In our second experiment, we probe the scaling of the average relative error between the t-th diagonal frame potentials of $\mathcal{D}_\rho$ and $\mathcal{S}_\rho$ as a function of $|\partial A|$. We focus on $t = 2$ and the row-by-row tripartition with which $|\partial A| = |A|$. For this experiment, we set width(B) $\in \{1, 3, 5\}$ and let $|A|$ range from 20 to 200. We still set C to always contain 6 rows of qubits below B, but note that $|C|$ becomes larger as $|A|$ increases. Again, we estimate each data point using 5000 independent ρ samples, and conditioned on each ρ, 10^4 independent state

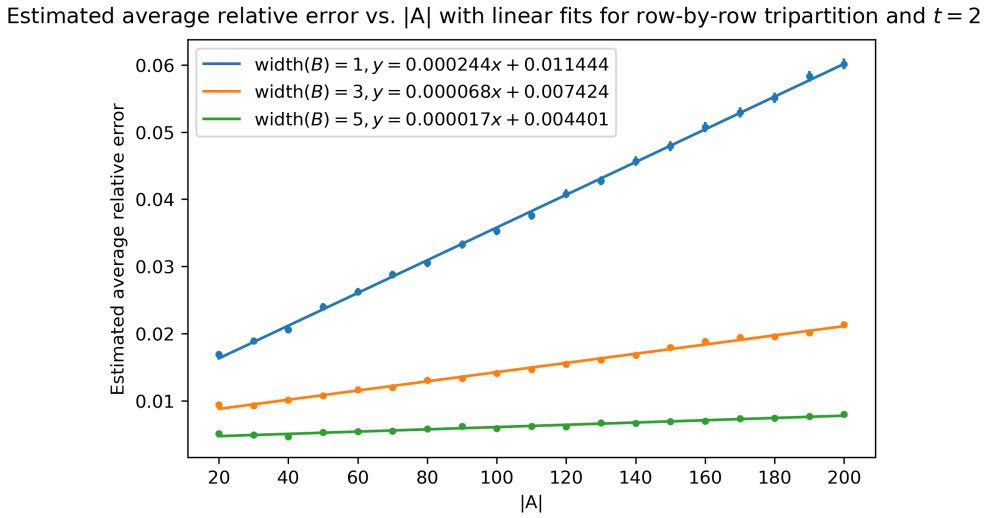


Figure 4.8: In this experiment, we probe the dependence of the average relative error on $|\partial A| = |A|$ for the row-by-row tripartition and $t = 2$. We fix C to contain 6 rows of qubits below B, but note that $|C|$ increases with $|A|$. In all three cases width$(B) \in \{1, 3, 5\}$, the data points are well-explained by a linear model with slope decreasing in width(B).

samples are used to estimate the t-th diagonal frame potentials of $\mathcal{S}_\rho$ and $\mathcal{D}_\rho$. The results are shown in Figure 4.8. As the values on the y-axis only spanned a narrow range, all we could conclude is that for each width$(B) \in \{1, 3, 5\}$, the estimated average relative error is well-modelled by the linear least squares fit. Nonetheless, note that as width(B) increases from 1 to 3 and from 3 to 5, the slopes of the best linear fits decay by a factor of 3.5 to 4, which is consistent with the inverse exponential in width(B) scaling predicted by the previous experiment shown in Figure 4.6.

In summary, all the scaling behaviours we observed numerically in Figures 4.6 to 4.8 show strong agreement with those conjectured by the Scrooge hypothesis.

4.5 Column-by-column simulation of shallow 2D random Clifford circuits

In Section 3.6, we showed how the presence of long-range tripartite MIE in the output states of shallow random Clifford circuits can be leveraged to achieve a quantum advantage over shallow classical circuits with bounded fan-in. At its core, GHZ-type MIE is able to thwart classical simulation because it encodes long-range correlations (see (3.103) and (3.109)) into any possible outcome drawn from the output distribution. A locality-constrained classical simulator is unable to perfectly capture this long-range correlation. In contrast, assuming the Scrooge hypothesis Conjecture 4.1, we also showed in Corollary 4.3 that with high probability, as soon as we trace out from the output state a hole of qubits of dimension $O(\log n) \times 1$, the remaining Z-type stabilizers of the reduced state all become column-wise decomposable. Therefore, with high probability, no long-range correlation persists in the output distribution of the reduced state.

In the paper [3], we show that assuming the Scrooge hypothesis, there is an efficient classical algorithm for simulating sampling from the output distributions of noisy shallow 2D brickwork random quantum circuits composed of Haar random gates, with independent single-qubit depolarizing noise at a rate of $\gamma = \Omega(\log n/n)$. To demonstrate the key idea of the algorithm in this section, we describe a simplified form of the algorithm for simulating noisy shallow 2D brickwork random Clifford circuits. For simplicity, the noise event is assumed to be depolarizing noise on a random hole of qubits of dimension $O(\log n) \times 1$ at the end of the circuit.

We assume $d > d^*$ and Conjecture 4.1 holds. Thus, Theorem 4.2 holds, and let a and b be the corresponding constants appearing in the statement of Theorem 4.2. The input-output behaviour of the algorithm is as follows. Suppose the algorithm takes as input a Good n-qubit depth-d 2D brickwork random Clifford circuit U with the n qubits living on a $\sqrt{n} \times \sqrt{n}$ grid. Let $|\psi\rangle = U|0^n\rangle$ be the output state of U. The algorithm picks a random hole H of dimension $a \log_2(n) \times 1$. By symmetry, assume that H lands on the left half of the grid. Define $\tilde{\rho} = \mathrm{Tr}_H(|\psi\rangle\langle\psi|)$. The output of the algorithm is a sample drawn according to the distribution $\tilde{q}(x) = \langle x|\tilde{\rho}|x\rangle$, $x \in \{0,1\}^{n-|H|}$ with $|H|$ independent uniformly random bits inserted at positions $i \in H$.

To describe the algorithm, we first establish some notations. Let $q(x) = |\langle x|\psi\rangle|^2$, $x \in \{0,1\}^n$ denote the output distribution of the full output state $|\psi\rangle$. For every $A \subseteq [n]$, let q_A denote the marginal distribution of q on A. For every $A, B \subseteq [n]$ with $A \cap B = \emptyset$ and $y \in \{0,1\}^{|B|}$, let $q_{A|B}(\cdot|y)$ denote the conditional marginal distribution of q on A conditioned on having sampled y on B, so for every $x \in \{0,1\}^{|A|}$, we have

$$q_{A|B}(x|y) = \frac{q_{A\cup B}(xy)}{q_B(y)}. \tag{4.85}$$

We use similar $\tilde{q}_A$ and $\tilde{q}_{A|B}$ notations to denote marginal and conditional marginal distributions of $\tilde{q}$. We label the columns of the grid 1 through $\sqrt{n}$ from left to right. For ease of notation, we define $k = b\log_2(n)$. For every $j \in \{k, \ldots, \sqrt{n}\}$, define $A_j \subseteq [n]$ to be the j-th column of qubits, $B_j \subseteq [n]$ to be the qubits in columns $j-k+1$ through j, and $L_j \subseteq [n]$ to be the qubits in columns 1 through j. For every $j \in \{k, \ldots, \sqrt{n}-1\}$, define $R_j \subseteq [n]$ to be the qubits in columns $j+1$ through n and define $\rho_j = \mathrm{Tr}_{R_j}(|\psi\rangle\langle\psi|)$, which is the reduced state of $|\psi\rangle$ on the j leftmost columns of qubits; we use $q_j = q_{L_j}$ to denote the marginal distribution of q on L_j, so for every $x \in \{0,1\}^{j\sqrt{n}}$, $q_j(x) = \langle x|\rho_j|x\rangle$.

The algorithm proceeds in a column-by-column fashion. First, the algorithm samples $y(k) \sim q_{B_k}$. We will explain how to perform this sampling later. Then, for every $j = k+1, \ldots, \sqrt{n}$, the algorithm samples

$$x(j) \sim q_{A_j|B_{j-1}}(\cdot|y(j-1)) \tag{4.86}$$

and then sets $y(j)$ to be the concatenation of the bits in the rightmost $k-1$ columns of $y(j-1)$ and $x(j)$. Lastly, the algorithm forms the n-bit string $z = y(k)x(k+1)\cdots x(\sqrt{n})$ and replaces the bits in z corresponding to the qubits in H with uniformly random bits to get an n-bit string z'. The output of the algorithm is z'.

We call the above algorithm the truncated column-by-column algorithm. To perform the first sampling step $y(k) \sim q_{B_k}$, it suffices to simulate the backward lightcone circuit of

B_k with respect to U, which amounts to simulating a constant-depth Clifford circuit with $O(\sqrt{n}\log n)$ qubits. Similarly, for every subsequent sampling step, it reduces to simulating the backward lightcone circuit of $A_j \cup B_{j-1}$ with respect to U, which is also a constant-depth Clifford circuit with $O(\sqrt{n}\log n)$ qubits. Thus, the whole algorithm requires simulating $O(\sqrt{n})$ constant-depth Clifford circuits in total, each acting on $O(\sqrt{n}\log n)$ qubits. We can simulate each backward lightcone circuit using the fast planar Clifford simulation algorithm of [91] in $\widetilde{O}((\sqrt{n}\log n)^{\omega/2})$ time, with ω being the matrix multiplication exponent. Then, the total runtime of the algorithm is $\widetilde{O}((\sqrt{n}\log n)^{\omega/2}\sqrt{n}) = \widetilde{O}(n^{(\omega+2)/4})$. As long as $\omega > 2$, this is faster than simulating the whole circuit U all at once using the fast planar Clifford simulation algorithm, which would take $\widetilde{O}(n^{\omega/2})$ time.

Next, we show the correctness of the algorithm. The strategy is to show that assuming the input circuit U is good, the truncated column-by-column algorithm samples from the same distribution as a full column-by-column algorithm whose correctness is easier to see. The full algorithm proceeds analogously to the truncated algorithm, with the differences being that in the step of (4.86) for every $j = k+1, \ldots, \sqrt{n}-1$, the full algorithm samples $x(j)$ from $q_{A_j|L_{j-1}}$ conditioned on all previously sampled bits, not just those in the k columns directly to the left of the j-th column. For $j = \sqrt{n}$, the full algorithm samples $x(j)$ from $\tilde{q}_{A_j|L_{j-1}\setminus H}$ conditioned on all previously sampled bits that do not correspond to qubits in H.

To see the correctness of the full column-by-column algorithm, notice that by (4.85), at the end of each step $j = k+1, \ldots, \sqrt{n}-1$, the concatenation of all the bits sampled so far is distributed according to q_j. For the step $j = \sqrt{n}$, the algorithm begins with $y \sim q_{j-1}$ as the concatenation of all previously sampled bits. Recall we assumed that H lands on the left half of the grid, so $H \subseteq L_{j-1}$. Let y' be obtained from y by deleting the bits y_i for all $i \in H$. Then y' is distributed according to $\tilde{q}_{L_{j-1}\setminus H}$. For the step $j = \sqrt{n}$, the full algorithm samples $x(j) \sim \tilde{q}_{A_j|L_{j-1}\setminus H}(\cdot|y')$, so the concatenation of y' with $x(j)$ is distributed according to $\tilde{q}$ as required.

We will need the following lemma, which is similar to Corollary 4.3. It is directly applicable to the analysis of each sampling step of both the truncated and full algorithms, except for $j = \sqrt{n}$.

Lemma 4.5. *Let $j \in \{k+1, \ldots, \sqrt{n}-1\}$. It holds that every non-identity Z-type stabilizer of ρ_j is column-wise decomposable using the Z-type stabilizers of ρ_j.*

Proof. Let P be a non-identity Z-type stabilizer of $\rho_j = \mathrm{Tr}_{R_j}(|\psi\rangle\langle\psi|)$. Then $P \otimes I^{\otimes|R_j|}$ is a Z-type stabilizer of $|\psi\rangle$ that has a hole. Thus, by Theorem 4.2, $P \otimes I^{\otimes|R_j|} = \prod_{i=1}^{r} Q_i$ for some $r \le cw(P)$ and Z-type stabilizers Q_i of $|\psi\rangle$ with $cw(Q_i) \le b\log_2 n + 1$. Let j^* be the index of the rightmost column that P has non-trivial support on. Let $i \in [r]$. By the way Q_i is constructed, we deduce that Q_i is the identity on all qubits in columns $\{j^*+1, \ldots, \sqrt{n}\}$, which includes R_j since $j^* \le j$. Thus, Q_i has the form $Q_i = Q_i' \otimes I^{\otimes|R_j|}$ for some Z-type Pauli operator Q_i', so Q_i' is a Z-type stabilizer of ρ_j. Therefore, $P = \prod_{i=1}^{r} Q_i'$ is a column-wise decomposition of P in terms of the Z-type stabilizers of ρ_j. □

We will show that the truncated algorithm samples from the same distribution as the full algorithm in every step. In the first step, both algorithms sample $y(k) \sim q_{B(k)}$, so the two algorithms behave identically by definition. Now let $j \in \{k+1, \ldots, \sqrt{n}-1\}$. We first recall how q_j is related to a system of linear equations. Let $S_{\rho_j}^Z$ be the set of all non-identity Z-type

stabilizers of ρ_j. Every element $P \in S^Z_{\rho_j}$ can be identified with a string $w \in \{0,1\}^{j\sqrt{n}} \setminus \{0^{j\sqrt{n}}\}$ and a sign $s \in \{1,-1\}$ so that $P = sZ(w)$, and P specifies a linear equation $w \cdot x = (1-s)/2$ for the variables $x \in \{0,1\}^{j\sqrt{n}}$ where $\cdot$ denotes the inner product in $\mathbb{F}_2^n$. The whole set $S^Z_{\rho_j}$ defines a (possibly redundant) system of linear equations E_j over $\mathbb{F}_2^n$ with one linear equation per element of $S^Z_{\rho_j}$, and the distribution q_j is the uniform distribution over the set of all solutions of E_j.

For every $P \in S^Z_{\rho_j}$, by Lemma 4.5, we obtain $P = \prod_{i=1}^r Q_i$ where $r \leq cw(P)$ and each Q_i is a Z-type stabilizer of ρ_j satisfying $cw(Q_i) \leq k+1$. Thus, we can construct an equivalent system of linear equations F_j where for every $P \in S^Z_{\rho_j}$, we replace the single equation defined by P in E_j with r equations, one for each Q_i. The distribution q_j is now the uniform distribution over the set of all solutions of F_j. Notice that in F_j, every linear equation that involves a variable in $\{x_i : i \in A_j\}$ can only involve other variables in $\{x_i : i \in B_{j-1}\}$.

During step j of the full algorithm, it has $y \sim q_{j-1}$ from the previous step and samples from the conditional marginal distribution $q_{A_j|L_{j-1}}(\cdot|y)$. The distribution $q_{A_j|L_{j-1}}(\cdot|y)$ is also the uniform distribution over the set of all solutions of a system of linear equations that is related to F_j in the following way. We start with the system F_j. For every $i \in L_{j-1}$ and every equation involving the variable x_i, we substitute and fix the value $x_i = y_i$. After all the substitutions, all linear equations that do not involve a variable in $\{x_i : i \in A_j\}$ are already satisfied since $y \sim q_{j-1}$. Thus, performing the substitution for all $i \in L_{j-1}$ produces a system of linear equations F'_j in the variables $\{x_i : i \in A_j\}$, and $q_{A_j|L_{j-1}}(\cdot|y)$ is the uniform distribution over the set of all solutions of F'_j. Notice that the coefficients of F'_j are fully determined by the bits in $\{y_i : i \in B_{j-1}\}$. Therefore, by induction on $j = k+1, \ldots, \sqrt{n}-1$, $q_{A_j|L_{j-1}}(\cdot|y) = q_{A_j|B_{j-1}}(\cdot|y(j-1))$. Therefore, the truncated algorithm behaves identically to the full algorithm for every step j such that $j < \sqrt{n}$.

The step for $j = \sqrt{n}$ follows from exactly the same reasoning as for the previous steps, except that we need to consider $\tilde{\rho}$ and $\tilde{q}$. We again analyze the behaviour of the full column-by-column algorithm. Let $S^Z_{\tilde{\rho}}$ be the set of all non-identity Z-type stabilizers of $\tilde{\rho}$. For every $P \in S^Z_{\tilde{\rho}}$, by Corollary 4.3, we get $P = \prod_{i=1}^r Q_i$ where $r \leq cw(P)$ and each Q_i is a Z-type stabilizer of $\tilde{\rho}$ satisfying $cw(Q_i) \leq k+1$, except for possibly one $i^* \in [r]$ such that H is supported on a column between the leftmost and rightmost columns that Q_{i^*} is supported on and $cw(Q_{i^*}) \leq 2k+1$. We obtain a system of linear equations F for $\tilde{q}$ where for every $P \in S^Z_{\tilde{\rho}}$, we replace the equation defined by P with r equations defined by each Q_i. Then, we condition on $y' \sim \tilde{q}_{L_{j-1} \setminus H}$, which amounts to substituting $x_i = y'_i$ for every $i \in L_{j-1} \setminus H$. Since H lands on the left half of the grid and $j = \sqrt{n}$, we obtain a system of linear equations F' on variables $\{x_i : i \in A_j\}$ such that every equation in F' has all the coefficients specified by the bits in $\{y'_i : i \in B_{j-1}\}$. Therefore, $\tilde{q}_{A_j|L_{j-1}\setminus H}(\cdot|y') = \tilde{q}_{A_j|B_{j-1}}(\cdot|y(j-1)) = q_{A_j|B_{j-1}}(\cdot|y(j-1))$.

Since every corresponding step of the truncated and full column-by-column algorithms samples from the same distribution, the final outputs of both algorithms are also sampled from the same distribution. Since the full algorithm samples from the computational-basis output distribution of $\tilde{\rho} \otimes (I/2)^{\otimes |H|}$, the truncated column-by-column algorithm also samples from the correct distribution.

Chapter 5

Alternating optimization methods for robust Lindbladian estimation

In this chapter, we consider the problem of characterizing noisy 2-qubit gates implemented on current quantum computing hardware. This chapter is based on the paper [4].

Suppose we have in our possession a noisy 2-qubit quantum computer, and we program it to implement a CNOT gate. How can we determine the actual quantum channel (i.e., CPTP map) executed by the quantum computer? Quantum process tomography (QPT) [92, 93] is a standard technique that solves this problem in principle and yields a full description of the channel in some matrix representation. While the quantum and classical resource costs of QPT grow exponentially with the number of qubits, it is feasible for 2-qubit gates.

A standard QPT protocol for characterizing a 2-qubit channel $\mathcal{E}$ involves estimating the Pauli expected values $\mathrm{Tr}(P_i\mathcal{E}(\rho_j))$ for 15 non-identity 2-qubit Pauli matrices P_i and 16 linearly independent initial states ρ_j, which constitutes sufficient information to solve for $\mathcal{E}$. As stated, this protocol involves running $15 \times 16 = 240$ circuits, each a large number of times (e.g., 10^4 times) to estimate one expected value $\mathrm{Tr}(P_i\mathcal{E}(\rho_j))$. For example, the following circuit will be one of the circuits that need to be executed during the standard 2-qubit QPT protocol.

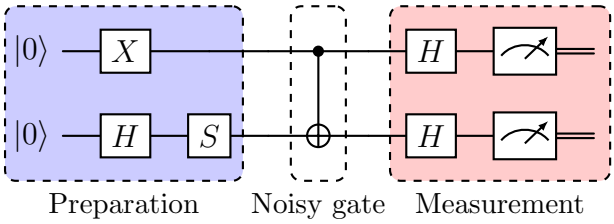


Figure 5.1: Circuit diagram for estimating $\mathrm{Tr}((X \otimes X)\mathcal{E}(|1\rangle\langle 1| \otimes |+_i\rangle\langle +_i|))$ where $\mathcal{E}$ corresponds to a noisy CNOT gate.

Notice that in the above circuit diagram, the input state $|1\rangle\langle 1| \otimes |+_i\rangle\langle +_i|$ to the noisy CNOT gate is itself prepared by applying noisy single-qubit gates to fixed $|0\rangle$ states, whose initialization is also imperfect. Likewise, the measurement of each qubit in the X basis is realized by first applying a Hadamard gate, which is noisy, and then measuring in the computational basis, which is also noisy. As a result, while the theoretical treatments of QPT tend to assume that the state preparations and measurements are noiseless, since in practice they are inevitably carried out on the same noisy device as the noisy gate, the channel

estimates obtained from QPT may be skewed by state preparation and measurement (SPAM) errors. Moreover, since each circuit is only run a finite number of times, some amount of statistical error will always be present in the outputs of QPT in a real experiment.

Gate set tomography (GST) [94–101] is an extension of QPT designed to address the problem of SPAM errors. Rather than analyzing individual gates in isolation, it attempts to self-consistently estimate SPAM and gate errors of a set of gates all at once. We give a brief overview of GST in Section 5.1.3.

Ultimately, both QPT and GST output a set of quantum channel estimates returned in some matrix representation, one of which is called the transfer matrix. Whenever a superoperator $\mathcal{E}$ can be expressed in the form $\mathcal{E}(\rho) = \sum_k A_k \rho B_k$ for some complex matrices $\{A_k\}$ and $\{B_k\}$, its transfer matrix representation is $E = \sum_k A_k \otimes B_k^T$. While the transfer matrix contains full information about the noisy gate, they can be difficult to interpret and act on. It is often more useful to build a model of the underlying physical process that generates the quantum channel. One approach tries to understand the channel as a solution to a quantum master equation [102]. In particular, it is often assumed that noisy processes on quantum devices are well approximated by memoryless, or Markovian, dynamics. Extracting the generator of a Markovian evolution, which is called the *Lindbladian*, delivers qualitative information about the noise that is otherwise obscured in the transfer matrix picture.

A number of methods have been proposed and demonstrated for fitting Lindbladian models to full (i.e., transfer matrix) or partial tomographic data [22, 103–110]. Previous works on the task of Lindbladian estimation can be broadly divided into three strands: (i) logarithm search methods, (ii) direct numerical optimization methods, and (iii) ad-hoc methods tailored to particular experimental scenarios. Logarithm search methods start by taking the logarithm of the transfer matrix representation of the channel, which was first studied formally for transfer matrices with distinct eigenvalues in [106, 107] and extended to the degenerate case in [111, 112]. These methods must contend with the fact that the matrix logarithm is non-unique and is sensitive to statistical errors in collecting the tomographic data. The initial theoretical works [106, 107] did not address SPAM or statistical errors at all. The direct precursor to the work presented in this chapter [111] proposed an implementable algorithm for fitting a Lindbladian model to a transfer matrix. However, it did not actively address SPAM errors, and the runtime of the algorithm as stated was extremely long for the 2-qubit case, making it impractical for use in practice.

Prior practical demonstrations of Lindbladian estimation from noisy experimental data have tended to fall under the heading of direct numerical optimization. The general approach is to define the problem in terms of an objective function that compares a parameterized Markovian channel with the tomographic data and pass this to an off-the-shelf numerical optimization solver. A key observation is that the set of all d^2-by-d^2 Lindbladian matrices can be exactly parameterized by $d^2(d^2-1)$ real variables. Furthermore, gradients of the objective function can often be obtained mechanically via automatic differentiation, and in such cases, gradient-based solvers can be used. Finally, a good initial guess required by the solvers is often available based on the experimenter's prior knowledge about the input channel. Prior works in this direction include [97, 105, 109, 110, 113].

Approaches to Lindbladian estimation that fall outside these general methods tend to rely on some assumption about the system being investigated, so that the Lindbladian model is restricted in some way, and the number of parameters to estimate is reduced. Such

techniques have been applied in NMR [103], quantum optics [104], and superconducting-qubit [100, 108, 114] settings for small systems. A method for studying larger systems under the assumption that the Markovian noise process is generated solely by local Pauli operators has also been studied [22, 23]. This method is scalable, but the assumption of a Pauli-Lindblad model is rather stringent. A more detailed discussion of prior work is deferred to Section 5.6.

In this work, we build on the method of [111], introducing algorithmic improvements to reduce the runtime and strengthen robustness to statistical errors in practical cases. Additionally, we use ideas from gate set tomography [94, 96, 97] to show that Lindbladian fitting techniques can also be made robust against realistic SPAM errors. We demonstrate the application of this approach on tomographic data collected from real noisy quantum computing hardware accessed via the cloud [115]. Our methods complement other Lindbladian estimation approaches [97, 109, 116, 117] and can be used to cross-validate AI-based heuristic algorithms to study the interplay between Markovian and non-Markovian dynamics [118, 119].

In the following sections, we give practical algorithms for Lindbladian fitting in two settings: (i) characterization of an individual noisy process, neglecting SPAM errors, and (ii) simultaneous and self-consistent characterization of a set of processes, allowing for SPAM errors. We do not attempt to tackle these problems in full generality but rather focus on the problem of fitting Lindbladian models to noisy 2-qubit quantum logic gates, where the noise is not arbitrarily strong. We build on the logarithm search methods developed in [106, 107, 111] in three main respects.

1. We clarify the conditions under which good solutions can already be found by running convex optimization over a few logarithm branches of the input transfer matrix and show that this includes classes of practically relevant quantum channels larger than previously thought. In particular, this includes cases whose transfer matrix has a degenerate spectrum. Instead, our new conditions merely require that the transfer matrix does not have eigenvalues close to being real negative. We call this simple algorithm the *Convex Solve* method.

2. We develop a novel algorithm to tackle the cases where some continuous degrees of freedom cannot be avoided. The method leverages the fact that in practice, we can start with an initial best guess for the Markovian channel; namely, the ideal gate that we program the quantum computer to implement. At an intuitive level, we use the eigendecomposition of the best-guess Lindbladian to guide the construction of a matrix logarithm and iteratively improve this best guess by projecting back to the set of Lindbladians. This leads to runtimes reduced by orders of magnitude compared to the algorithm given in [111]. We call this the *Alternating Projections* method.

3. We propose a protocol that synthesizes Lindbladian fitting and gate set tomography [96] techniques by self-consistently alternating between a SPAM estimation step and a Lindbladian fitting step. For the Lindbladian fitting step, we use the Convex Solve and Alternating Projections methods from points 1 and 2 as subroutines in our implementation; however, we note that the template for the protocol is generic, and

any method for fitting a Lindbladian to a transfer matrix could be substituted. We call this algorithm *Gate Set Flip-Flop*.

The remainder of the chapter is structured as follows. We explain the necessary technical background in Section 5.1, covering Markovian channels, Lindbladians, quantum process tomography, and gate set tomography. We describe our novel algorithms in detail in Section 5.2: we present in Section 5.2.1 our Convex Solve and Alternating Projections algorithms and in Section 5.2.2 Gate Set Flip-Flop. In Sections 5.3 and 5.4, we benchmark the performance of the new methods extensively using synthetic data with realistic levels of statistical error and a wide range of physically plausible gate noise types and strengths. Finally, in Section 5.5, we demonstrate the application of the algorithms to data obtained from real superconducting-qubit hardware. In particular, we focus on characterizing the noise associated with 2-qubit entangling gates and idling processes and assess whether and how the effective noise processes change when other gates are activated in parallel elsewhere on the device.

5.1 Preliminaries

In this section, we provide an overview of the necessary background and establish common notation used throughout this chapter. If not specified otherwise, all norms $\|\cdot\|$ denote the Frobenius norm.

5.1.1 Markovian channels and Lindblad generators

The noisy dynamics that we wish to understand can be described by a quantum channel $\mathcal{E} : \rho \to \mathcal{E}(\rho)$ acting on $d \times d$ density matrices. We can represent density matrices ρ as vectors using stacked density matrix notation

$$|\rho\rangle\!\rangle_{j,k} = \langle\!\langle e_j, e_k|\rho\rangle\!\rangle \coloneqq \langle e_j|\rho|e_k\rangle, \tag{5.1}$$

where $|e_j\rangle = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ with a 1 in the j-th position. A quantum channel $\mathcal{E}$ can then be represented by a $d^2 \times d^2$ matrix as follows

$$E_{(j,k),(l,m)} = \langle\!\langle e_j, e_k|E|e_l, e_m\rangle\!\rangle \coloneqq \mathrm{Tr}[|e_k\rangle\langle e_j|\mathcal{E}(|e_l\rangle\langle e_m|)]. \tag{5.2}$$

The matrix E is referred to as the *transfer matrix* or the elementary basis representation of $\mathcal{E}$. For example, for $\mathcal{E}$ representing a noisy 2-qubit quantum logic gate, $d = 4$, so E is a 16×16 matrix. It is straightforward to verify using (5.1) and (5.2) that $E|\rho\rangle\!\rangle = |\mathcal{E}(\rho)\rangle\!\rangle$. Thus, in this representation, the action of the channel on the density matrix is given by matrix-vector multiplication, and the composition of channels corresponds to matrix-matrix multiplication. In practice, whenever $\mathcal{E}$ is expressed in the form

$$\mathcal{E}(\rho) = \sum_k A_k \rho B_k \tag{5.3}$$

for some complex matrices $\{A_k\}$ and $\{B_k\}$, we have

$$E = \sum_k A_k \otimes B_k^T. \tag{5.4}$$

In this picture, we can conveniently express the expectation value of a Hermitian observable O on the state obtained by applying the channel $\mathcal{E}$ to the initial state ρ as

$$\langle\langle O|E|\rho\rangle\rangle = \mathrm{Tr}[O\mathcal{E}(\rho)], \tag{5.5}$$

where the operator O is vectorized in the same way as in (5.1).

Throughout this chapter, we are interested in Markovian channels. These are a subset of channels that are generated by Lindbladians, i.e., by generators of the form

$$\mathcal{L}(\rho) \coloneqq i[\rho, H] + \sum_\alpha \gamma_\alpha \left[J_\alpha \rho J_\alpha^\dagger - \frac{1}{2}\left(J_\alpha^\dagger J_\alpha \rho + \rho J_\alpha^\dagger J_\alpha \right) \right] \tag{5.6}$$

where H is a Hamiltonian, each J_α is called a jump operator, and each γ_α is a real positive scalar [120]. In this canonical form, H is traceless, all jump operators are traceless and normalized to Frobenius norm 1, and $\mathrm{Tr}(J_\alpha^\dagger J_{\alpha'}) = 0$ whenever $\alpha \neq \alpha'$. The first term in the Lindblad form describes the unitary part of the evolution, while the second term represents the dissipative part of the process. In this work, we restrict our analysis to time-independent Lindbladian generators for which the transfer matrix of the corresponding channel at time t is given by $E(t) = e^{Lt}$ where L is the elementary basis representation of $\mathcal{L}$.

The set of all Lindbladians in the elementary basis representation forms a closed convex cone with a simple known characterization [106, 107] through the Γ-involution, which is defined by linearly extending

$$(|e_j, e_k\rangle\rangle\langle\langle e_l, e_m|)^\Gamma \coloneqq |e_j, e_l\rangle\rangle\langle\langle e_k, e_m|. \tag{5.7}$$

A matrix L is a Lindbladian if and only if

1. L^Γ is Hermitian, i.e., the Choi matrix of L is Hermitian;
2. L^Γ is *conditionally completely positive*, i.e., $\omega_\perp L^\Gamma \omega_\perp \succeq 0$ where $\omega_\perp = (I - |\omega\rangle\rangle\langle\langle\omega|)$, and $|\omega\rangle\rangle \coloneqq \frac{1}{\sqrt{d}} \sum_{j=1}^d |e_j, e_j\rangle\rangle$;
3. $\langle\langle\omega|L = 0$, corresponding to the trace-preserving property.

For a small Hilbert space dimension like the $d = 4$ case considered in this work, the problem of projecting a matrix to the set of all Lindbladians can be solved efficiently using convex optimization. The projection problem is minimizing $\|X - L\|_F$ subject to L being a Lindbladian, with X being the input and L being the optimization variable. In software, performing the projection amounts to setting up the problem using CVXPY [121] and then calling an off-the-shelf convex optimizer such as the SCS solver [122].

Given a Lindbladian L in the elementary basis representation, it is possible to solve for a canonical decomposition comprised of a Hamiltonian H, jump operators $\{J_\alpha\}$, and real positive scalars $\{\gamma_\alpha\}$ such that the elementary basis representation of the Lindbladian form

(see (5.6)) defined by H, $\{J_\alpha\}$, and $\{\gamma_\alpha\}$ is L. For completeness, we include a derivation of this fact in Section 5.6.3.

For $E = e^L$, L is a matrix logarithm of E. The set of logarithms of E is defined as $\log(E) = \{A : e^A = E\}$, and $\log(E)$ has infinite cardinality even for 1×1 matrices since for every $\lambda \in \mathbb{C}$, $\lambda + 2\pi k i$ is a logarithm of e^λ for every integer k. Assume E is diagonalizable with not necessarily unique complex eigenvalues $\mu_1, \ldots, \mu_{d^2}$. For every $j \in \{1, \ldots, d^2\}$, we can fix a logarithm λ_j of μ_j so that $e^{\lambda_j} = \mu_j$ and $|\Im(\lambda_j)| \leq \pi$. If μ_j is not real negative, then the choice of λ_j is unique. Then every $A \in \log(E)$ can be written in the form

$$A = \sum_{j=1}^{d^2} (\lambda_j + 2m_j\pi i)|r_j\rangle\rangle\langle\langle l_j| \tag{5.8}$$

where $m = (m_1, \ldots, m_{d^2}) \in \mathbb{Z}^{d^2}$ is an integer vector and $|r_j\rangle\rangle$ ($|l_j\rangle\rangle$) is a unit right (left) eigenvector of E corresponding to the eigenvalue μ_j. Choosing an $m \in \mathbb{Z}^{d^2}$ corresponds to choosing a branch of the matrix logarithm. We call a branch satisfying $|\Im(\lambda_j) + 2m_j\pi| \leq \pi$ for every $j \in \{1, \ldots, d^2\}$ a principal branch, and it is uniquely specified by $m = (0, \ldots, 0)$ if E has no real negative eigenvalues. When $\mu_1, \ldots, \mu_{d^2}$ are all distinct, E has d^2 unique eigenspace projectors $|r_1\rangle\rangle\langle\langle l_1|, \ldots, |r_{d^2}\rangle\rangle\langle\langle l_{d^2}|$, so $\log(E)$ is a countably infinite set whose discrete degrees of freedom are fully parameterized by the choice of m.

The case where E has degenerate eigenvalues is more intricate but not uncommon: this always occurs if E corresponds to a unitary gate. A degenerate eigenspace of E could split into different eigenspaces of A, and we call this phenomenon *eigenspace splitting*. For instance, if $\mu_j = \mu_k$ for some $j \neq k$, then in specifying an $A \in \log(E)$, $|r_j\rangle\rangle$ and $|r_k\rangle\rangle$ could be any of the uncountably many unit eigenvectors of E corresponding to the eigenvalue $\mu_j = \mu_k$ while $\lambda_j + 2m_j\pi i \neq \lambda_k + 2m_k\pi i$ is possible. Therefore, there are continuous degrees of freedom in choosing an $A \in \log(E)$ if E has degenerate eigenvalues. It is not hard to show that for a Lindbladian L, the condition L^Γ is Hermitian (see (5.7)) implies that all the complex eigenvalues of L come in complex conjugate pairs [106, 107, 111]. As a result, the quintessential example of this eigenspace splitting phenomenon occurs for us when -1 is a 2-fold degenerate eigenvalue of E, and this 2-dimensional eigenspace gets split into two 1-dimensional eigenspaces of A corresponding to the eigenvalues πi and $-\pi i$.

5.1.2 Linear inversion quantum process tomography

In quantum process tomography [92, 93], a quantum channel is estimated by implementing the same process on a quantum device a large number of times on different input states and measuring each resulting output state in different bases. More concretely, let the $d^2 \times d^2$ transfer matrix for the process of interest be denoted E^*, and let $\{|\rho_j\rangle\rangle : j \in [d^2]\}$ and $\{\langle\langle F_i| : i \in [d^2]\}$ be the chosen tomographically complete set of preparation states and measurement operators, respectively. Typically, E^* is assumed to be the noisy implementation of some ideal process E^{ideal}. Without loss of generality, we assume that each F_i represents one outcome of a two-outcome POVM $\{F_i, I - F_i\}$. The quantum circuit that prepares state ρ_j, implements the target process and then carries out the measurement corresponding to F_i is executed many times to obtain sufficient statistics to estimate the value of $\langle\langle F_i|E^*|\rho_j\rangle\rangle = \text{Tr}[F_i\mathcal{E}^*(\rho_j)]$. This must be repeated for each pairing of measurement setting

F_i with preparation ρ_j. We can organize the chosen state preparations and measurements into matrices. The measurement matrix is given by a matrix in which each row is the effect of a possible measurement outcome

$$A^{\text{ideal}} = \begin{pmatrix} \langle\langle F_1| \\ \langle\langle F_2| \\ \vdots \\ \langle\langle F_i| \\ \vdots \\ \langle\langle F_{d^2}| \end{pmatrix} \tag{5.9}$$

while the state preparation matrix is given by a matrix in which each column is one of the possible input states

$$B^{\text{ideal}} = \begin{pmatrix} |\rho_1\rangle\rangle & |\rho_2\rangle\rangle & \cdots & |\rho_j\rangle\rangle & \cdots & |\rho_{d^2}\rangle\rangle \end{pmatrix}. \tag{5.10}$$

In the case of an ideal process E^{ideal}, the probabilities of obtaining outcome F_i after preparing in state ρ_j can be collected into a matrix

$$P^{\text{ideal}} = A^{\text{ideal}} E^{\text{ideal}} B^{\text{ideal}}. \tag{5.11}$$

Note that A^{ideal}, E^{ideal}, and B^{ideal} are all matrices chosen by the experimenter, so P^{ideal} is known. However, in reality, the circuit elements used to implement A^{ideal}, E^{ideal}, and B^{ideal} will be noisy. Thus, the true probabilities being estimated in quantum process tomography will be

$$P^* = A^* E^* B^* \tag{5.12}$$

for some unknown P^*, A^*, E^*, and B^* different from P^{ideal}, A^{ideal}, E^{ideal}, and B^{ideal} respectively. In quantum process tomography, it is assumed that the state preparation and measurement settings are implemented exactly, whereas the true process E^* is not known. Namely, $A^* = A^{\text{ideal}}$ and $B^* = B^{\text{ideal}}$. Then, the task is to estimate E^* from P^*. There are a number of ways to do this. In linear inversion tomography, A^{ideal} and B^{ideal} are chosen to be square and invertible, so under the assumption that $A^* = A^{\text{ideal}}$ and $B^* = B^{\text{ideal}}$, E^* can be estimated via the identity

$$E^* = (A^{\text{ideal}})^{-1} P^* (B^{\text{ideal}})^{-1}. \tag{5.13}$$

There are a number of more refined methods to estimate E^* in the quantum process tomography literature, but a shared feature of conventional approaches is the assumption that $A^* = A^{\text{ideal}}$ and $B^* = B^{\text{ideal}}$ [123, 124].

In this work, we focus on E^{ideal} being a 2-qubit quantum logic gate, so A^{ideal} and B^{ideal} will each contain 16 measurement and state preparation settings respectively. Measuring each setting independently therefore requires a suite of $16^2 = 256$ different circuits to estimate P^*. In practice, the circuit overhead can be reduced by parallelizing the measurement of compatible observables. For example, by measuring both qubits in the Z basis, we can infer mean values for $Z \otimes I$, $I \otimes Z$, and $Z \otimes Z$ using the same circuit. Using this scheme and

measuring in the Pauli basis, the number of circuits can be reduced to 144. In a real experiment, since each matrix element $P^*_{ij} = \langle\langle F_i|E^*|\rho_j\rangle\rangle$ of P^* is estimated by collecting statistics from a finite number of circuit executions (called tomographic shots), the experimental estimate for P^* (and in turn the estimate for E^*) will inevitably differ from its true value, even when preparations and measurements are ideal. The amount of statistical error can be reduced by increasing the number of tomographic shots; however, doing so to a significant extent may be prohibitive on current quantum computing hardware. The practicality of taking a large number of shots is limited not just by the overall resource cost but also by the instability of noise processes on the device. In other words, the duration of the experiment increases with the number of shots taken, increasing the risk that E^* can drift while data is being taken [125]. Statistical error can also be reduced by projecting the output of process tomography to CPTP, and there are known algorithms for doing so [123, 124, 126].

5.1.3 Gate set tomography

In this subsection, we give a brief overview of the simplest gate set tomography (GST) protocol, linear inversion GST [94]. A more comprehensive review including more sophisticated methods can be found in [96]. In quantum process tomography (whether linear inversion tomography or some more sophisticated method), assuming $A^* = A^{\text{ideal}}$ and $B^* = B^{\text{ideal}}$ is a prerequisite to computing an estimate of E^*. Operationally, this corresponds to the assumption that there are no SPAM errors in the tomography experiment, and the experimenter knows the A^{ideal} and B^{ideal} matrices because they are chosen by the experimenter. This is a strong assumption that will never be satisfied exactly and may not even be approximately satisfied in near-term devices.

In linear inversion GST, this assumption is replaced by two weaker assumptions:

- when we try to implement the (known) informationally complete state preparations B^{ideal} and measurements A^{ideal}, the state preparations B^* and measurements A^* that are actually implemented (due to SPAM errors) are still informationally complete;
- the SPAM errors are independent of which quantum process is being implemented, and remain fixed for a given preparation or measurement setting.

With these two assumptions, the GST method is as follows. Choose a set of k distinct quantum gates $E_1^{\text{ideal}}, \ldots, E_k^{\text{ideal}}$ (here each E_i^{ideal} is the transfer matrix of a unitary gate) and together with the (instantaneous) identity gate, form a gate set

$$\mathcal{G} = \{I, E_1^{\text{ideal}}, \ldots, E_k^{\text{ideal}}\}. \tag{5.14}$$

Then carry out quantum process tomography for every gate in $\mathcal{G}$. For each gate $i \in \{1, \ldots, k\}$, let E_i^* denote the actual noisy gate implemented by the quantum device. This enables the construction of the matrices

$$P_i^* = A^* E_i^* B^*. \tag{5.15}$$

Since A^* and B^* are unknown, it is not possible to reconstruct E_i^* from P_i^*. However, by implementing the identity gate instantaneously, we can construct what is known as the *Gram matrix*

$$g^* = A^* B^*. \tag{5.16}$$

We can therefore construct an estimate

$$E_i = B(g^*)^{-1} P_i^* B^{-1} = B(B^*)^{-1} E_i^* B^* B^{-1} \tag{5.17}$$

for E_i^* up to some *similarity transform* B.

This may seem unfinished—since we do not know B, how can we construct an estimate of E_i^*? One of the key insights of GST is that it is only ever possible to reconstruct each E_i^* up to some similarity transform since these similarity transforms are unobservable. Thus, (5.17) is the best we can hope to achieve for each gate. To see this, notice that for every invertible B and $A = g^* B^{-1}$,

$$A^* B^* = g^* = AB \tag{5.18}$$

and for every gate i,

$$A^* E_i^* B^* = A(B(B^*)^{-1} E_i^* B^* B^{-1}) B = A E_i B. \tag{5.19}$$

Therefore, assuming A, B, and every E_i are all physical, it is impossible to distinguish whether the true quantum process tomography experiments have occurred with measurement, preparation, and noisy gates being A^*, B^*, and E_i^* respectively or A, B, and E_i respectively since they both lead to identical Gram and probability matrices. This implies that the similarity transform B is in fact just picking out a particular gauge. The different choices of gauge give equivalent gate sets, so when considering the entire gate set, any invertible (and physical) choice of matrix is as good as any other for B. However, the choice of gauge clearly has a significant effect when we are interested in characterizing the individual gates within the gate set. If, for example, we take a perfect set of gates (with no errors) but choose a gauge that rotates between the different gates in the set, then this will imply that the gate fidelities are all very low and that there are large coherent errors in the device. However, this is merely an artifact of the chosen gauge.

In gate set tomography, the gauge is chosen to be the one that minimizes the distance to the ideal gate set (in some metric - see [96] for a discussion of different choices of metric). More precisely, in GST the gauge is chosen to be a minimizer of the optimization problem in some given norm $\|\cdot\|_*$

$$\min_B \left(\sum_{i=1}^{k} a_i \| B(g^*)^{-1} P_i^* B^{-1} - E_i^{\text{ideal}} \|_* + \sum_{j=1}^{d^2} b_j \| B_j - |\rho_j\rangle\rangle \|_* + \sum_{i=1}^{d^2} c_i \| A_i - \langle\langle F_i | \|_* \right) \tag{5.20}$$

where $A = gB^{-1}$, B_j denotes the j-th column of B, A_i denotes the i-th row of A, and $\{a_i\}$, $\{b_j\}$, and $\{c_i\}$ are non-negative weights. Intuitively, the weights should be chosen to be proportional to the believed fidelity of a gate, preparation, or measurement setting. For example, if $|\rho_1\rangle\rangle$ corresponds to preparing a known high-fidelity $|0\rangle$ state and the other $|\rho_j\rangle\rangle$'s are prepared from $|0\rangle$ by applying a short-depth circuit, it may be sensible to choose $b_1 = 1$ and $b_j = 0$ for all $j \geq 2$. The simplest choices for the weights are $a_i = 1$ for all $i \in \{1, \ldots, k\}$ and $b_j = c_j = 0$ for all $j \in \{1, \ldots, d^2\}$, hence simplifying the gauge optimization problem to

$$\min_B \sum_{i=1}^{k} \| B(g^*)^{-1} P_i^* B^{-1} - E_i^{\text{ideal}} \|_*. \tag{5.21}$$

In summary, in its simplest form, the input to GST is a set of matrices

$$\{g^*, P_1^*, \ldots, P_k^*\} = \{A^*B^*, A^*E_1^*B^*, \ldots, A^*E_k^*B^*\} \tag{5.22}$$

obtained from $k+1$ runs of quantum process tomography, and the desired output is a matrix B that minimizes (5.21). Note that B^* may not be an optimal solution for (5.21) if $E_i^{\text{ideal}} \neq E_i^*$ for at least one gate i.

5.2 Our work

5.2.1 The Lindbladian fitting problem

The input to the Lindbladian fitting problem is the transfer matrix E of a CPTP map (recall (5.2)), and the goal is to fit a Lindbladian model to E. More concretely, we wish to find a Lindbladian L that minimizes $\|e^L - E\|$.

In this work, we restrict ourselves to special cases of the Lindbladian fitting problem that satisfy two additional assumptions. Firstly, it is assumed that the input E is close to some unknown Markovian channel $E^* = e^{L^*}$. That is, $\|E - E^*\| \leq c_1$ for some small constant c_1. We further assume that E^* is not far from a known Markovian channel $E^{\text{ideal}} = e^{L^{\text{ideal}}}$. In fact, we assume $\|L^* - L^{\text{ideal}}\| \leq c_2$ for some small constant c_2 which is stronger than assuming E^* is not far away from E^{ideal}. As an example, suppose L is a Lindbladian generator for the 1-qubit Pauli Z gate, then $3L$ would be a Lindbladian generator for Z^3, which is identical to Z as a unitary gate, but $\|L - 3L\|$ is large. The algorithms we consider will search over the branches of the complex matrix logarithm $\log(E) = \{A : e^A = E\}$ in the outermost loop[1]. The bound $\|L^* - L^{\text{ideal}}\| \leq c_2$ implies a bound on the differences between the imaginary parts of the eigenvalues of L^* and L^{ideal}, which in turn limits the number of physically relevant branches of $\log(E)$ to search over. The assumption $\|L^* - L^{\text{ideal}}\| \leq c_2$ is also intended to rule out cases like $L^{\text{ideal}} = 0$, which generates the identity gate, while L^* generates a far away unitary gate such as $X \otimes I$.

In an experiment,

- E would be the output of quantum process tomography (optionally projected to CPTP),
- E^{ideal} would correspond to an ideal gate,
- L^{ideal} is a known Lindbladian generator that the experimenter has chosen to execute,
- and E^* would be the Markovian channel closest to the true noisy gate being implemented in place of E^{ideal}.

The difference between E^* and E^{ideal} captures the total amount of time-independent Markovian gate noise while the discrepancy between E and E^* accounts for possible non-Markovian or time-dependent Markovian gate noise as well as inevitable statistical error

[1] Throughout this work, we do assume E and E^{ideal} are both diagonalizable.

from performing process tomography. The two assumptions translate to the physical assumptions that the gate noise is approximately time-independent Markovian and not too strong (e.g., some noise turning Z into Z^3). In this work, the two assumptions are stated heuristically, and we do not have formal bounds on the precise values of c_1 and c_2 required by our algorithms. Nonetheless, we believe our weak approximately Markovian gate noise assumption is satisfied when running process tomography on individual gates on today's quantum computing hardware. In Section 5.3, we give numerical evidence that c_2 does not need to be too small.

The Convex Solve method

To set the stage for the Alternating Projections method and emphasize the effect of statistical error, we first describe the simpler Convex Solve method (Algorithm 1). Due to statistical error, the eigenvalues of E are (almost surely) all distinct (regardless of what E^{ideal} or E^* are). This implies $\log(E) = \{A : e^A = E\}$ is a countably infinite set, which we can enumerate over. The Convex Solve method then simply enumerates over each $A \in \log(E)$ and finds the Lindbladian L closest to A. The latter task can be accomplished by solving a positive semidefinite program (recall Section 5.1.1). Finally, the algorithm returns the L ever obtained that achieves the smallest $\|e^L - E\|$. Again, by the assumption $\|L^* - L^{\text{ideal}}\| \leq c_2$, we can terminate the search over the branches of $\log(E)$ as soon as $\|A - L^{\text{ideal}}\|$ becomes too large (recall the Z vs. Z^3 example). In practice, rather than explicitly checking $\|A - L^{\text{ideal}}\|$, we can restrict the search to a few low-lying branches of the matrix logarithm. We note that close variants of the Convex Solve method have been considered in prior works [106,107,111], specifically for the case where the eigenvalues of E^* are all distinct. As such, we do not claim that our main contribution here lies in proposing the Convex Solve method; rather, it is an improved understanding of the effectiveness of this strategy, which we explain next.

The simplest special case of the Convex Solve method simply returns the Lindbladian closest to a principal branch of $\log(E)$, which is unique as long as E does not have real negative eigenvalues even when E has degenerate eigenvalues. In Section 5.6.4, we prove the following result, which applies in the absence of statistical error. Let $\mathcal{L}$ denote the set of all Lindbladians.

Theorem 5.1. *Suppose there exists a Lindbladian L^* such that $E^* = \exp(L^*)$ and*

$$\|L^* - L^{ideal}\| < \frac{\pi - \rho(L^{ideal})}{\kappa(V)}, \tag{5.23}$$

where V is an invertible matrix that diagonalizes L^{ideal}, $\kappa(V)$ is its condition number, and $\rho(L^{ideal})$ is the spectral radius of L^{ideal}. Then given $E = E^*$ as input, the Convex Solve method outputs an $L \in \mathcal{L} \cap \log(E)$.

This tells us that under the assumption of weak, time-independent Markovian noise, even when E^* contains degenerate eigenvalues, the simplest special case of the Convex Solve method already solves the problem in the absence of statistical error (i.e., $E = E^*$), provided $|\Im(\lambda)| < \pi$ for every eigenvalue λ of L^{ideal}. The last condition can be satisfied when the transfer matrix $\exp(L^{\text{ideal}})$ has no eigenvalues close to the negative real line, which holds

with a meaningful gap for many important gates, including $\sqrt{X} \otimes I$, $T \otimes I$, and most notably, the identity gate $I \otimes I$. Of course, in reality, statistical error will always be present, but empirically we observe that the result is surprisingly robust to shot noise. Our synthetic testing in Section 5.3 suggests that the Convex Solve method also already works at the level of statistical noise induced by 10^4 tomographic shots for these gates.

The Convex Solve method is not the whole story because the transfer matrices E^{ideal} of an important family of gates that we would like to characterize do contain real negative eigenvalues, so under the weak noise assumption, E^* and E will have eigenvalues close to being real negative. This family includes CNOT, ISWAP, and the tensor product of Pauli and Hadamard gates. Under a realistic level of statistical noise induced by 10^4 tomographic shots, we observe that for gates from this family, both $\|L - A\|$ and $\|e^L - E\|$ could be significantly larger than $\|E - E^*\|$ for every closest Lindbladian L to every $A \in \log(E)$ (see Section 5.3 for a CNOT example). Hence, the Convex Solve method fails on these inputs.

Algorithm 1: Convex Solve Method

input : Estimated (diagonalizable) transfer matrix E; max branch number $m_{\max}$
output: Closest Lindbladian generator L_{est} to all checked branches of logarithm

Compute (some) set of eigenvalues $\{\mu_j\}_j$ and right $\{|r_j\rangle\rangle\}_j$ and left $\{\langle\langle l_j|\}_j$ eigenvectors of E
for $j = 1$ **to** d^2 **do**
  $\lambda_j \leftarrow$ principal log of μ_j
end
for $m \in \{-m_{max}, \ldots, 0, \ldots, m_{max}\}^{d^2}$ **do**
  $A_m \leftarrow \sum_{j=1}^{d^2} (\lambda_j + 2\pi i m_j) |r_j\rangle\rangle\langle\langle l_j|$
  Solve the convex optimization problem
    minimize $\|L - A_m\|$
    subject to L^Γ Hermitian
    $\omega_\perp L^\Gamma \omega_\perp \succeq 0$
    $\langle\langle \omega | L = 0$
  and let L_m be the returned optimal solution
  $\Delta_m \leftarrow \|e^{L_m} - E\|$
end
$L_{\text{est}} \leftarrow L_{m_{\text{opt}}}$ such that m_{opt} is the branch that minimizes Δ_m
return L_{est}

The Alternating Projections method

Now, we are ready to describe the Alternating Projections method. The algorithm given in [111] operates under similar assumptions but is too slow in practice for the 2-qubit $d = 4$ case. In this work, we build upon the understanding developed in [111] and give a new algorithm that has significantly better performance in terms of both accuracy and runtime. One insight in the new algorithm is that we do not have to solve the Lindbladian fitting problem "blind". We have knowledge about what a "good" initial guess is because we know what gate we tried to implement. We can start from the Lindbladian generator of the ideal

gate L^{ideal}.

Here, we sketch the main idea of the algorithm (Algorithm 2) and defer a complete technical description of our implementation to Section 5.6.5. The input to the algorithm is E and an initial best-guess Lindbladian L_0. The default choice sets $L_0 = L^{\text{ideal}}$, but a better estimate, if available, could be used instead. Let $\mathcal{L}$ denote the set of all Lindbladians and define $\log(E) = \{A : e^A = E\}$. The goal is to compute

$$\min_{L \in \mathcal{L}} \|e^L - E\| = \min_{L \in \mathcal{L}, A \in \log(E)} \|e^L - e^A\| \tag{5.24}$$

and output an optimal Lindbladian L. We approximate the latter minimization problem by the problem

$$\min_{L \in \mathcal{L}, A \in \log(E)} \|L - A\|. \tag{5.25}$$

Note that if E is Markovian, i.e., $E = E^*$, then the optimal values of (5.24) and (5.25) are both zero, and the solution set is precisely $\mathcal{L} \cap \log(E)$. If both $\mathcal{L}$ and $\log(E)$ were closed convex sets, then the problem in (5.25) can be solved using alternating projections, which start from a given starting point and alternately project onto each of the two sets until convergence. In our case, $\mathcal{L}$ is a closed convex cone and projecting onto it can be accomplished by solving a positive semidefinite program (recall Section 5.1.1). However, the set $\log(E)$ is not convex, and we do not know how to solve the problem $\min \|A - L_0\|$ subject to $A \in \log(E)$ directly. To this end, we devise a procedure to construct an $A \in \log(E)$ that approximately minimizes $\|A - L_0\|$, inspired by the assumption that $\|E - e^{L_0}\|$ is small to begin with. Suppose E has $n \leq d^2$ distinct eigenvalues associated with n unique eigenspace projectors $\Pi_1, \ldots \Pi_n$. Let $v_1, \ldots, v_{d^2}$ be a set of eigenvectors of L_0. For every pair of eigenvector v_j and projector Π_k, we compute the cost $\|v_j - \Pi_k v_j\|$. We then solve a minimum-cost maximum-flow problem to assign each eigenvector v_j to a best-fit eigenspace Π_k. A matrix $\bar{A} \in \log(E)$ is reassembled by pairing up the logarithm of the eigenvalues of E (in some branch) with the eigenvectors $\Pi_k v_j$. The algorithm then projects $\bar{A}$ onto $\mathcal{L}$ to get a Lindbladian $\bar{L}$ and checks whether

$$\|e^{\bar{L}} - E\| < \|e^{L_0} - E\|. \tag{5.26}$$

If yes, then we update the current best-guess Lindbladian to be $\bar{L}$ and repeat from $\bar{L}$.

Eigenvalue clustering

Since the eigenvectors of E tend to differ non-trivially from the eigenvectors of E^* under realistic levels of statistical error, we adopt the eigenvalue clustering pre-processing step proposed in [111]. That is, the user needs to specify a precision parameter β so that the eigenvalues of E differing by less than β are treated as being identical, and their eigenspaces are merged into a single degenerate eigenspace. Note that this causes the matrix $\bar{A}$ constructed during the algorithm to belong to some enlarged set $\widetilde{\log}(E) = \{A : e^A \approx E\}$ as opposed to $\log(E)$, where the $\approx$ is controlled by the precision parameter β. Note that if $\beta = 0$, then $\widetilde{\log}(E) = \log(E)$. Also, notice that our Alternating Projections method reduces to the Convex Solve method if we set the precision parameter $\beta = 0$. Hence, the Alternating Projections method includes the Convex Solve method as a special case.

Algorithm 2: Alternating Projections (informal)

input : Estimated transfer matrix E; initial guess Lindbladian L_0; precision $\beta \geq 0$; maximum search depth T

output: Estimated Lindbladian L^{est} such that $\exp(L^{\text{est}}) \approx E$

Compute eigendecomposition of input tomographic transfer matrix, $E = \sum_j \mu_j |r_j\rangle\rangle\langle\langle l_j|$

Cluster tomographic eigenvalues into approximate eigenspaces S_k according to precision β

Compute principal log eigenvalues λ_j such that $e^{\lambda_j} = \mu_j$

for *branches m and random perturbations* $\tilde{L}_0$ **do**

  Compute log eigenvalues of tomographic data according to the chosen branch, $\hat{\lambda}_j = \lambda_j + 2\pi i m_j$

  for $t = 1$ **to** T **do**

    Get eigenvalues σ_j and eigenvectors $|\tilde{v}_j\rangle\rangle$ of current Lindbladian model $\tilde{L}_{t-1}$

    Match and project model eigenvectors $|\tilde{v}_j\rangle\rangle$ onto tomographic subspaces S_k

    Pair log-eigenvalues $\hat{\lambda}_{j'}$ with projected eigenvectors $\Pi_k|\tilde{v}_j\rangle\rangle$, by matching to the associated model eigenvalues σ_j, so as to construct approximate matrix logarithm $\bar{A}_m$

    Run convex optimization to project $\bar{A}_m$ back onto the set of Lindbladians, outputting an updated model $\tilde{L}_t$

  end

  Append $\tilde{L}_T$ to list of candidate solutions $\mathbb{L}$

end

return $L \in \mathbb{L}$ that minimizes $\|e^L - E\|$

Consider a case where L^{ideal} has a k-fold degenerate eigenspace Π that gets split into two (possibly still degenerate) eigenspaces Π_1 and Π_2 of L^* due to gate noise. When our algorithm computes a set of eigenvectors $v_1, \ldots, v_k$ for Π, there is no guarantee which vectors in Π the eigensolver will return. Indeed, L^{ideal} is indifferent to the choices of $v_1, \ldots, v_k$ as all of them lead to the same eigenspace projector. However, it is possible that $v_1, \ldots, v_k$ are all far away from Π_1 or Π_2. For example, consider $\Pi = \mathbb{C}^2$, $v_1 = |0\rangle$, $v_2 = |1\rangle$, $\Pi_1 = \text{span}\{|+\rangle\}$, and $\Pi_2 = \text{span}\{|-\rangle\}$. This will not cause trouble if $\|E - E^*\|$ is sufficiently small to allow us to choose β small enough so that the eigenvalues of E corresponding to Π_1 and Π_2 are treated as distinct. However, in the parameter regimes feasible today, it is likely that the algorithm would need to choose a large β, which results in the merging of Π_1 and Π_2, so that the matrix $\bar{A} \in \widetilde{\log}(E)$ constructed is not too far away from $\mathcal{L}$. In other words, choosing a large β blurs the goal of finding vectors close to Π_1 and Π_2. To address this issue, instead of initializing at precisely L^{ideal}, the algorithm tries to roughly guess the correct "directions" by starting at random perturbations $\tilde{L}^{\text{ideal}}$ of L^{ideal}. We empirically observe that for the 1-qubit case, a handful of uniformly random perturbed starts would suffice. However, for the 2-qubit case, the search space becomes too vast to cover with a small number of uniformly random perturbations. As a heuristic solution, we perturb diagonally

$$\tilde{L}^{\text{ideal}} = L^{\text{ideal}} + D \tag{5.27}$$

and diagonally "along the X-axis"

$$\tilde{L}^{\text{ideal}} = L^{\text{ideal}} + H^{\otimes 4} D H^{\otimes 4} \tag{5.28}$$

for random diagonal matrices D with small norms.

Due to the mismatch between (5.24) and (5.25) and the fact that $\log(E)$ (or for that matter $\widetilde{\log}(E)$) is non-convex, standard convergence guarantees for alternating projection methods do not apply to our algorithm. For the Markovian $E = E^*$ case, the arguments given in Section 5.6.4 directly translate to the Alternating Projections method when choosing the precision parameter $\beta = 0$. We leave proving stronger formal guarantees for our algorithm as future work.

Criteria for success

We have implemented our algorithm in software and tested extensively on synthetic data obtained using simulated quantum process tomography. We focus on the case where the gate noise is Markovian, so E^* corresponds to the ground truth channel while E differs from E^* solely due to statistical error. In this case, on the one hand, it is almost certain that

$$\min_{L \in \mathcal{L}} \|e^L - E\| < \|E^* - E\| \tag{5.29}$$

while on the other hand, it is pointless to search for a Lindbladian L whose objective value $\|e^L - E\|$ is much smaller than $\|E^* - E\|$. We introduce two criteria for the positive retrieval of Lindbladians in our testing:

- **Success 1**: The output Lindbladian L satisfies
$$\|e^L - E\| \leq \|E - E^*\|.$$
- **Success 2**: The output Lindbladian L satisfies
$$\|e^L - E^*\| \leq \|E - E^*\|.$$
This is a more stringent success criterion, where the output Lindbladian L closely fits the ground truth transfer matrix E^*.

We view passing Success 2 as a bonus since the algorithm is not intended to remove statistical error from E. Instead, we would realistically expect $\|e^L - E^*\| \approx \|E - E^*\|$ and passing Success 1 guarantees at a bare minimum that $\|e^L - E^*\| \leq \|e^L - E\| + \|E - E^*\| \leq 2\|E - E^*\|$. We tested our Alternating Projections method on 600 test cases split across 30 different gate and gate noise combinations, all under a realistic level of statistical error induced by 10^4 tomographic shots. Our algorithm achieved success rates of 100% and 61.4% with respect to Success 1 and Success 2 respectively. See Table 5.1 for more detail and Section 5.3 for additional test results.

5.2.2 The gate set Lindbladian fitting problem

Removing SPAM errors under ideal statistics

We first describe the gate set Lindbladian fitting problem in the theoretical setting where there is no statistical error and all gate noises are Markovian. In this case, the input to the gate set Lindbladian fitting problem consists of a set of $k + 1$ matrices

$$\{g^*, P_1^*, \ldots, P_k^*\} = \{A^* B^*, A^* E_1^* B^*, \ldots, A^* E_k^* B^*\} \tag{5.30}$$

where each $E_i^* = e^{L_i^*}$ is the transfer matrix of a Markovian channel representing a chosen unitary gate $E_i^{\text{ideal}} = e^{L_i^{\text{ideal}}}$ perturbed by some small gate noise, each row of A^* encodes a 2-outcome POVM, and each column of B^* is a vectorized density matrix. It is assumed that A^* and B^* originate from effecting some SPAM errors on some chosen ideal measurement (2-outcome POVMs) settings A^{ideal} and preparation (density matrices) settings B^{ideal} respectively. Note that the SPAM errors are not assumed to be Markovian. The Gram matrix $g^* = A^*B^*$ is assumed to be invertible. By left multiplying $(g^*)^{-1}$ to each P_i^*, we obtain a set of matrices

$$\{(B^*)^{-1}E_1^*B^*, \ldots, (B^*)^{-1}E_k^*B^*\}, \tag{5.31}$$

one for each gate in the gate set. A solution to the gate set Lindbladian fitting problem consists of, and hence the goal is to find, a physical B and Lindbladians $L_1, \ldots, L_k$ such that

$$B(g^*)^{-1}P_i^*B^{-1} = B(B^*)^{-1}E_i^*B^*B^{-1} = e^{L_i} \tag{5.32}$$

for every gate $i \in \{1, \ldots, k\}$ and $A = g^*B^{-1}$ is physical. We define A to be physical if every row of A encodes a 2-outcome POVM while B is physical if B is invertible and every column of B is a vectorized density matrix. Let $(B, L_1, \ldots, L_k)$ be a solution and for every gate i, define $E_i = e^{L_i}$. While it is clear that $(B^*, L_1^*, \ldots, L_k^*)$ is a solution, it is important to note that $(B, L_1, \ldots, L_k)$ need not be equal to $(B^*, L_1^*, \ldots, L_k^*)$. In fact, given the input

$$\{A^*B^*, A^*E_1^*B^*, \ldots, A^*E_k^*B^*\}, \tag{5.33}$$

it is impossible to distinguish whether the underlying physical process is generated by $(A^*, B^*, L_1^*, \ldots, L_k^*)$ or $(A, B, L_1, \ldots, L_k)$ since $AB = A^*B^*$ and for every gate $i \in \{1, \ldots, k\}$,

$$AE_iB = A(B(B^*)^{-1}E_i^*B^*B^{-1})B = A^*E_i^*B^*. \tag{5.34}$$

In other words, $(A^*, B^*, L_1^*, \ldots, L_k^*)$ and $(A, B, L_1, \ldots, L_k)$ generate the same input. Nevertheless, under the same weak noise assumptions that we made in Section 5.2.1, B^* is not far away from B^{ideal} and E_i^* is not far away from E_i^{ideal} for every gate i. Thus, we would prefer solutions closer to $(B^{\text{ideal}}, L_1^{\text{ideal}}, \ldots, L_k^{\text{ideal}})$.

The gate set Lindbladian fitting problem is precisely captured by the optimization problem of minimizing the objective function

$$f(B, L_1, \ldots, L_k) = \sum_{i=1}^{k} \left\| B(g^*)^{-1}P_i^*B^{-1} - e^{L_i} \right\| \tag{5.35}$$

subject to the constraints that $L_1, \ldots, L_k$ are Lindbladians, B is physical, and $A = g^*B^{-1}$ is physical. Clearly, $(B, L_1, \ldots, L_k)$ is a solution to the gate set Lindbladian fitting problem if and only if $(B, L_1, \ldots, L_k)$ is feasible and $f(B, L_1, \ldots, L_k) = 0$. Comparing to standard gauge optimization in GST, note that the objective function (5.35) is similar to (5.21), except for replacing each E_i^{ideal} with e^{L_i}, and that B^* *is* an optimal gauge for (5.35).

Accepting statistical error

Now consider the realistic scenario where the input matrices are not precisely

$$\{g^*, P_1^*, \ldots, P_k^*\} \tag{5.36}$$

but are some nearby matrices

$$\{\tilde{g}, \tilde{P}_1, \ldots, \tilde{P}_k\} \tag{5.37}$$

due to the inevitable statistical error in performing quantum process tomography. In this case, there is unlikely to be a physical B that makes $B\tilde{g}^{-1}\tilde{P}_i B^{-1}$ exactly Markovian for any gate i. Thus, the goal of the gate set Lindbladian fitting problem needs to be relaxed to accommodate statistical error. To this end, we relax the goal to finding a near-physical B such that $A = \tilde{g}B^{-1}$ is near-physical and $B\tilde{g}^{-1}\tilde{P}_i B^{-1}$ is close to being Markovian for every gate i. The amount of slack allowed for the physical constraints is an input parameter selected by the user and should be chosen based on the number of tomographic shots used to estimate the Gram matrix $\tilde{g}$. We model this problem by the optimization problem of minimizing

$$f_{\max}(B, L_1, \ldots, L_k) = \max_{i=1,\ldots,k} \|B\tilde{g}^{-1}\tilde{P}_i B^{-1} - e^{L_i}\| \tag{5.38}$$

subject to the constraints that $L_1, \ldots, L_k$ are Lindbladians, B is near-physical, and $A = \tilde{g}B^{-1}$ is near-physical. We minimize the max over the gate set in the definition of $f_{\max}$ as opposed to the sum in f to enforce the requirement of finding a single gauge B that simultaneously fits all gates in the gate set. Under the sum formulation, a gauge B that fits some of the gates in the gate set extremely well but the remaining gates poorly may have a small objective value. Such a B clearly does not explain the true physical process and hence is undesirable. Adding more gates to the gate set may alleviate this problem. However, in a real experiment, it is common for the experimenter to want to characterize only a particular gate, such as CNOT, but GST is performed on a gate set consisting of CNOT and other auxiliary gates chosen solely for the purpose of fitting the SPAM errors. Thus, it is also undesirable to require a large gate set. The max formulation in (5.38) tries to address this trade-off. Furthermore, since some gates could be noisier than others in the gate set, substituting max in the objective function of standard gauge optimization (see (5.21)) incorrectly incentivizes pulling gate errors in the poorest-fit gate into SPAM errors. However, the aforementioned problem does not appear in the formulation of (5.38) since the gate errors are self-consistently fitted jointly with SPAM errors.

Applying gradient-based local minimization algorithms to optimize (5.38) is tricky since the max function is not differentiable everywhere. One common workaround is to approximate the max function by a smooth approximation, and this is the approach we take. The LogSumExp (LSE) function is defined as

$$\mathrm{LSE}(x_1, \ldots, x_k) = \ln(e^{x_1} + \cdots + e^{x_k}), \tag{5.39}$$

and it is easy to verify that for every $t > 0$,

$$\max\{x_1, \ldots, x_k\} \leq \frac{1}{t}\mathrm{LSE}(tx_1, \ldots, tx_k) \leq \max\{x_1, \ldots, x_k\} + \frac{\ln(k)}{t}. \tag{5.40}$$

Moreover, the LSE function is monotonically increasing in all of its inputs. We fix the scaling factor t to be some constant that depends on k and the number of tomographic shots. Our final objective function for the gate set Lindbladian fitting problem is the following approximation of (5.38) based on the LSE function:

$$h(B, L_1, \ldots, L_k) = \frac{1}{t}\mathrm{LSE}(t\|B\tilde{g}^{-1}\tilde{P}_1 B^{-1} - e^{L_1}\|, \ldots, t\|B\tilde{g}^{-1}\tilde{P}_k B^{-1} - e^{L_k}\|). \tag{5.41}$$

The Gate Set Flip-Flop algorithm

The idea of the Gate Set Flip-Flop algorithm is to minimize (5.41) via local minimization by alternating between minimizing B with $L_1, \ldots, L_k$ fixed and minimizing $L_1, \ldots, L_k$ with B fixed; see Algorithm 3 for pseudocode. More specifically, when $L_1, \ldots, L_k$ are fixed, we minimize the (restricted) objective function over B subject to physical constraints on A and B using gradient-based methods. When B is fixed, minimizing $L_1, \ldots, L_k$ amounts to fitting a Lindbladian to each $E_i = B\tilde{g}^{-1}\tilde{P}_i B^{-1}$, which we can handle using either the Convex Solve or the Alternating Projections methods described in Section 5.2.1. Note that we adopt L_i at the current iteration of Gate Set Flip-Flop as the input best-guess Lindbladian for the next run of Alternating Projections. Since E_i being Markovian for all $i \in \{1, \ldots, k\}$ correspond to a solution with objective value zero, which is essentially impossible in the presence of statistical error, the E_i's will be non-Markovian throughout the algorithm.

Algorithm 3: Gate Set Flip-Flop

input : Tomographic measurement data $(\tilde{g}, \tilde{P}_1, \ldots, \tilde{P}_k)$ where $\tilde{g}$ is the measured Gram matrix and $\tilde{P}_i$ is the data for the i-th quantum process; initial-guess Lindbladians for each gate $(L_1^{\text{ideal}}, \ldots, L_k^{\text{ideal}})$ following the same indexing as for tomographic data; initial guess B^{ideal} for tomographic preparation settings; stopping condition S

output: Estimated Lindbladians $(L_1, \ldots, L_k)$, preparation settings B, and measurement settings A.

$B \leftarrow B^{\text{ideal}}$, $L_1 \leftarrow L_1^{\text{ideal}}$, $\ldots$, $L_k \leftarrow L_k^{\text{ideal}}$
while *Stopping condition S not met* **do**
    `// Here we optimize the Lindbladians first. This order can be swapped, as discussed in the main text.`
    for $i = 1$ **to** k **do**
        $E_i \leftarrow B\tilde{g}^{-1}\tilde{P}_i B^{-1}$
        $L_i \leftarrow$ `LindbladFit(`E_i, L_i`)`
        `// LindbladFit is a function that takes as input a tomographic channel estimate (and optionally a best-guess Lindbladian) and outputs an estimate for a generating Lindbladian.`
    end
    $B \leftarrow$ **minimize** $h(B, L_1, \ldots, L_k)$ **subject to** B and $A = gB^{-1}$ physical
    where h is given by (5.41)
end
return $B, L_1, \ldots, L_k$

The initial guess used to seed the optimization is $(B^{\text{ideal}}, L_1^{\text{ideal}}, \ldots, L_k^{\text{ideal}})$. This is a sensible choice because we seek solutions in the vicinity of $(B^{\text{ideal}}, L_1^{\text{ideal}}, \ldots, L_k^{\text{ideal}})$ in accordance with the weak noise assumption. In addition, we need to choose between starting with minimizing B or $L_1, \ldots, L_k$. Minimizing B first against the ideal gates E_i^{ideal} is similar to gauge optimization in standard GST, except with physical constraints on B and a different objective function. For each gate i, minimizing L_i first against $B^{\text{ideal}}\tilde{g}^{-1}\tilde{P}_i(B^{\text{ideal}})^{-1}$ resembles linear inversion process tomography where SPAM errors are assumed to be non-existent. The choice that is more preferable depends on whether we have prior knowledge about the

noise characteristics in the input. If we believe that state preparation errors are stronger than gate noises, then minimizing B first could perform the best. Conversely, if we have prior knowledge that state preparation errors are considerably weaker than gate noises, then minimizing the Lindbladians first could be preferable.

Recall the fact that the LSE function is monotonically increasing in each of its inputs. This implies that the objective values of the iterates generated by Gate Set Flip-Flop form a monotonically non-increasing sequence that is bounded from below. Thus, the sequence of objective values necessarily converges to a fixed point. However, by the nature of local minimization, the Gate Set Flip-Flop algorithm is not guaranteed to find an optimal solution.

The main difference between the gate set Lindbladian fitting problem and standard gauge optimization in GST (see (5.21)) lies in replacing each E_i^{ideal} with e^{L_i} in the objective function. Operationally speaking, during each iteration of Gate Set Flip-Flop, the gauge B is minimized with respect to a set of more accurate estimates of the noisy gates than the ideal gates E_i^{ideal}. This can improve the accuracy of the estimated SPAM errors encoded in B by reducing the amount of gate noise erroneously pulled into SPAM errors.

5.3 Synthetic testing of Alternating Projections and Convex Solve

We first describe how we generate our test cases. We choose a 2-qubit gate and Markovian gate noise combination, for example, CNOT with amplitude damping noise. This specifies the ideal and ground truth transfer matrices E^{ideal} and E^* respectively. We then feed E^* through simulated quantum process tomography to obtain a matrix $\tilde{E}$. Finally, $\tilde{E}$ is projected to CPTP to generate an input transfer matrix E. Multiple runs of process tomography on the same E^* will generate different $\tilde{E}$ and E instances since the instantiation of statistical error in each run will be different. Since the gate noise is Markovian, the difference between E and E^* is solely due to statistical error. In this section, all inputs are generated using 10^4 tomographic shots, which is also the setting we use for our tests on real quantum hardware. We emphasize that during synthetic testing, both E^{ideal} and E^* are chosen and known.

Before presenting aggregate test results, we first highlight the difference between Convex Solve and Alternating Projections using a CNOT gate example. For CNOT gates, it is possible that at the level of statistical error induced by 10^4 tomographic shots and under weak gate noises, the closest Lindbladian to every physically relevant $A \in \log(E) = \{A : e^A = E\}$ exponentiates to a channel very far away from E. Figures 5.2 to 5.4 present a synthetic noisy CNOT gate example where (1) the Lindbladian fitting problem can be solved exactly using Convex Solve when there is no statistical error in the input transfer matrix, (2) the Convex Solve method estimates none of the Hamiltonian, the γ, or the jump operator accurately when the input suffers from a realistic level of statistical error, and (3) Alternating Projections outputs a Lindbladian that estimates all of the Hamiltonian, the γ, and the jump operator accurately for the same statistically noisy input. For the Alternating Projections method to work for this example, it is necessary to choose the precision parameter $\beta > 0$ to merge all the eigenvalues of E close to being real negative into one degenerate eigenspace, hence projecting to an enlarged search space $\widetilde{\log}(E)$.

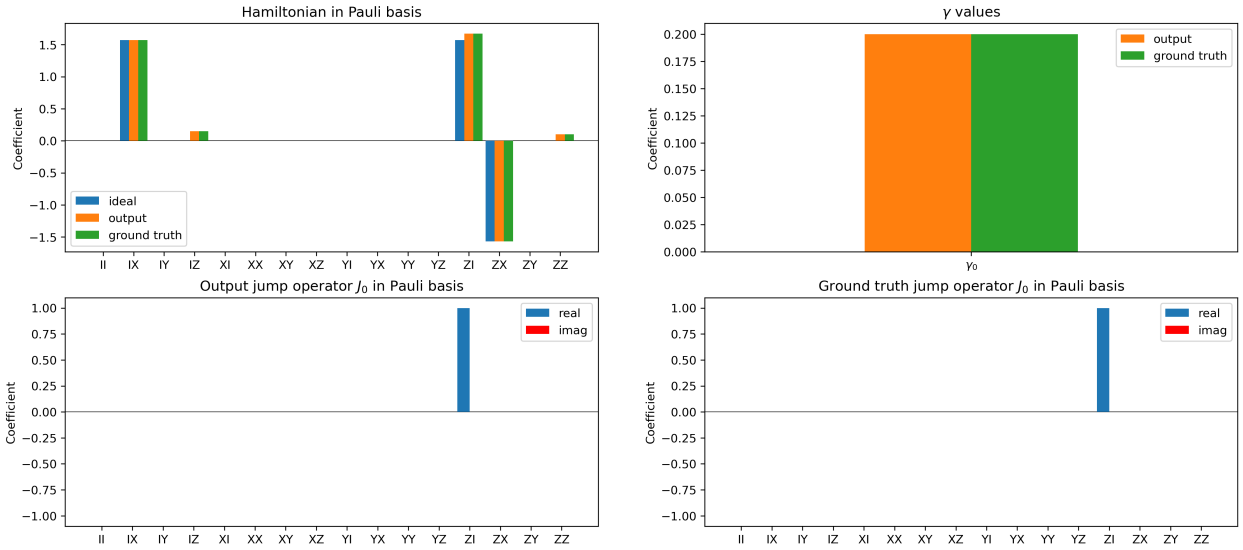


Figure 5.2: CNOT gate with coherent ZI, IZ, and ZZ errors and dissipative ZI error. E^{ideal} is the transfer matrix of the ideal noiseless CNOT gate, and $\|E - E^{\text{ideal}}\|$ quantifies the total amount of gate noise in the input. In this example, there is no statistical error, so $\|E - E^*\| = 0$. An exhaustive search over the branches of $\log(E)$ finds a Lindbladian L satisfying $\|e^L - E\| = \|e^L - E^*\| = 0$ up to numerical errors. The canonical decompositions are computed using the procedure described in Section 5.1.1.

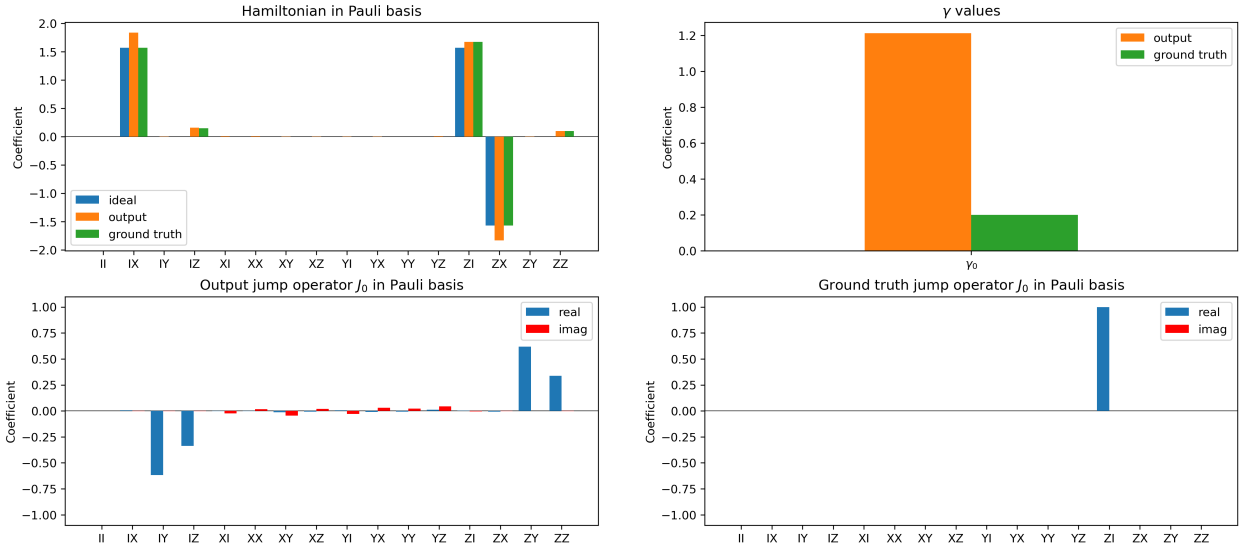


Figure 5.3: For the same noisy CNOT gate considered in Figure 5.2, at the level of statistical noise induced by 10^4 shots, an exhaustive search over the branches of $\log(E)$ is no longer able to find a Lindbladian that exponentiates closely to the input. The output Lindbladian L fails to capture accurately any of the Hamiltonian, the γ, or the jump operator. Notice that e^L is further away from the input E than the zero-knowledge estimate E^{ideal} where E^{ideal} is the transfer matrix of the ideal noiseless CNOT gate.

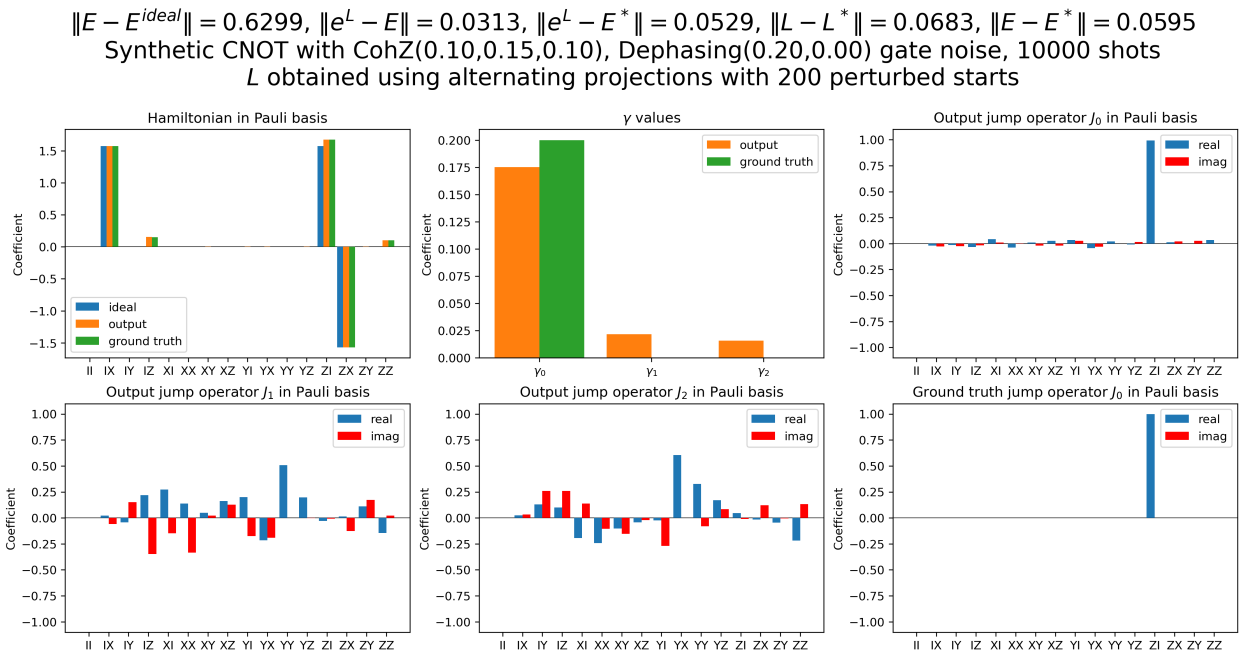


Figure 5.4: For the same input E considered in Figure 5.3, the alternating projections algorithm finds a Lindbladian that exponentiates very closely to the input. The output Lindbladian L captures all of the Hamiltonian, γ_0, and the jump operator J_0 accurately. While the algorithm erroneously predicts the jump operators J_1 and J_2 due to statistical error, their corresponding γ_1 and γ_2 coefficients are small.

For our synthetic benchmark, we use the two success criteria defined in Section 5.2.1. Recall that we say a test run passes Success 1 if $\|e^L - E\| \leq \|E - E^*\|$ holds, i.e., whether the output Lindbladian L closely fits the input transfer matrix E. Since $\|e^L - E\|$ is the objective function that the algorithm explicitly seeks to minimize, $\|E - E^*\|$ is the objective value attained by the ground truth Lindbladian L^* besides also quantifying the amount of statistical error in the input. Success 2, the second more stringent success criterion, checks for $\|e^L - E^*\| \leq \|E - E^*\|$, i.e., whether the output Lindbladian L closely fits the ground truth transfer matrix E^*. We view passing Success 2 as a bonus since the algorithm is not intended to remove statistical error from E—we would realistically expect $\|e^L - E^*\| \approx \|E - E^*\|$. Passing Success 1 guarantees at a bare minimum that $\|e^L - E^*\| \leq 2\|E - E^*\|$. See Table 5.1 for a table of aggregate results and see Figure 5.5 for a histogram of the distribution of the $\|e^L - E\|$ and $\|e^L - E^*\|$ values for the 20 instances of CNOT gate with coherent X and dephasing noise tested.

By examining the canonical decompositions of Lindbladians returned by our algorithm, we found cases where our algorithm's predicted jump operators visibly differed from the ground truth jump operators even when the output passed Success 2 (see Figure 5.6 for an example). Strictly speaking, this does not mean that the algorithm has failed since the algorithm's goal is to find *a* Lindbladian that approximately generates the input dynamics, and the output Lindbladian L does attain a small $\|e^L - E\|$ value.

Next, we investigate the weak noise assumption $\|L^* - L^{\text{ideal}}\| \leq c_2$ made near the beginning of Section 5.2.1. How small does c_2 need to be to ensure that the Alternating Projections method succeeds? To put the $\|L^* - L^{\text{ideal}}\|$ values in perspective, we can evaluate the average

Synthetic testing of the Alternating Projections method

Success 1, 2	Overrotation	Coh$X$	Bitflip	Coh$X$ Dephasing	Coh$X$ AmpDamp Bitflip
CNOT	20, 12	20, 20	20, 20	20, 0	20, 2
ISWAP	20, 8	20, 16	20, 17	20, 8	20, 2
$X \otimes H$	20, 13	20, 18	20, 19	20, 16	20, 20

Success 1, 2	AmpDamp	Coh$Z$	Dephasing	Coh$Z$ Bitflip	Coh$Z$ AmpDamp Dephasing
CNOT	20, 6	20, 20	20, 19	20, 5	20, 8
ISWAP	20, 10	20, 17	20, 20	20, 9	20, 20
$X \otimes H$	20, 20	20, 17	20, 16	20, 1	20, 6

Synthetic testing of the Convex Solve method

Success 1, 2	Overrotation	Coh$X$	Bitflip	Coh$X$ Dephasing	Coh$X$ AmpDamp Bitflip
$\sqrt{X} \otimes I$	20, 16	20, 18	20, 12	20, 20	20, 20
$T \otimes I$	20, 20	20, 20	20, 20	20, 20	20, 20
$I \otimes I$	/, /	20, 20	20, 20	20, 20	20, 20

Success 1, 2	AmpDamp	Coh$Z$	Dephasing	Coh$Z$ Bitflip	Coh$Z$ AmpDamp Dephasing
$\sqrt{X} \otimes I$	20, 20	20, 19	20, 20	20, 19	20, 20
$T \otimes I$	20, 20	20, 20	20, 20	20, 20	20, 20
$I \otimes I$	20, 20	20, 20	20, 20	20, 20	20, 20

Table 5.1: Each gate and noise model combination is tested on 20 input instances generated using simulated quantum process tomography with 10^4 shots. Each entry records the number of runs (out of 20) that passed the Success 1, 2 criteria $\|e^L - E\| \leq \|E - E^*\|$, $\|e^L - E^*\| \leq \|E - E^*\|$. When using Alternating Projections, the algorithm tries 500 perturbed starts. The Alternating Projections method consistently succeeds with respect to the Success 1 criterion for all the gate and noise combinations tested. The combined success rates are $600/600 = 100\%$ and $385/600 = 61.4\%$ for Success criteria 1 and 2 respectively. The Convex Solve method also consistently succeeds for all the cases tested, in particular the idling gate $I \otimes I$. The noise strengths $\|L^* - L^{\text{ideal}}\|$ range from 0.089 to 0.355.

gate fidelity between E^{ideal} (which corresponds to a unitary gate) and E^* using the formula

$$F_{\text{avg}}(E^{\text{ideal}}, E^*) = \frac{\frac{1}{d}\text{Tr}[(E^{\text{ideal}})^\dagger E^*] + 1}{d+1} \tag{5.42}$$

adapted from [127, 128] for transfer matrices where $d = 4$ is the dimension of the system. In Figure 5.7 (a), we consider running the Alternating Projections algorithm on CNOT with increasing strengths of coherent X, amplitude damping, and dephasing noise. The results indicate that our algorithm succeeds with respect to the Success 1 criterion for this specific gate and noise model combination as long as $\|L^* - L^{\text{ideal}}\| \leq 0.45$, at which point $F_{\text{avg}}(E^{\text{ideal}}, E^*)$ is roughly 94%. For this gate and noise model combination, the inputs E always have eigenvalues close to being real negative. Hence, we expect the Convex Solve method to fail on this test case, which is confirmed in Figure 5.7 (b). In Figure 5.7 (c) and (d), we repeat the same experiment on CNOT with overrotation and dephasing noise. The Alternating Projections method's performance is less sensitive to the noise strength in this case, and it succeeds even when $\|L^* - L^{\text{ideal}}\| = 1.8$, which translates to an average

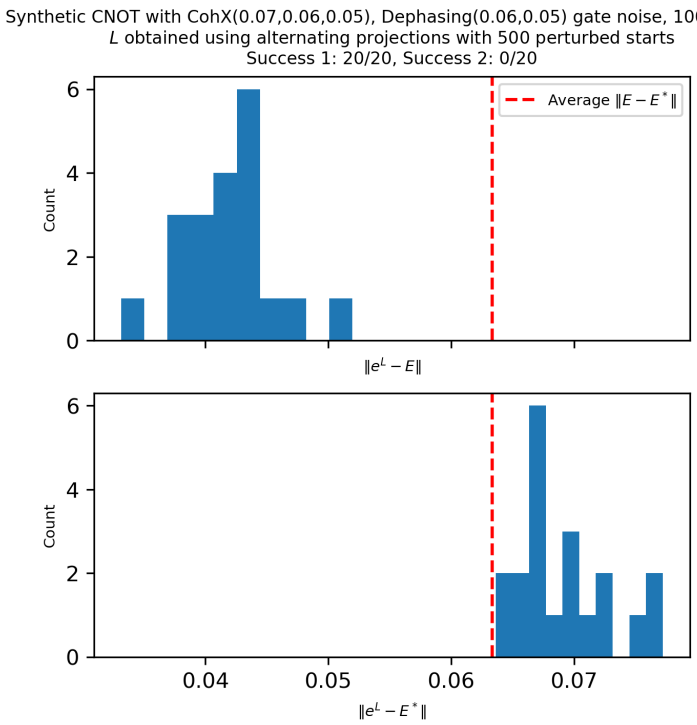


Figure 5.5: Distribution of the $\|e^L - E\|$ (top) and $\|e^L - E^*\|$ (bottom) values for the 20 instances of CNOT with coherent X and dephasing noise tested. The algorithm is overfitting statistical error to some extent.

gate fidelity of 83.5% between E^{ideal} and E^*. For CNOT with overrotation and dephasing noise, the -1 eigenvalues of E^{ideal} steadily rotate away from the real negative axis as the noise strength increases, and the Convex Solve method begins working when the average gate fidelity between E^{ideal} and E^* drops below roughly 96%.

Lastly, we examine to what extent does the Convex Solve method solve the Lindbladian fitting problem with statistical error when E^{ideal} corresponds to the identity gate $I \otimes I$. In Figure 5.8, we plot the results of running the Convex Solve method on noisy identity gates with statistical error with various noise models and increasing noise strengths where the average gate fidelities between E^{ideal} and E^* range from 99.9% to 44.6%. For each noise model and noise strength combination, we run the Convex Solve method on 20 input instances generated using the procedure described at the beginning of this section. The obtained $\|e^L - E\|$ values are averaged to calculate each data point. Our results indicate that the Convex Solve method succeeds consistently in the parameter regimes tested.

Our implementation runs significantly faster than the algorithm implemented in prior work [111]. For comparison, the same ISWAP analysis reported in [111], which took two weeks to run, can be performed in under two minutes using our implementation. The speedups mainly come from parallelizing the search over the branches of complex logarithm, skipping over clearly unphysical branches entirely, and while in a correct branch, the Alternating Projections method requiring fewer calls to the convex optimizer to find a good fitting Lindbladian. We use the SCS package [122] to solve the convex optimization problem of projecting a matrix to a Lindbladian. Our current implementation spends by far the most CPU time in the convex optimization step in the innermost loop.

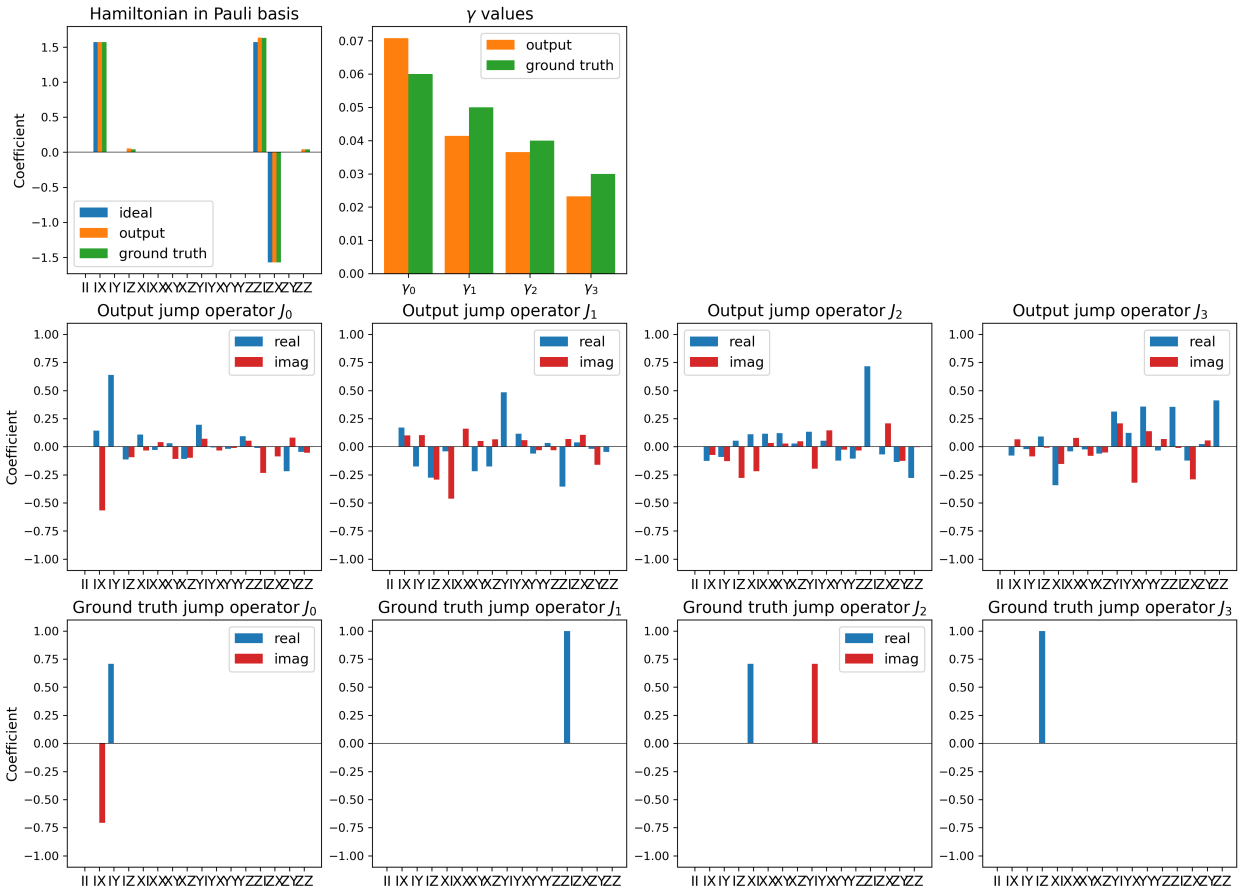


Figure 5.6: Canonical decomposition of a CNOT with coherent Z, amplitude damping, and dephasing noise instance. The algorithm succeeded on the input according to both of our success criteria. The Lindbladian found by the algorithm matches the ground truth Hamiltonian accurately and the γ values are well-estimated. However, the predicted jump operators are difficult to interpret.

5.4 Synthetic testing of Gate Set Flip-Flop

In this section, we report synthetic test results for the Gate Set Flip-Flop algorithm. We first describe how we generate our test cases. The gate set consists of the six 2-qubit gates $\{\text{CNOT}, \sqrt{X} \otimes I, I \otimes \sqrt{X}, T \otimes I, I \otimes T, \text{ISWAP}\}$. A SPAM error is chosen randomly from amplitude damping, bitflip, incoherent Y, and dephasing for each of the 16 state preparation settings and 16 measurement settings. For state preparation errors, the noise strength coefficient γ for each jump operator in (5.6) is drawn from the normal distribution with mean 0.07 and standard deviation 0.007. For measurement errors, the γ value for each jump operator is sampled randomly from the normal distribution with mean 0.12 and standard deviation 0.012. In effect, we try to model a quantum device whose measurements are noisier than state preparations. For each gate in the gate set, a gate noise with normally distributed noise strengths is randomly chosen from overrotation with bitflip, overrotation with dephasing, co-

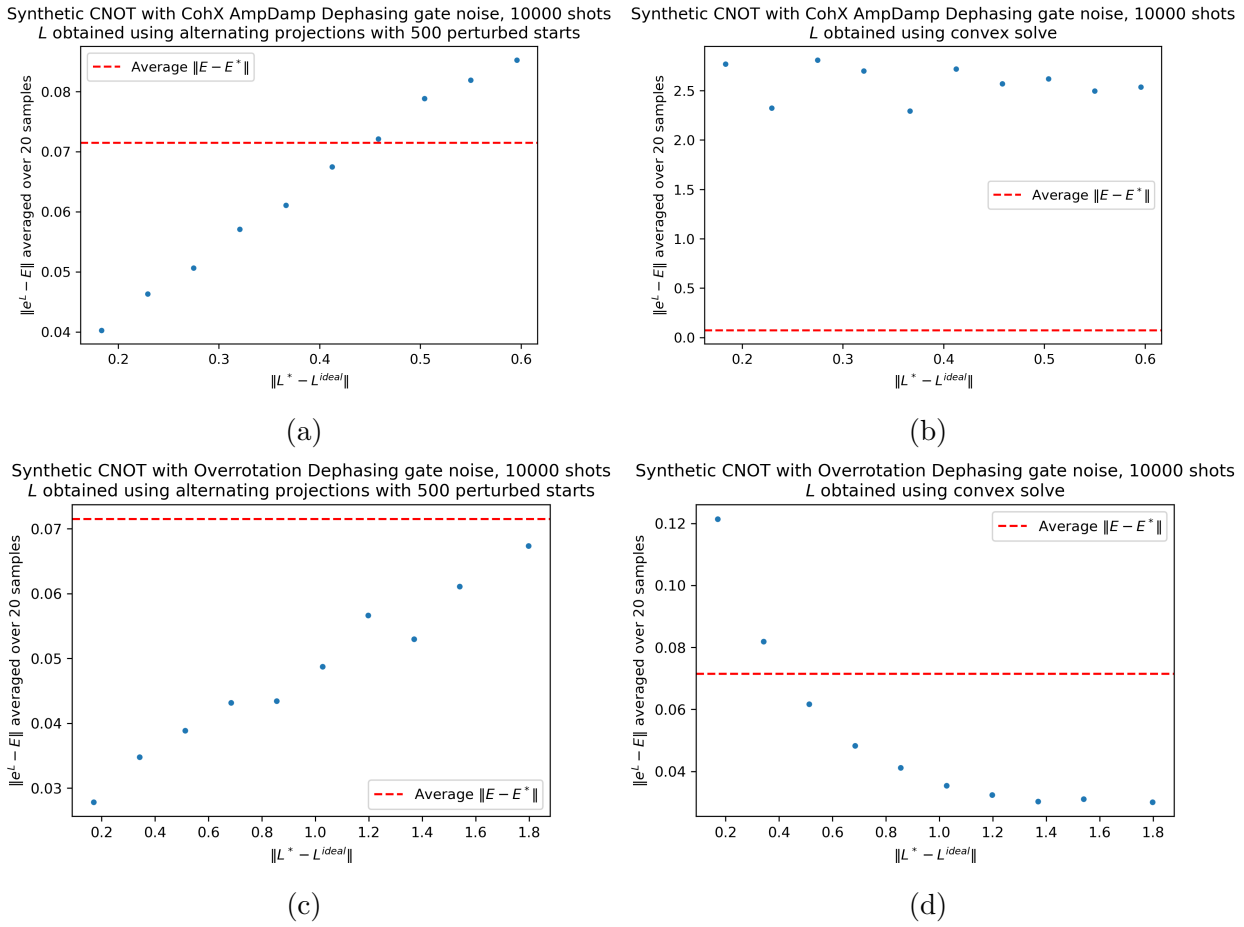


Figure 5.7: CNOT with increasing strengths of (a)(b) coherent X, amplitude damping, and dephasing gate noise and (c)(d) overrotation and dephasing gate noise. (a) The Alternating Projections method passes Success 1 for this test case until $\|L^* - L^{\text{ideal}}\|$ reaches roughly 0.45. (b) Even as the noise becomes stronger, the inputs E always have eigenvalues close to being real negative, and the Convex Solve method consistently fails. (c) For this noise model, the Alternating Projections method succeeds for the entire range of noise strengths tested. (d) As the noise becomes stronger, E stops having eigenvalues close to being real negative, and the Convex Solve method starts working.

herent Z with amplitude damping and dephasing, coherent Z with bitflip, and coherent X with dephasing. For every gate in the gate set, a simulated quantum process tomography experiment is carried out with 10^4 shots while the Gram matrix is estimated using 10^5 shots. We randomly generated 1000 test cases using the settings described above.

For each test case, the Gate Set Flip-Flop algorithm is run for three iterations on the objective function (5.41). We start by minimizing the Lindbladians first since the state preparation errors are weaker than the gate noises. The Lindbladian fitting problem is solved using the Alternating Projections method for the CNOT and ISWAP gates and the Convex Solve method for the other gates. Each Alternating Projections run tries 400 perturbed starts. The constrained minimization with respect to B is performed using the SLSQP algorithm [129] implemented in NLopt [130]. The physical constraints on the eigenvalues and trace of each column of B and the eigenvalues of each row of A are enforced with a slack of 0.001.

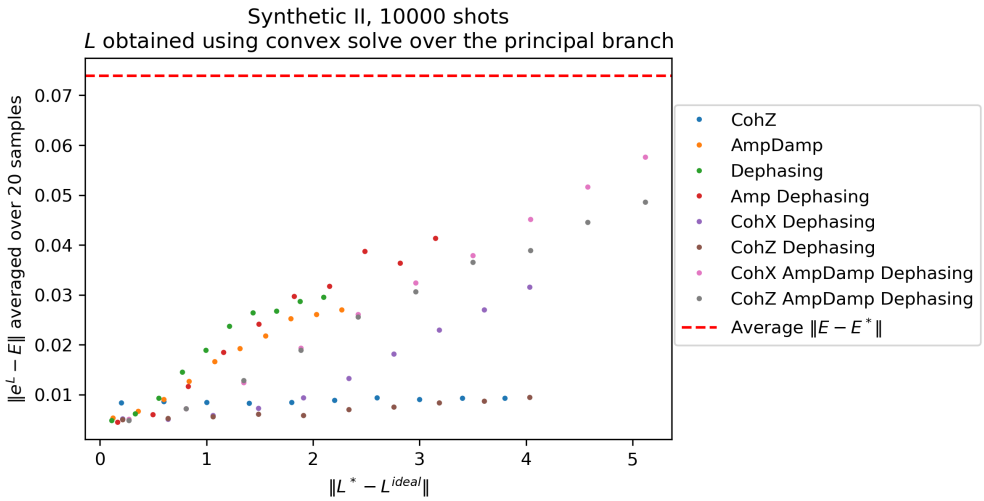


Figure 5.8: Each noise model is tested with 10 levels of increasing noise strengths. Each data point corresponds to the average $\|e^L - E\|$ value of 20 input instances. Across all the data points, the average gate fidelities between E^{ideal} and E^* range from 99.9% to 44.6%.

Analogous to the criterion Success 1 defined in Section 5.2.1, we consider a run of the Gate Set Flip-Flop algorithm to be successful if the algorithm's output $(B, L_1, \ldots, L_k)$ satisfies $f_{\max}(B, L_1, \ldots, L_k) \leq f_{\max}(B^*, L_1^*, \ldots, L_k^*)$ where B^* encodes the ground truth state preparation settings and $L_1^*, \ldots, L_k^*$ are the ground truth Lindbladians for the noisy gates. In Figure 5.9, we plot the distributions of $f_{\max}(B, L_1, \ldots, L_k)$ versus $f_{\max}(B^*, L_1^*, \ldots, L_k^*)$ values for the 1000 randomly generated test cases after the zeroth, first, second, and third iterations. Note that $f_{\max}(B, L_1, \ldots, L_k)$ after the zeroth iteration is just $f_{\max}(B^{\text{ideal}}, L_1^{\text{ideal}}, \ldots, L_k^{\text{ideal}})$. Each flip-flop iteration consists of one round of gauge optimization and one round of Lindbladian fitting. We observe that our Flip-Flop algorithm succeeds consistently after three iterations. It failed on just a handful of test cases because the slack for the physical constraints was too tight for these inputs, causing the SLSQP optimizer to get stuck at B^{ideal}. Raising the slack to a larger value such as 0.005 resolves the issue.

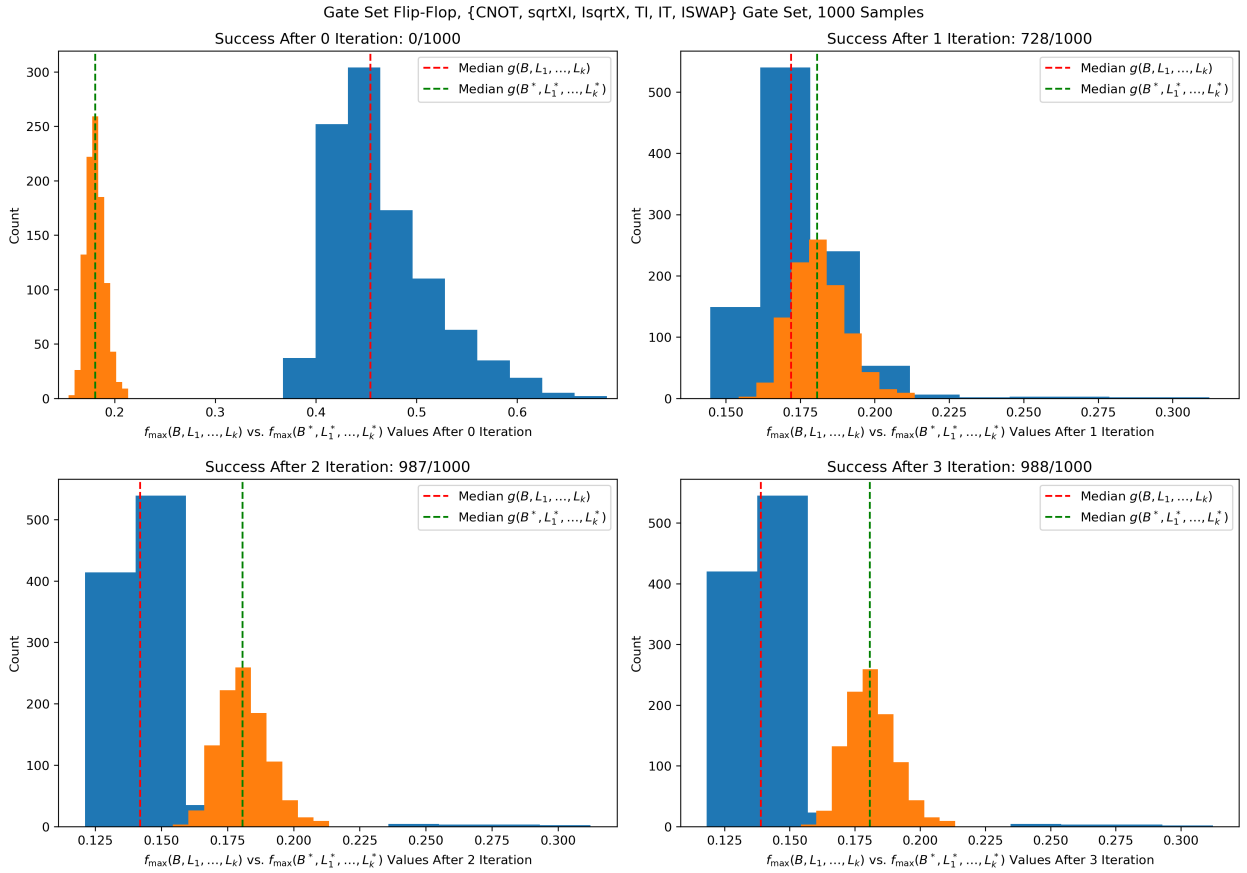


Figure 5.9: Synthetic test results for the Gate Set Flip-Flop algorithm on 1000 randomly generated test cases. The blue histograms plot the $f_{\max}(B, L_1, \ldots, L_k)$ values while the orange histograms plot the $f_{\max}(B^*, L_1^*, \ldots, L_k^*)$ values. A run of the algorithm is deemed successful if $f_{\max}(B, L_1, \ldots, L_k) \leq f_{\max}(B^*, L_1^*, \ldots, L_k^*)$.

5.5 Analysis of tomographic data from superconducting qubit hardware

In this section, we present example results produced by our algorithms when run on data obtained from real quantum computing hardware. The tomography circuits were executed on devices available via the cloud-based IBM Quantum platform [115] using the *Qiskit* software development kit [131]. Before describing the demonstration in detail, we note that in general, there can be several reasons why noisy gates might not be exactly generated by a time-independent Lindbladian. For one, the process may be more accurately modelled by a time-dependent Lindbladian, since some gates, particularly those involving 2-qubit interactions, may be implemented via a composite pulse sequence [132]. Even for simple pulses, the amplitude must be ramped on and off in some finite time, so the true Lindbladian of the process must be time-dependent. Also, the drift in the relative strength of noise components can yield apparent non-Markovianity in tomographic data, since this can lead to a distortion in the eigenvalue structure of the process. It has also been observed [114] that charge-parity noise occurring on the microsecond timescale can manifest as apparent non-Markovianity.

Finally, there may of course be genuine non-Markovian noise effects, for example, if there is an unwanted coupling to a defect acting as a two-level system [133, 134]. Nevertheless, a noiseless unitary operation is always compatible with a time-independent Lindbladian, so in the case where noise is relatively weak, an effective time-independent Lindbladian may be useful as a model for quantitatively and qualitatively understanding the noise profile for a given target process. Moreover, if our algorithm does not return a Lindbladian that exponentiates close to the tomographic data, it provides evidence of plausible non-Markovian noise effects present in the device.

In Section 5.5.1, we demonstrate a typical gate set analysis on a 7-qubit device. In particular, we characterize the continuously parameterized R_{ZX} pulse gate and find consistency between the Lindbladians estimated independently for different repetitions of the gate. We find the R_{ZX} to be both less noisy and closer to a time-independent Markovian process when compared to 2-qubit rotations decomposed in terms of CNOT gates and single-qubit rotations. In the following two subsections, we demonstrate how the protocol can be used to investigate cross-talk in different scenarios. In Section 5.5.2, we characterize a CNOT gate operated in isolation while the rest of the 27-qubit device remains idle and compare this with the same gate applied in parallel, finding that parallel application induces quantitative and qualitative changes to both the Hamiltonian and dissipative parts of the Lindbladian model. In particular, there is an apparent increase in ZX overrotation, the rate of dissipative evolution increases, and the Pauli decomposition of the dominant jump operator changes. Finally, Section 5.5.3 shows the application of our method to interrogate idling processes on pairs of qubits on the 7-qubit processor. Our results suggest that qubit proximity and activity elsewhere on the device both play a part in determining the effective noise model.

All experimental data are available at https://doi.org/10.5281/zenodo.18759706.

5.5.1 Two-qubit gate analysis on *ibm_perth*

We first show results obtained from the 7-qubit device known as *ibm_perth*. Gate set tomography was carried out for the set

$$\mathcal{G} = \{\text{CNOT}, R_{ZX}(0.5), R_{ZX}(0.5)^3, R_{ZX}(0.5)^5, \tag{5.43}$$

$$R_{ZZ}(0.5), R_{ZZ}(0.5)^3, R_{ZZ}(0.5)^5, T \otimes I, I \otimes T, \sqrt{X} \otimes I, I \otimes \sqrt{X}\} \tag{5.44}$$

implemented on the connected qubits labelled 3 and 5 using IBM's indexing (see Figure 5.19). As suggested above, there are different factors explaining why we might expect the ground truth to depart from a time-independent Markovian process, and we exemplify this with this gate set. On this device, the CNOT gate was implemented as an echoed cross-resonance gate [132, 135, 136], composed with shorter single-qubit pulses—in effect an $R_{ZX}(\pi/2)$ gate with an internal dynamical decoupling sequence to reduce coherent error and uncontrolled couplings with other subsystems, where

$$R_{ZX}(\theta) = \exp\left(-i\frac{\theta}{2} Z \otimes X\right). \tag{5.45}$$

The CNOT was provided as a native gate on *ibm_perth*, meaning that the device has been calibrated to maximize the fidelity of this gate. However, *Qiskit* enables the implementation

of custom pulse sequences, and an R_{ZX} gate can be implemented for arbitrary angles by extracting the pulse sequence calibrated for the $R_{ZX}(\theta)$ rotation underlying the CNOT gate and modifying the pulse duration (or amplitude) accordingly. In the gate set $\mathcal{G}$, we contrast this with $R_{ZZ}(\theta) = \exp(-i(\theta/2)Z \otimes Z)$ gates, which we decompose as a Z-rotation sandwiched by two CNOT gates. The single-qubit $\sqrt{X}$ and T gates are implemented by a simple pulse and a change of frame respectively, and are mainly included in the set to assist with gauge optimization. In each case, the noiseless process is unitary and can therefore be generated by a time-independent Lindbladian with no dissipative part. In reality, the tomographed channel could depart from this model to varying degrees depending on the gate. We may consider this a stress test of our algorithms, as it can be the case that there is no time-independent Lindbladian model that exponentiates close to the data.

For our tomographically complete set of initial states, we prepare all 2-qubit tensor products of four of the single-qubit Pauli eigenstates $\{|0\rangle, |1\rangle, |+\rangle, |+_i\rangle\}$, defined in the ideal case as

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad |+\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |+_i\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}. \tag{5.46}$$

On the IBM devices studied, the standard fiducial state on all qubits is $|0\rangle$, prepared via a non-unitary process. The other states in our set can then be prepared by a sequence of single-qubit gates. For our measurement settings, we measure all 2-qubit Pauli observables and take our measurement operators to be the projector onto the $+1$ eigenspace of each Pauli operator. The standard measurement on IBM devices is the Z measurement, so to measure other Pauli operators, we use single-qubit gates to rotate into the appropriate frame. To collect the tomographic data, we parallelize Pauli measurements, running 144 independent circuits per gate in the set $\mathcal{G}$ at 10^4 shots per circuit. All circuits are shuffled together in a random order to average out any drift and avoid correlations with real time within the data. After initial processing, the data for the $R_{ZZ}(0.5)^5$ process was found to have an eigenvalue structure incompatible with a time-independent Markovian process so was excluded from subsequent analysis. No further error mitigation is applied to the data, and the raw tomographic matrices are taken as input for the Gate Set Flip-Flop algorithm.

In Section 5.6.4, we have proved that in the absence of statistical error, when gate noise is sufficiently weak, and the ideal Lindbladian has no eigenvalues close to the real negative axis, a good solution can always be found by Convex Solve in the principal branch alone. This principle was also borne out in our numerical testing of synthetic data with simulated shot noise in Section 5.3 (see Figure 5.8). To illustrate that this principle also works on real data, during the Lindbladian fitting step, we use this simple method for all gates except CNOT and $R_{ZX}(0.5)^5 = R_{ZX}(2.5)$. The latter two gates respectively have -1 eigenvalues and eigenvalues close to the real negative axis, so for these we use the Alternating Projections method.

Since the ideal process E^{ideal} for each gate is unitary, we can use the formula (5.42) to evaluate the average gate fidelity for the estimated noisy Markovian process at each iteration. In Figure 5.10, we show figures of merit for the full gate set (excluding $R_{ZZ}(0.5)^5$) over iterations of the flip-flop algorithm. In this case, Gate Set Flip-Flop was initialized by first running the Lindbladian fit before any gauge optimization. The solution is seen to be well converged in 8 iterations. In Table 5.2 (a), we compare the distances between the data

Process	$\|E - E^{\text{ideal}}\|$	$\|E - \exp(L)\|$
CNOT	0.42	0.33
$R_{ZX}(0.5)$	0.31	0.25
$R_{ZX}(0.5)^3$	0.51	0.30
$R_{ZX}(0.5)^5$	0.66	0.36
$R_{ZZ}(0.5)$	0.44	0.28
$R_{ZZ}(0.5)^3$	0.79	0.34
$T \otimes I$	0.25	0.22
$I \otimes T$	0.29	0.26
$\sqrt{X} \otimes I$	0.38	0.31
$I \otimes \sqrt{X}$	0.34	0.29

(a)

j	k	$\|\frac{L_j}{j} - \frac{L_k}{k}\|$	$\|\exp(\frac{L_j}{j}) - \exp(\frac{L_k}{k})\|$
1	3	0.097	0.094
1	5	0.12	0.12
3	5	0.081	0.079

(b)

Table 5.2: (a) Distances between the tomographic data and corresponding ideal transfer matrices E^{ideal} and estimated Markovian channels $\exp(L)$, respectively, from gate set analysis on *ibm_perth* qubits 3 and 5. For each process P^*, the tomographically estimated transfer matrix estimate $E = B(g^*)^{-1}P^*B^{-1}$ is evaluated in the final gauge B obtained from the flip-flop algorithm. We find that in all cases, the Markovian channel estimate gives a better fit to the data than assuming the gate is ideal. (b) Distances between time-normalized Lindbladians and the exponentiated transfer matrices, where L_k is the estimated Lindbladian for the noisy implementation of the pulse gate sequence $R_{ZX}(0.5)^k$ on *ibm_perth*.

E, as evaluated in the final gauge, and the ideal transfer matrices E^{ideal} and the Markovian estimate $\exp(L)$ respectively. In all cases, the estimated Lindbladian gives at least as good a fit as assuming the ideal, and we find the fit to be significantly improved for processes where we expect the most gate noise, namely those involving repeated application of 2-qubit gates. In contrast, the single-qubit gates only show a slightly improved fit, in keeping with the belief that single-qubit gates have good fidelity on modern superconducting-qubit devices. The remaining distance $\|E - \exp(L)\|$ between the data and the estimated process is likely at least in part due to statistical noise and residual SPAM errors that were not filtered by the algorithm.

Note that for an ideal gate we would have $R_{ZX}(0.5)^k = R_{ZX}(k/2)$, and in the channel picture $\exp(L)^k = \exp(kL)$ for any time-independent Lindbladian L. By normalizing the Lindbladians with respect to the number of applications of the gate, we can test the consistency of the estimated Markovian process for each sequence. In Table 5.2 (b), we show the distances between the time-normalized Lindbladians for each pair of processes, as well as the distances between the exponentiated transfer matrices. Figures 5.11 to 5.13 show the estimated Lindbladian decomposition for the $R_{ZX}(0.5)^k$ process for $k \in \{1, 3, 5\}$, comparing initial and final estimates. We note that the Hamiltonians in Figures 5.11 to 5.13 consistently show an overrotation about the ZX axis.

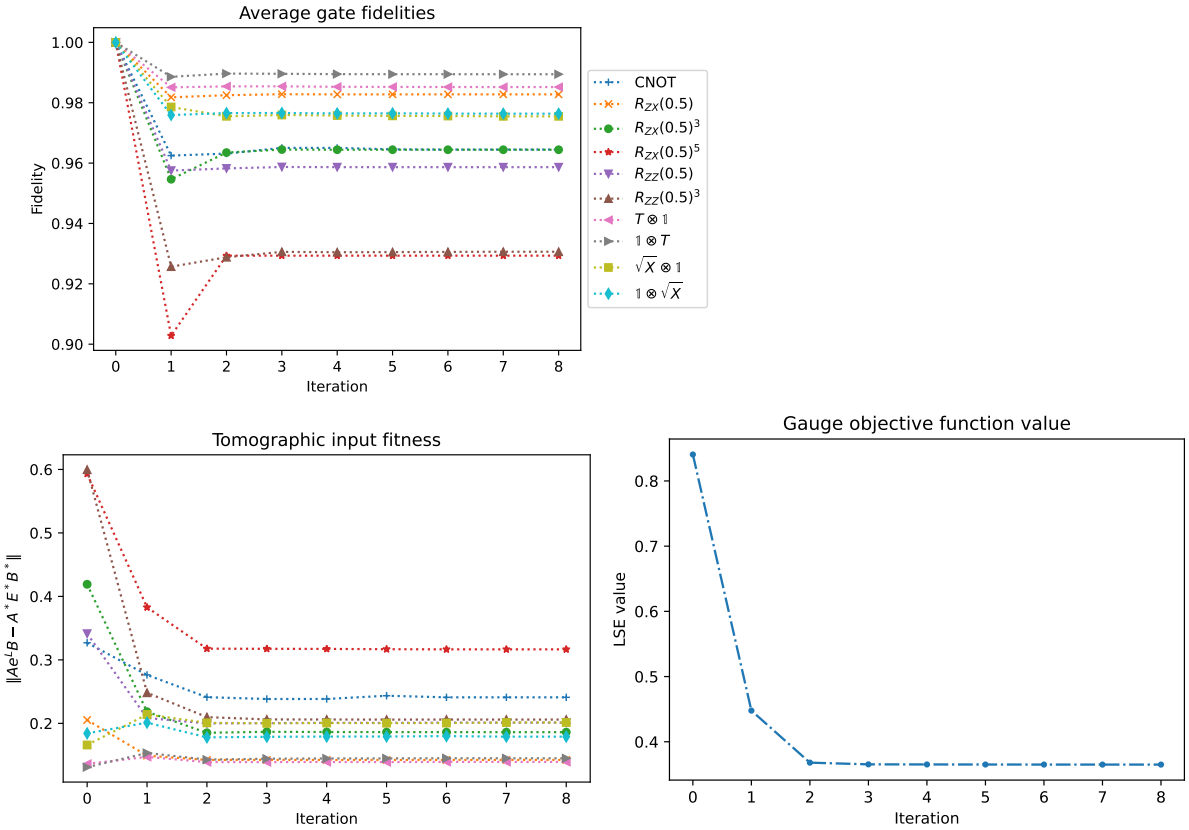


Figure 5.10: Gate Set Flip-Flop figures of merit for data collected from *ibm_perth*. (Top panel) The average gate fidelity between the ideal gate and the estimated Markovian process at each iteration, computed using (5.42). (Bottom left panel) Distance between raw tomographic data and the equivalent data for the estimated Markovian process assuming the current gauge. (Bottom right panel) Gauge objective function as defined in (5.41).

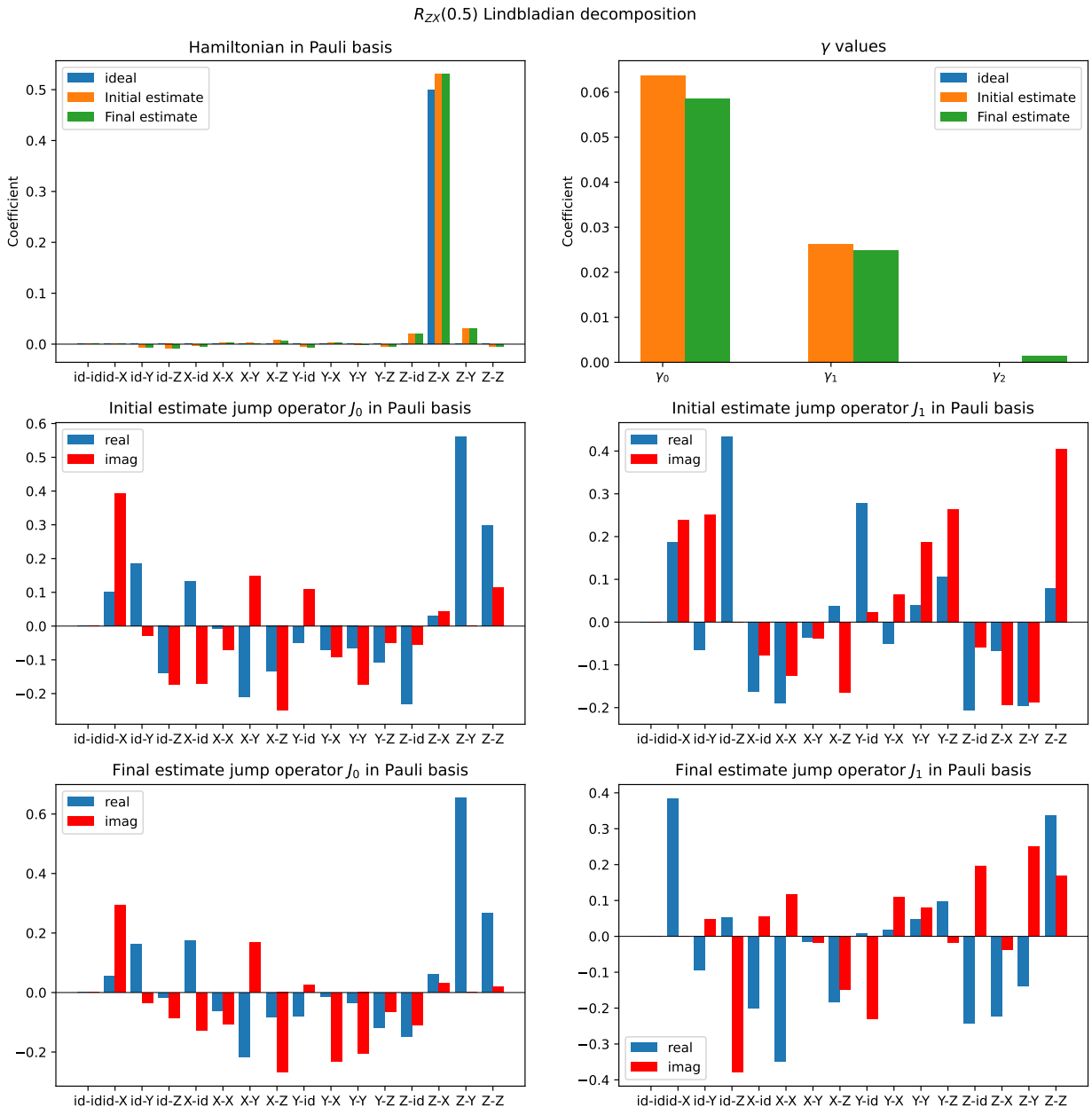


Figure 5.11: Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth*. We compare the ideal unitary process with the initial estimate prior to any gauge optimization, and the final estimate after convergence of the Gate Set Flip-Flop algorithm. Pauli decompositions are given only for the two dominant jump operators.

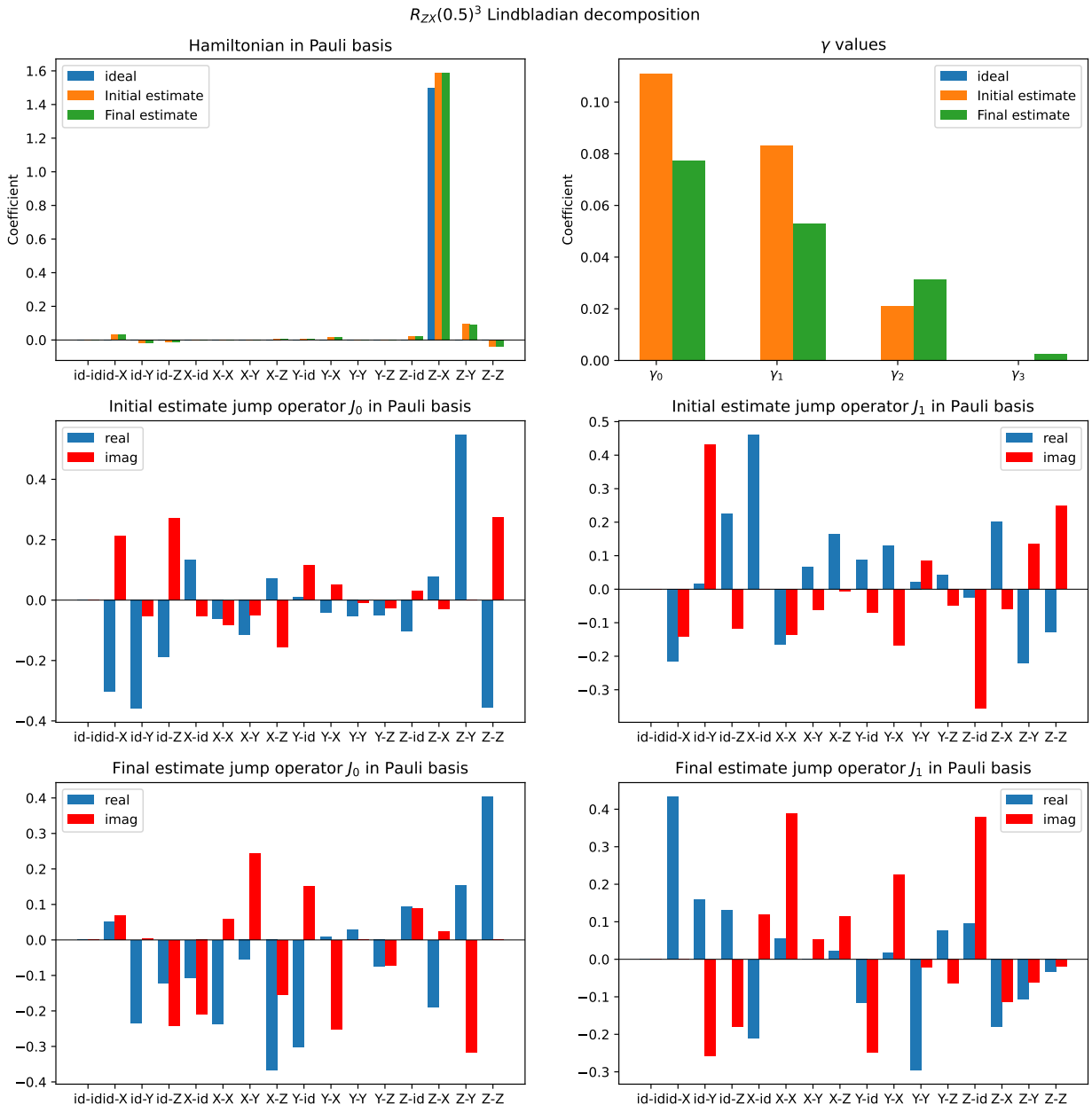


Figure 5.12: Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)^3 = R_{ZX}(1.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth*. The ideal process is compared with the initial estimate before gauge optimization, and the final fitted Lindbladian after convergence of Gate Set Flip-Flop. Pauli decompositions are given only for the two dominant jump operators.

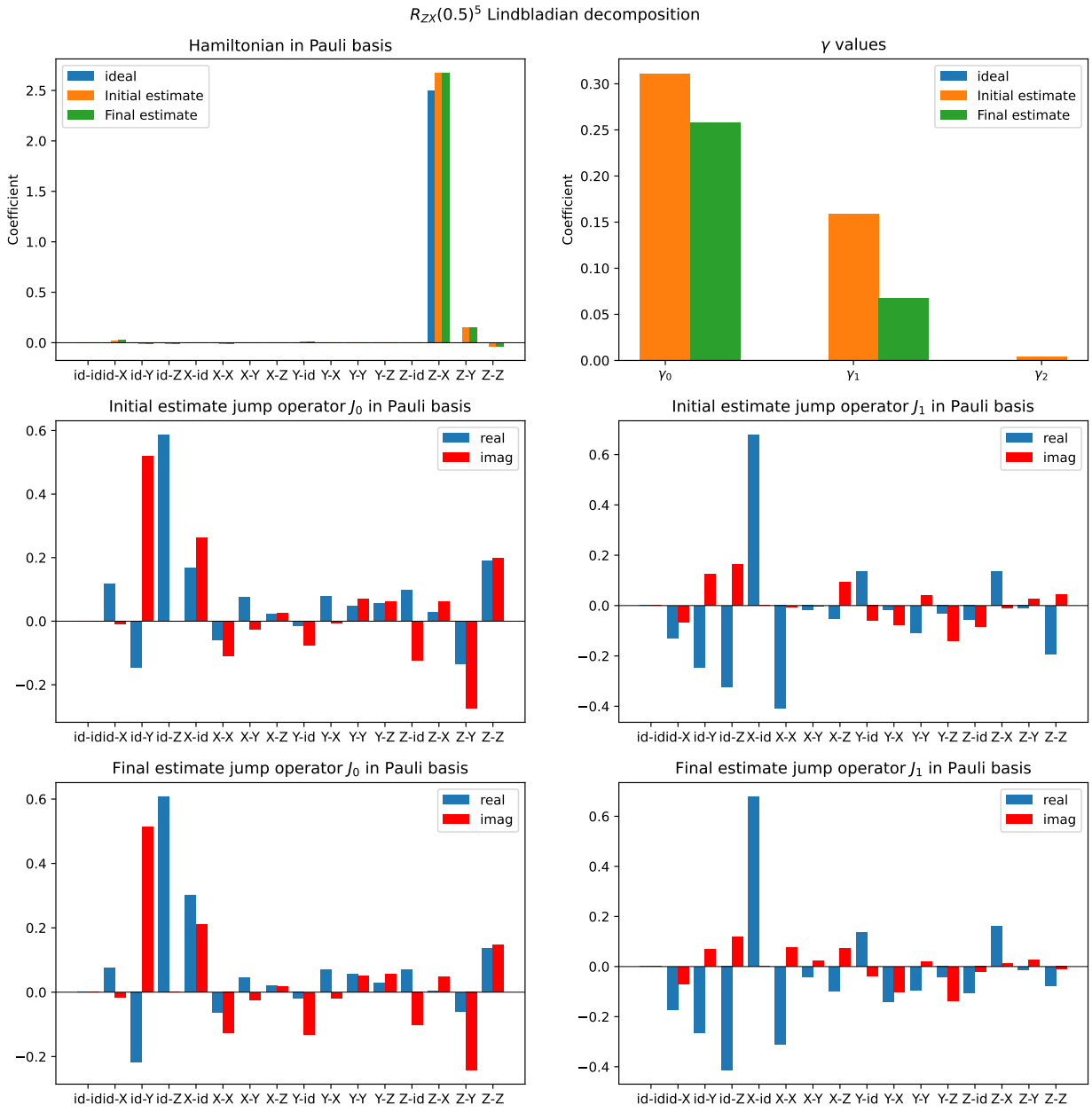


Figure 5.13: Canonical decomposition of the fitted Lindbladian for the $R_{ZX}(0.5)^5 = R_{ZX}(2.5)$ process applied to qubit pair $(3, 5)$ on *ibm_perth*. We compare the ideal unitary process with fitted estimates before and after running the Gate Set Flip-Flop routine until converged. Pauli decompositions are given only for the two dominant jump operators.

5.5.2 Parallel gate cross-talk experiment on *ibm_cairo*

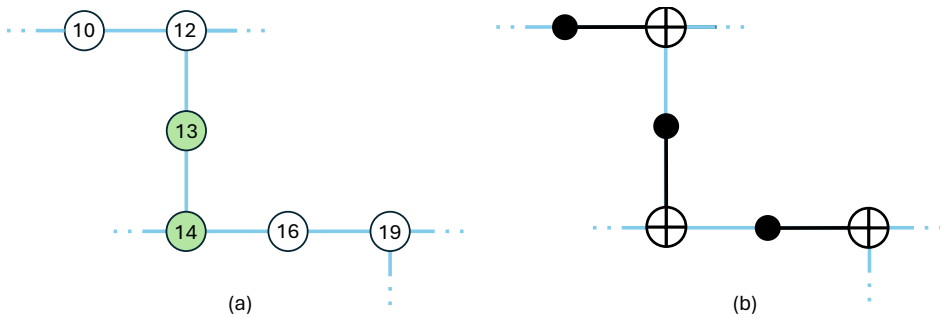


Figure 5.14: (a) Partial qubit layout on *ibm_cairo*. Numbered vertices show qubit locations, edges show connections where CNOT gates are available. Dashed lines show connections to qubits not used in the experiment. The qubit pair targeted for gate set tomography is shown shaded green. (b) Placement of CNOT gates during tomography of the $\mathrm{CNOT}_{\mathrm{parallel}}$ process.

For this second set of data, we focus on the analysis of a particular gate and consider its susceptibility to cross-talk. We perform gate set tomography of the set

$$\mathcal{G} = \{\mathrm{CNOT}_{\mathrm{single}}, \mathrm{CNOT}_{\mathrm{parallel}}, T \otimes I, I \otimes T, \sqrt{X} \otimes I, I \otimes \sqrt{X}\} \tag{5.47}$$

where $\mathrm{CNOT}_{\mathrm{single}}$ and $\mathrm{CNOT}_{\mathrm{parallel}}$ are CNOT gates executed either in isolation or in parallel with CNOT gates executed on other qubit pairs. The data was collected from IBM's 27-qubit device *ibm_cairo*. The topology of the 6-qubit patch used for the tomography is shown in Figure 5.14, along with the placement of 2-qubit gates for the $\mathrm{CNOT}_{\mathrm{parallel}}$ process. As in the previous demonstration, the single-qubit gates are provided to aid gauge optimization, shots were taken at a rate of 10^4 per setting, circuits were executed in random order, and readout mitigation was not applied to the data before running the algorithm. For this gate set, Convex Solve on the principal branch was again used to fit Lindbladians for all processes except the two CNOT processes, for which Alternating Projections was used.

We illustrate the convergence of the solution in Figure 5.15. Again, we observe good convergence within around 10 iterations. Comparing with Figure 5.10, we see that we are able to obtain a closer fit to a set of time-independent Lindbladians, and the estimated Markovian processes generally have better fidelity. Final distances from the data are shown in Table 5.3, where again we see only minimal improvement in fit for single-qubit gates, which are expected to already be close to ideal, with more significant improvement for the noisy 2-qubit processes. We note that the difference between the tomographic input fitness for the converged solutions for single and parallel CNOTs is relatively small, suggesting that in this scenario, cross-talk does not induce a dramatic increase in apparent non-Markovianity. On the other hand, we see that the algorithm suggests a reduced average gate fidelity for the parallel CNOT, and we observe qualitative differences in the estimated Lindbladian decompositions (Figure 5.16). In particular, we see increased overrotation about the $Z \otimes I$ axis for the parallel case and an increase in the strength of dissipative noise.

For completeness, in Figures 5.17 and 5.18, we compare the initial and final estimates for each implementation of the CNOT. We notice that in both cases, the Hamiltonian part of the Lindbladian remains relatively stable, but the final choice of gauge output by the flip-flop algorithm yields an apparent reduction in the strength of the dissipative part.

Process	$\|E - E^{\text{ideal}}\|$	$\|E - \exp(L)\|$
$\text{CNOT}_{\text{single}}$	0.23	0.18
$T \otimes I$	0.19	0.18
$I \otimes T$	0.20	0.17
$\text{CNOT}_{\text{parallel}}$	0.37	0.17
$\sqrt{X} \otimes I$	0.20	0.17
$I \otimes \sqrt{X}$	0.19	0.17

Table 5.3: Distances between the tomographic channel estimates and corresponding ideal transfer matrices E^{ideal} and estimated Markovian channels $\exp(L)$, respectively, from gate cross-talk analysis on *ibm_cairo*. For each process P^*, the tomographically estimated transfer matrix estimate $E = B(g^*)^{-1}P^*B^{-1}$ is evaluated in the final gauge B obtained from the flip-flop algorithm.

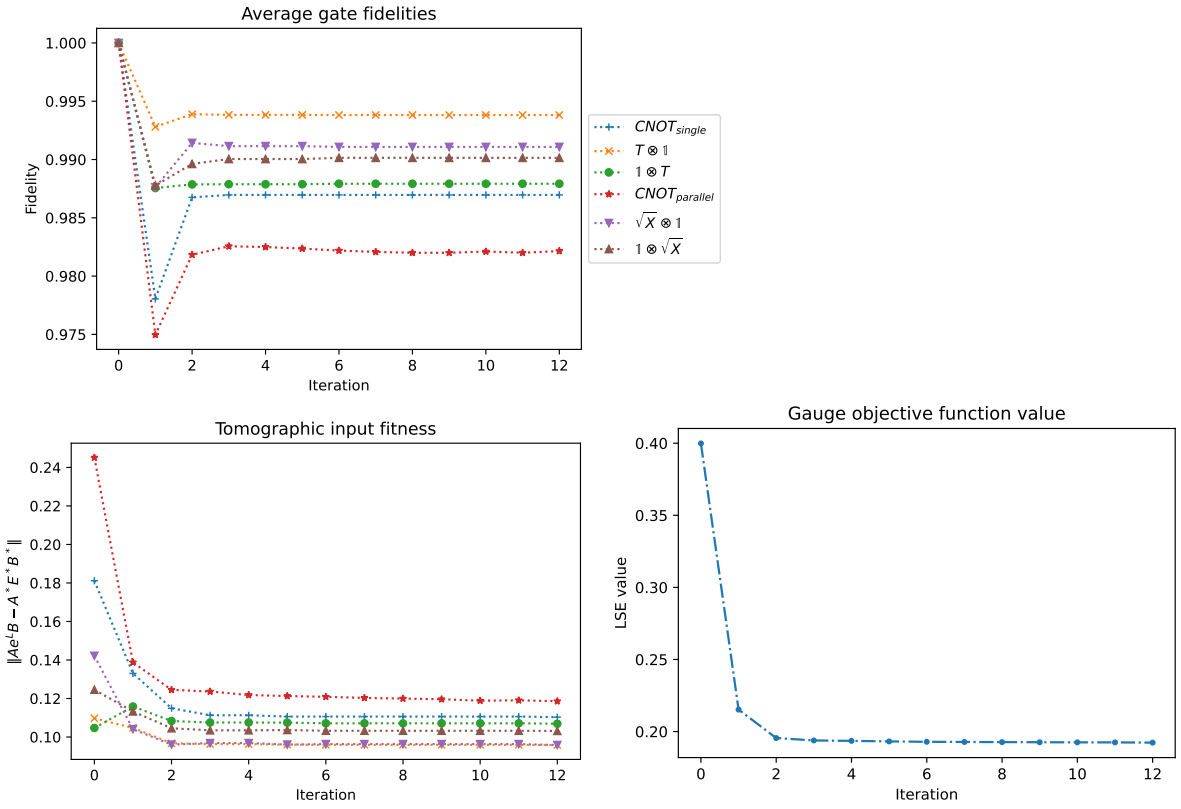


Figure 5.15: Gate Set Flip-Flop figures of merit for data collected from *ibm_cairo*, tomographing qubits 13 and 14. (Top left) The average gate fidelity between the ideal gate and the estimated Markovian process at each iteration, computed using (5.42). (Bottom left) Distance between raw tomographic data and the equivalent data for the estimated Markovian process assuming the current gauge. (Bottom right) Gauge objective function as defined in (5.41).

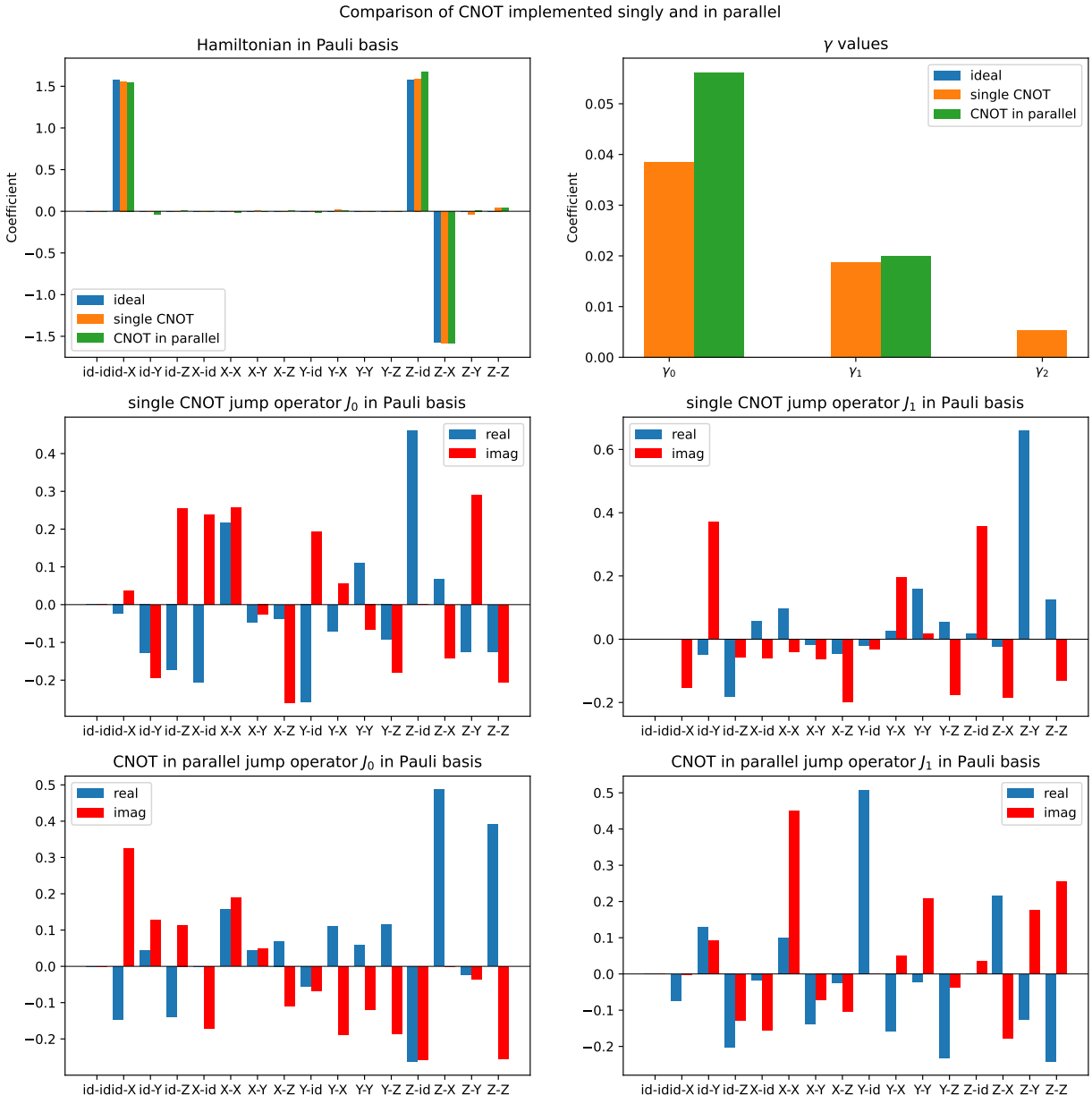


Figure 5.16: Comparison of the canonical decompositions of the ideal unitary CNOT gate with the fitted Lindbladians after Gate Set Flip-Flop for the noisy CNOT gate implemented on qubit pair $(13, 14)$ on *ibm_cairo*. Note that the estimated decomposition for the noisy CNOT differs depending on whether the CNOT was implemented in isolation while the rest of the device was left idle (orange) or in parallel with CNOT gates implemented elsewhere (green).

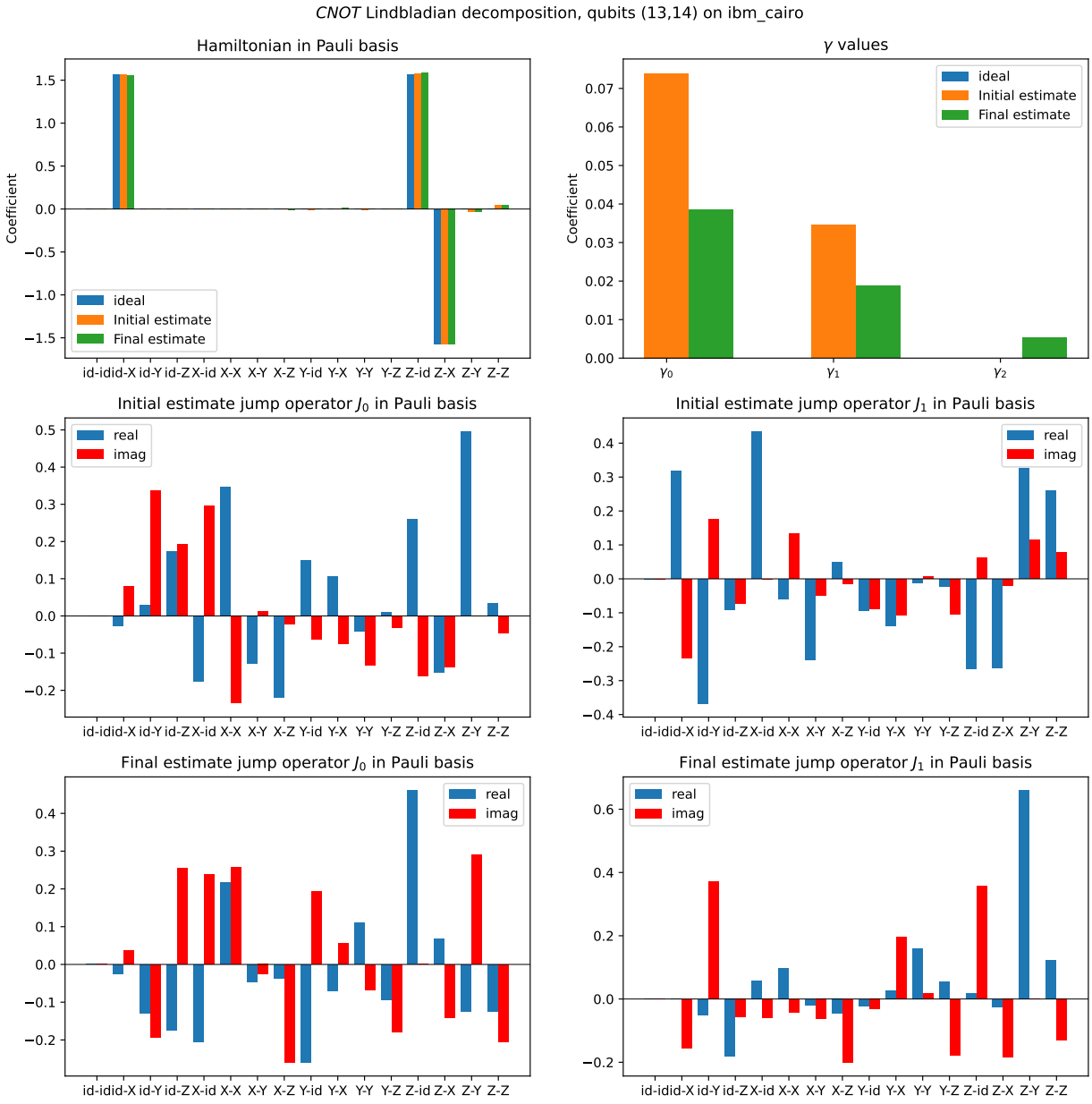


Figure 5.17: The estimated canonical decomposition (Section 5.1.1) of the estimated Lindbladian generators for the CNOT implemented on qubit pair $(13, 14)$ on *ibm_cairo*. Here we compare the ideal unitary process with experimental estimates before and after Gate Set Flip-Flop, for the case where the gate was implemented on $(13, 14)$ alone while the rest of the device was left idle.

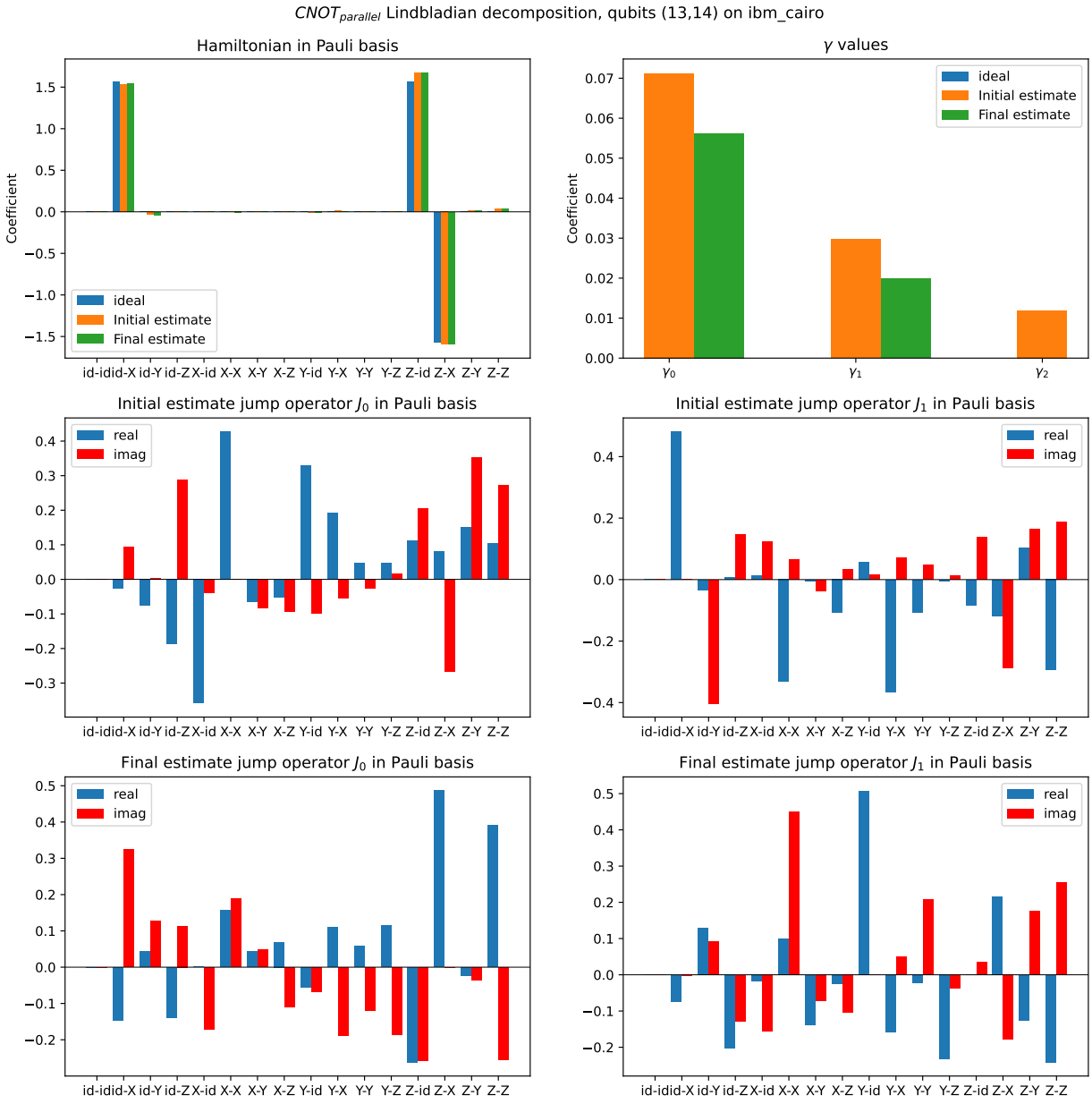


Figure 5.18: Estimated canonical decomposition (Section 5.1.1) of the estimated Lindbladian generators for the CNOT implemented on qubit pair $(13, 14)$ on *ibm_cairo*. Here we compare the ideal unitary process with experimental estimates before and after Gate Set Flip-Flop, for the case where the gate was implemented in parallel with CNOT gates being applied simultaneously to pairs $(10, 12)$ and $(16, 19)$.

5.5.3 Idling cross-talk experiment on *ibm_perth*

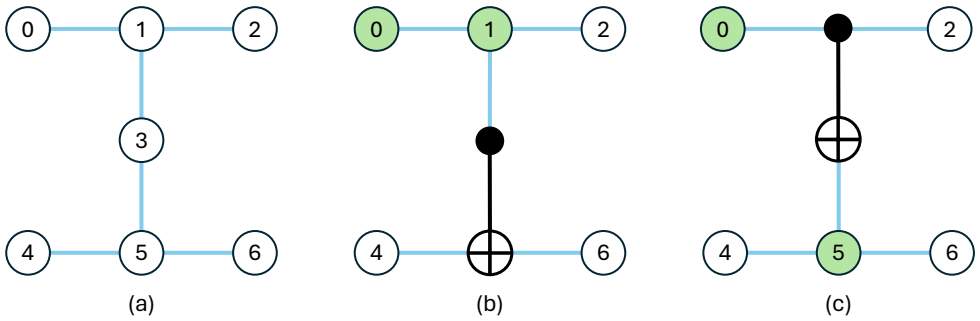


Figure 5.19: (a) Qubit layout on *ibm_perth*. Numbered vertices show qubit locations, edges show connections where CNOT gates are available. (b) Shaded green qubits show the position of the pair targeted for process tomography in the $(0,1)$ run. For the $\mathcal{D}_{\mathrm{CNOT}}$ process, the CNOT is applied on pair $(3,5)$ as shown. (c) During tomography of the $\mathcal{D}_{\mathrm{CNOT}}$ process on qubits $(0,5)$, the CNOT is applied between qubits 1 and 3.

Here we investigate the noisy idling process on pairs of qubits. We take the gate set

$$\mathcal{G} = \{\mathrm{CNOT}, T\otimes I, I\otimes T, \sqrt{X}\otimes I, I\otimes\sqrt{X}, \mathcal{D}, \mathcal{D}^5, \mathcal{D}_X, \mathcal{D}_X^5, \mathcal{D}_{\mathrm{CNOT}}, \mathcal{D}_{\mathrm{CNOT}}^5\}, \tag{5.48}$$

where $\mathcal{D}$ is the 2-qubit channel obtained by sending a delay instruction to the quantum device. We used a delay length of 380 ns to approximately match the time taken to implement a CNOT gate on the device. $\mathcal{D}_X$ and $\mathcal{D}_{\mathrm{CNOT}}$ are obtained by sending the same length of delay to the target qubit pair, while either X or CNOT gates, respectively, are applied to nearby qubits. In this section, we give a demonstration of our algorithm for two different pairs on the *ibm_perth* device, namely the pairs labelled $(0,1)$ and $(0,5)$. In Figure 5.19, we show the device layout and highlight which qubits are targeted for tomography, and where the CNOT is applied in the case of the $\mathcal{D}_{\mathrm{CNOT}}$ process. For the $\mathcal{D}_X$ processes, X gates are applied to all qubits not targeted for tomography. For example, when characterizing the $\mathcal{D}_X$ idling process on qubit pair $(0,1)$, the X gate is applied to qubits 2, 3, 4, 5, and 6. For the $(0,1)$ pair, the CNOT used to aid gauge optimization is available on the device as an elementary gate. However, for the gate set targeting $(0,5)$, a direct connection is not available. We instead implement the CNOT for this pair using the circuit decomposition shown in Figure 5.20. Once again, we use the Alternating Projections method for fitting the CNOT, and Convex Solve on the principal branch for all others.

We show the convergence for pairs $(0,1)$ and $(0,5)$ in Figure 5.21. We remark that the approach of using Convex Solve on a single branch is still successful in achieving convergence to a good set of solutions, despite the fact that several of the idling processes turn out to be very far from the identity channel as evidenced in Table 5.4. The distances of the 1900 ns idling processes from the ideal were observed to be much larger than any seen for the 2-qubit gates analyzed, whereas distances to the estimated Markovian channel were comparable to other cases. The Lindbladian decompositions are plotted in Figures 5.22 and 5.23. We find that the noise differs significantly in character depending on the qubit choice and the gates applied in parallel.

Remarkably, for both pairs, the noise is reduced dramatically when X gates are applied to the other qubits, compared to the case where all qubits are left to idle. One possible

explanation is that repeatedly applying X pulses to neighbouring qubits might act as a rudimentary dynamical decoupling sequence [137]. For the $(0,1)$ pair, repeatedly applying a CNOT on qubits 3 and 5 appears to induce the buildup of a coherent Z error on qubit 1, while both the $\mathcal{D}$ and $\mathcal{D}_{\mathrm{CNOT}}$ processes exhibit a ZZ coupling between the target qubits. In contrast, the more separated $(0,5)$ pair shows no significant ZZ coupling. Meanwhile, the application of CNOTs on the pair $(1,3)$, situated between the target qubits $(0,5)$, appears to reduce the effective coherent Z error on qubit 5 while increasing it on qubit 0.

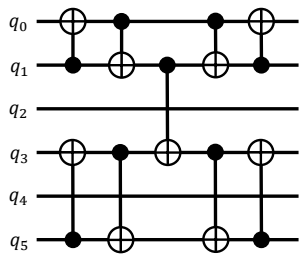


Figure 5.20: Circuit decomposition for CNOT between qubits 0 and 5 on *ibm_perth*.

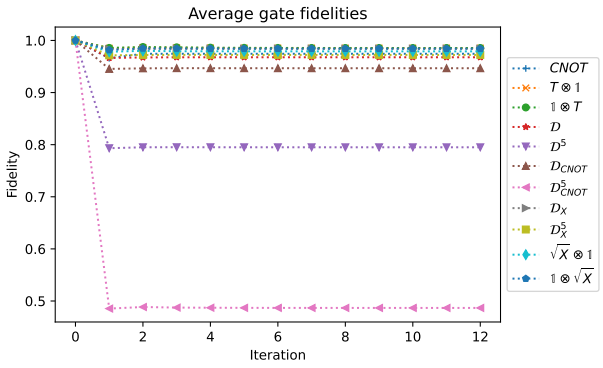
Process	$\|E - E^{\mathrm{ideal}}\|$	$\|E - \exp(L)\|$
CNOT	0.28	0.23
$T \otimes I$	0.25	0.22
$I \otimes T$	0.25	0.23
$\mathcal{D}$	0.57	0.22
$\mathcal{D}^5$	2.08	0.18
$\mathcal{D}_{\mathrm{CNOT}}$	1.08	0.21
$\mathcal{D}^5_{\mathrm{CNOT}}$	4.07	0.22
$\mathcal{D}_X$	0.33	0.24
$\mathcal{D}^5_X$	0.75	0.21
$\sqrt{X} \otimes I$	0.27	0.22

(a) Qubit pair (0,1)

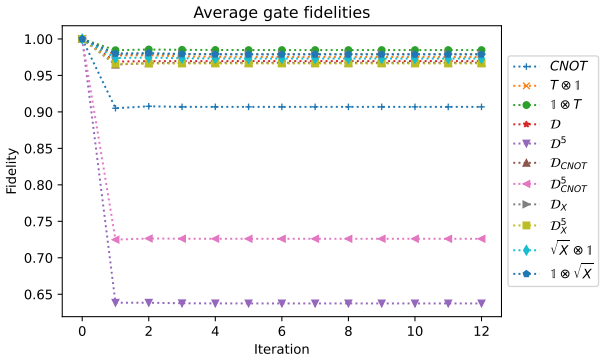
Process	$\|E - E^{\mathrm{ideal}}\|$	$\|E - \exp(L)\|$
CNOT	0.76	0.24
$T \otimes I$	0.30	0.23
$I \otimes T$	0.28	0.25
$\mathcal{D}$	0.77	0.21
$\mathcal{D}^5$	3.23	0.23
$\mathcal{D}_{\mathrm{CNOT}}$	0.66	0.26
$\mathcal{D}^5_{\mathrm{CNOT}}$	2.63	0.18
$\mathcal{D}_X$	0.35	0.24
$\mathcal{D}^5_X$	0.55	0.20
$\sqrt{X} \otimes I$	0.31	0.24

(b) Qubit pair (0,5)

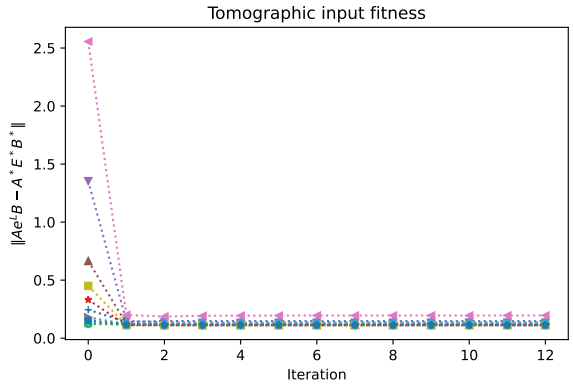
Table 5.4: Distances between the tomographic channel estimates and corresponding ideal transfer matrices E^{ideal} and estimated Markovian channels $\exp(L)$, respectively, from idling analysis on *ibm_perth*, on qubit pairs (a) (0,1), and (b) (0,5). For each process P^*, the tomographically estimated transfer matrix estimate $E = B(g^*)^{-1}P^*B^{-1}$ is evaluated in the final gauge B obtained from the flip-flop algorithm.



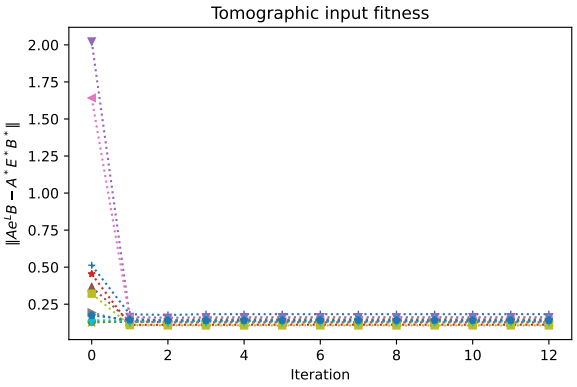
(a) Qubit pair $(0, 1)$



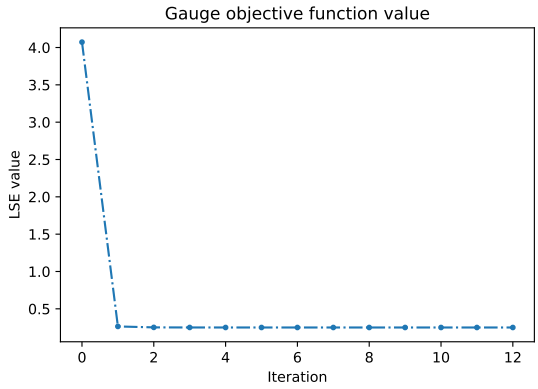
(b) Qubit pair $(0, 5)$



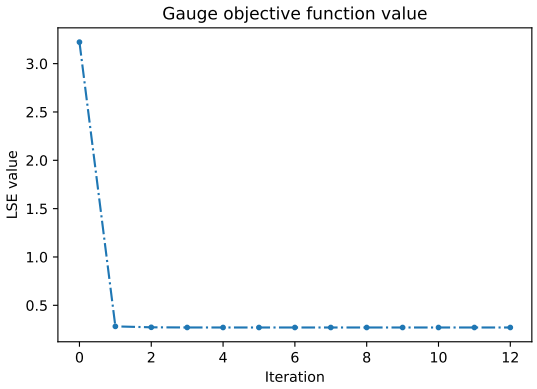
(c) Qubit pair $(0, 1)$



(d) Qubit pair $(0, 5)$



(e) Qubit pair $(0, 1)$



(f) Qubit pair $(0, 5)$

Figure 5.21: Gate Set Flip-Flop figures of merit for the idling experiment on qubits $(0, 1)$ and $(0, 5)$ from *ibm_perth*. (a,b) The average gate fidelity between the ideal gate and the estimated Markovian process at each iteration, computed using (5.42). (c,d) Distance between raw tomographic data and the equivalent data for the estimated Markovian process assuming the current gauge. (e,f) Gauge objective function as defined in (5.41).

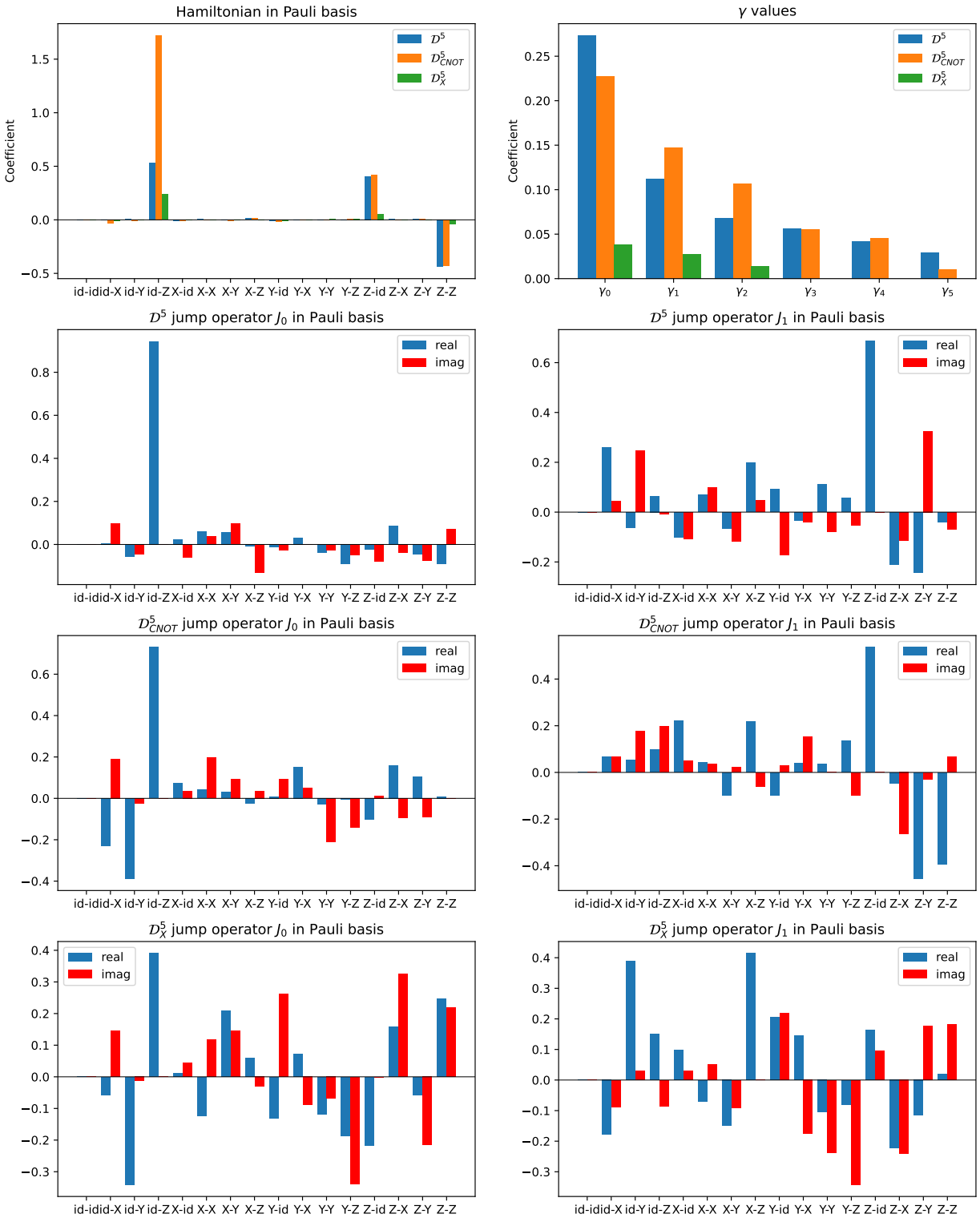


Figure 5.22: Canonical decompositions of the estimated Lindbladians for three variants of the idling process on the adjacent qubit pair $(0, 1)$ on *ibm_perth* (see Figure 5.19). The ideal process (i.e., the generator of the identity channel) is not shown since the Hamiltonian and jump operators would all be zero. The estimates are output from the Gate Set Flip-Flop algorithm once converged. (Blue) The whole device is left to idle for 5 repetitions of the 380 ns delay instruction; (Orange) pair $(0, 1)$ idles for the same length of time while a CNOT is applied in parallel to pair $(3, 5)$ for 5 repetitions (see Figure 5.19 (b)); (Green) pair $(0, 1)$ idles while the X gate is applied 5 times to all other qubits.



Figure 5.23: Canonical decompositions of idling processes for the separated qubit pair $(0, 5)$ on *ibm_perth* (see Figure 5.19), estimated using Gate Set Flip-Flop. (Blue) The whole device is left to idle for 5 repetitions of the 380 ns delay instruction; (Orange) Pair $(0, 5)$ idles for the same length of time while a CNOT is applied in parallel to pair $(2, 3)$ for 5 repetitions (see Figure 5.19 (c)); (Green) Pair $(0, 5)$ idles while the X gate is applied 5 times to all other qubits.

5.6 Deferred discussions

5.6.1 Direct numerical estimation

A simple example of direct numerical estimation would be taking $f(x) = \|\exp(L(x)) - E\|$, where $L(x)$ is a Lindbladian parameterized by the optimization variables x, E is the input data, and a solver (e.g., MATLAB `fminsearch`) is called to minimize f with respect to x. For this particular formulation, the parametrization x can be taken to have length $d^2(d^2-1)$, gradients of the objective function f can be obtained via automatic differentiation so that gradient-based solvers (e.g., MATLAB `fminunc`) can be used, and L^{ideal}, which is known to the experimenter, serves as a good initial guess to seed the optimization. An early example of this approach estimated the Markovian evolution for a 2-qubit nuclear magnetic resonance (NMR) system [105] by minimizing a time-series generalization of the objective f using MATLAB `fminsearch`, which implements the classic Nelder-Mead simplex search algorithm [138]. More recently, state-of-the-art demonstrations of Lindbladian tomography have used an objective function based on maximum likelihood estimation (MLE) [109, 110]. This has been applied to a 2-qubit idling process on a superconducting-qubit device in [109]. Quantum process tomography is carried out for a series of different time delays to build up a time series of tomographic snapshots. SPAM errors are filtered out by preparing a fiducial state and immediately measuring in the standard basis. The resulting SPAM-filtered channel estimates are then fit into a log-likelihood function that is maximized using MATLAB `fminsearch` to extract the noisy Lindblad generator of the idling channel. An evolution of this MLE technique was proposed in [110]. Under the assumption of small noise, the log-likelihood function is linearized to simplify the task into a convex-optimization problem. The procedure is also combined with the compressed sensing technique to reduce the number of required measurements. It is worth emphasizing that despite being observed to be effective in practice in prior works, local direct numerical optimization approaches (just like ours) are heuristic in nature and have little formal guarantee of returning optimal solutions.

5.6.2 Ad-hoc methods

Another approach to Lindbladian estimation is to impose some assumptions on the system being investigated and to fit the system using a restricted Lindbladian model with a reduced number of parameters. Such techniques have been demonstrated in small-scale NMR [103], quantum optics [104] and superconducting-qubit [100, 108, 114] systems. We note that in the latter example the authors use Lindbladian fitting as one part of a protocol used to analyze non-Markovian dynamics via post-Markovian master equations [139]. The characterization of such non-Markovian processes is beyond the scope of the present work. One restricted form of Lindbladian that has recently proved useful in the context of error mitigation on quantum computing devices with limited connectivity is the so-called sparse Pauli-Lindblad model [22, 23]. The method assumes that a single layer of 2-qubit gates has an associated Markovian noise process generated solely by local Pauli operators, where the locality is with respect to the connectivity of the device. This yields a Lindbladian that can be specified by a number of parameters that grows only linearly with the system size. The parameters

can then be estimated efficiently by fitting the decay of Pauli observables after repeated applications of the gate layer. The scalability of this method makes it appealing for practical implementation, and indeed it has been demonstrated successfully on intermediate-sized superconducting-qubit devices as a step in probabilistic error cancellation [22] and zero noise extrapolation [23] error mitigation protocols. It should be noted, however, that the assumption of a Pauli-Lindblad model is a rather strong one, resting on the assumption that the noisy gate layer can be well-modelled by the ideal gate layer composed with a stochastic Pauli channel with local generators. The stochastic Pauli channel assumption was justified in [22, 23] by the fact that gate layers are implemented under a particular randomized compiling protocol known as Pauli twirling [140–144]. This, as well as the noise characterization step, in turn relied on the fact that the 2-qubit gates being analyzed were Clifford gates, and so preserve the Pauli group. These considerations mean that the method of [22, 23] is only applicable to particular quantum computing devices with 1- and 2-qubit gates that are already relatively well calibrated, in which case it can provide a good effective noise model that could be used in certain error mitigation techniques. It is not aimed at estimating the underlying noise process for specific few-qubit processes.

5.6.3 Lindbladian canonical decomposition

Given a Lindbladian L in the elementary basis representation, it is possible to solve for a canonical decomposition comprised of a Hamiltonian H, jump operators $\{J_\alpha\}$, and real positive scalars $\{\gamma_\alpha\}$ such that the elementary basis representation of the Lindbladian form (see (5.6)) defined by H, $\{J_\alpha\}$, and $\{\gamma_\alpha\}$ is L. Let $n = \log_2(d)$ and let $P_1, \ldots, P_{d^2-1}$ denote the non-identity n-qubit Pauli operators. Since H is traceless, H is determined by its expansion in the Pauli basis

$$H = \sum_{k=1}^{d^2-1} \frac{\mathrm{Tr}(P_k H)}{d} P_k. \tag{5.49}$$

Using the formula (5.4), we get

$$L = i(I \otimes H^T - H \otimes I) + \sum_\alpha \gamma_\alpha \left[J_\alpha \otimes \overline{J_\alpha} - \frac{1}{2}(J_\alpha^\dagger J_\alpha \otimes I + I \otimes J_\alpha^T \overline{J_\alpha}) \right]. \tag{5.50}$$

By the fact that H and $\{J_\alpha\}$ are traceless, for every $k \in \{1, \ldots, d^2 - 1\}$,

$$\mathrm{Tr}(L(P_k \otimes I)) = -i\mathrm{Tr}(P_k H)d - \frac{1}{2}\sum_\alpha \gamma_\alpha \mathrm{Tr}(P_k J_\alpha^\dagger J_\alpha)d. \tag{5.51}$$

Thus,

$$\mathrm{Tr}(L(P_k \otimes I)) - \overline{\mathrm{Tr}(L(P_k \otimes I))} = -2i\mathrm{Tr}(P_k H)d \tag{5.52}$$

$$\implies \mathrm{Tr}(P_k H) = \frac{i(\mathrm{Tr}(L(P_k \otimes I)) - \overline{\mathrm{Tr}(L(P_k \otimes I))})}{2d}. \tag{5.53}$$

Similarly, for every $j, k \in \{1, \ldots, d^2 - 1\}$,

$$\mathrm{Tr}(L(P_j \otimes \overline{P_k})) = \sum_\alpha \gamma_\alpha \mathrm{Tr}(P_j J_\alpha)\mathrm{Tr}(P_k J_\alpha^\dagger) \tag{5.54}$$

and we can organize these values into a $(d^2-1)\times(d^2-1)$ matrix

$$C = \begin{bmatrix} \mathrm{Tr}(L(P_1 \otimes \overline{P_1})) & \cdots & \mathrm{Tr}(L(P_1 \otimes \overline{P_{d^2-1}})) \\ \vdots & \ddots & \vdots \\ \mathrm{Tr}(L(P_{d^2-1} \otimes \overline{P_1})) & \cdots & \mathrm{Tr}(L(P_{d^2-1} \otimes \overline{P_{d^2-1}})) \end{bmatrix}. \tag{5.55}$$

Since the jump operators are orthonormal,

$$C = \sum_\alpha \gamma_\alpha \begin{bmatrix} \mathrm{Tr}(P_1 J_\alpha) \\ \vdots \\ \mathrm{Tr}(P_{d^2-1} J_\alpha) \end{bmatrix} \begin{bmatrix} \mathrm{Tr}(P_1 J_\alpha^\dagger) & \cdots & \mathrm{Tr}(P_{d^2-1} J_\alpha^\dagger) \end{bmatrix} \tag{5.56}$$

is a spectral decomposition of C. Thus, the positive scalars $\{\gamma_\alpha\}$ and the jump operators $\{J_\alpha\}$ can be retrieved by computing a spectral decomposition of C.

5.6.4 Analyzing the Convex Solve method for the Lindbladian fitting problem without statistical error

Recall that the input to the Lindbladian fitting problem is a $d^2\times d^2$ transfer matrix E, and the goal is to search for an $L\in\mathcal{L}$ that minimizes $\|e^L - E\|$ where $\mathcal{L}$ is the set of all Lindbladians. In Section 5.2.1, we have restricted ourselves to special cases of the problem satisfying two additional assumptions. Namely, we assume there exists an unknown Lindbladian L^* such that for $E^* = e^{L^*}$, $\|E - E^*\| \le c_1$ for some small constant c_1, and there exists a known Lindbladian L^{ideal} such that $\|L^* - L^{\text{ideal}}\| \le c_2$ for some small constant c_2. In this appendix, we further assume $E = E^*$ (i.e., $c_1 = 0$), so the goal becomes outputting an $L\in\mathcal{L}\cap\log(E)$ which need not necessarily be L^* and where $\log(E) = \{A : e^A = E\}$.

We first establish a key lemma. For every diagonalizable matrix A, let $\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$ denote its spectral radius. For every invertible matrix B, let $\kappa(B) = \|B\|\|B^{-1}\|$ denote its condition number.

Lemma 5.1. *Write $L^{ideal} = VDV^{-1}$ for some diagonal matrix D and invertible matrix V. If $\|L^* - L^{ideal}\| < \frac{\pi-\rho(L^{ideal})}{\kappa(V)}$, then for every eigenvalue μ of L^*, $|\Im(\mu)| < \pi$.*

Proof. Let μ be an eigenvalue of L^*. By the Bauer-Fike theorem, there exists an eigenvalue λ of L^{ideal} such that

$$|\mu| - |\lambda| \le |\mu - \lambda| \le \kappa(V)\|L^* - L^{\text{ideal}}\| < \pi - \rho(L^{\text{ideal}}). \tag{5.57}$$

Thus,

$$|\Im(\mu)| \le |\mu| < \pi - \rho(L^{\text{ideal}}) + |\lambda| \le \pi. \tag{5.58}$$

□

Note that for L^{ideal} generating a unitary gate, all the eigenvalues of the ideal transfer matrix $E^{\text{ideal}} = e^{L^{\text{ideal}}}$ have modulus 1, which implies $\rho(L^{\text{ideal}}) \le \pi$ for L^{ideal} in the principal branch. It is worth emphasizing that Lemma 5.1 is not asymptotic in nature, and its assumption, which only depends on L^{ideal}, holds for realistic values of c_2 for many practically

relevant L^{ideal}. For example, for the identity gate with $L^{\text{ideal}} = 0$, we have $\rho(L^{\text{ideal}}) = 0$ and $\kappa(V) = 1$, so $c_2 < \pi$ suffices. As another example, for the $T \otimes I$ gate, we have $\rho(L^{\text{ideal}}) = \frac{\pi}{4}$ and $\kappa(V) = 1$, so we can choose $c_2 < \frac{3}{4}\pi$. Unfortunately, for both CNOT and ISWAP, we have $\rho(L^{\text{ideal}}) = \pi$, so Lemma 5.1 does not apply.

When E has degenerate eigenvalues, a degenerate eigenspace of E could split into different eigenspace of L^* due to the non-uniqueness of complex logarithms, which induces continuous degrees of freedom in specifying an $A \in \log(E)$. We call this phenomenon eigenspace splitting. It should be clear that conditioned on eigenspace splitting not occurring for L^*, the Convex Solve method will eventually locate L^* in $\log(E)$ since in this case, the Convex Solve method correctly enumerates over the remaining discrete degrees of freedom. Our main observation states that if the assumption of Lemma 5.1 holds, then eigenspace splitting cannot occur for L^*. We can now prove Theorem 5.1 from the main text, which stated that provided L^* and L^{ideal} satisfy $\|L^* - L^{\text{ideal}}\| < \frac{\pi - \rho(L^{\text{ideal}})}{\kappa(V)}$, then given the exact transfer matrix $E = e^{L^*}$, the Convex Solve method will output a solution $L \in \mathcal{L} \cap \log(E)$.

Proof of Theorem 5.1. First consider the case where E has d^2 distinct eigenvalues. Then Lemma 5.1 directly implies that L^* is the principal branch of $\log(E)$.

Now suppose $E = E^*$ contains degenerate eigenvalues. Suppose there exist $a, b \in \mathbb{R}$ and integer k with $|k| \geq 1$ such that L^* has two eigenvalues $a + bi$ and $a + bi + 2\pi ki$ which both exponentiate to the eigenvalue e^{a+bi} for E. Then either $|b| \geq \pi$ or $|b + 2\pi k| \geq \pi$, which contradicts Lemma 5.1. Therefore, L^* and E have the same set of eigenspace projectors and L^* is the principal branch of $\log(E)$. □

Although Theorem 5.1 does not apply if E^{ideal} has real negative eigenvalues, it remains true that if eigenspace splitting does not occur for L^*, then the Convex Solve method succeeds on E. Next, we describe two mechanisms by which eigenspace splitting could occur when E^{ideal} contains real negative eigenvalues for arbitrarily small $\|L^* - L^{\text{ideal}}\|$.

Consider a situation where -1 is a degenerate eigenvalue of E^{ideal} with degeneracy at least 4 (for example, the ideal transfer matrix of CNOT has a 6-dimensional degenerate -1 eigenspace). Then πi and $-\pi i$ are eigenvalues of L^{ideal} each with degeneracy at least 2. Then for every $\varepsilon > 0$, it is possible for L^* to have eigenvalues $(\pi + \varepsilon)i$, $(-\pi - \varepsilon)i$, $(\pi - \varepsilon)i$, and $(-\pi + \varepsilon)i$. Now, $(\pi + \varepsilon)i$ and $(-\pi + \varepsilon)i$ exponentiate to the same eigenvalue for E^* and similarly for $(\pi - \varepsilon)i$ and $(-\pi - \varepsilon)i$. Thus, a degenerate eigenspace of E^* corresponding to the eigenvalue $e^{(\pi+\varepsilon)i}$ gets split into different eigenspaces for L^* corresponding to the eigenvalues $(\pi + \varepsilon)i$ and $(-\pi + \varepsilon)i$. Note that since ε is arbitrary, this can happen for arbitrary $\|L^* - L^{\text{ideal}}\|$, so no condition on $\|L^* - L^{\text{ideal}}\|$ alone can guarantee eigenspace splitting does not occur for noisy CNOT gates. We note that the behaviour described in this paragraph has never been observed in practice during our numerical testing.

Another failure mode for the Convex Solve method occurs when E persists to have real negative eigenvalues. This can happen when precisely one out of the two qubits is noisy. We view such scenarios to be highly non-generic.

In the real world, the input E obtained from quantum process tomography suffers from finite statistical error controlled by the number of tomographic shots. In particular, $E \neq E^*$ and E will almost never correspond to a Markovian channel. At best, we can assume E is the transfer matrix of a CPTP map since it is cheap to project E to be CPTP. Hence, the

actual input E to our problem is always the transfer matrix of a non-Markovian channel, and the goal is to find a Lindbladian L that best fits E with the promise that E is close to some unknown Markovian channel E^*. The presence of statistical noise in the input E is necessary for the Lindbladian fitting problem to be non-trivial for the family of Markovian channels we consider as statistical noise causes the input to be non-Markovian.

5.6.5 Detailed description of Alternating Projections

In this appendix, we give a more complete description of our Alternating Projections method for solving the Lindbladian fitting problem. Here we report the specific implementation used to obtain the numerical results given in the main text, but we note that our approach is heuristic, and it is not necessary to be prescriptive about technical details such as the random perturbation and logarithm search steps. For clarity, we separate out the outer loop (Algorithm 4) that handles the pre-processing of the input tomographic data and the core subroutine (Algorithm 5) that alternates between projections of the Lindbladian eigenvectors onto the approximate eigenspaces of the tomographic data and projection of the resultant approximate matrix logarithm $\bar{A}$ onto the set of Lindbladians. We assume the input transfer matrix E and initial guess Lindbladian L_0 are both diagonalizable. In the presence of statistical error, the eigenvalues of E will be distinct, but this is not a requirement of the algorithm.

The enumeration of relevant branches of the logarithm in Line 7 of Algorithm 4 goes as follows. We assume for simplicity that $|\Im(\lambda)| \leq \pi$ for every eigenvalue λ of L^{ideal} i.e., L^{ideal} is in the principal branch of $\log(E^{\text{ideal}})$. We then start with an initial set of branches

$$M' = \{(m_1, \ldots, m_{d^2}) \in \mathbb{Z}^{d^2} : m_j \in \{-1, 0, 1\} \text{ for every } j \in \{1, \ldots, d^2\}\}, \tag{5.59}$$

where each entry corresponds to a shift in the j-th eigenvalue by $2\pi m_j i$, as per Line 9 of Algorithm 4. Recall that for $\hat{\lambda}_1, \ldots, \hat{\lambda}_{d^2}$ to be the set of eigenvalues of a $d^2 \times d^2$ Lindbladian matrix, they need to come in complex conjugate pairs [111]. M is defined to be the subset of elements of M' satisfying this condition. This constraint alone greatly reduces the number of branches of M' to try, and the search over $m \in M$ is embarrassingly parallelizable.

The minimum-cost maximum-flow problem in Line 5 of Algorithm 5 is formulated as follows. We construct a directed graph with $d^2 + n + 2$ vertices, where n is the number of approximate eigenspaces determined by the eigenvalue-clustering procedure in Line 5 of Algorithm 4. There is a vertex corresponding to each eigenvector $|\tilde{v}_j\rangle\!\rangle$, $j \in \{1, \ldots, d^2\}$ and a vertex corresponding to each eigenspace projector Π_k, $k \in \{1, \ldots, n\}$. For convenience, we shall use $\tilde{v}_j$ and Π_k to denote these vertices. In addition, there is a special source vertex s and a special sink vertex t. Every directed edge in the graph has an integer capacity and an integer weight. For every $j \in \{1, \ldots, d^2\}$, there is an edge from s to $\tilde{v}_j$ with capacity 1 and weight 0. For every $j \in \{1, \ldots, d^2\}$ and $k \in \{1, \ldots, n\}$, there is an edge from $\tilde{v}_j$ to Π_k with capacity 1 and weight[2] given by $\lfloor 10000 \|\tilde{v}_j - \Pi_k \tilde{v}_j\| \rfloor$. For every $k \in \{1, \ldots, n\}$,

[2]The purpose of the constant 10^4 here is to amplify differences between floating point values so that any pair of edges for which $\|\tilde{v}_j - \Pi_k \tilde{v}_j\|$ differs will have distinct integer weights. This is to avoid numerical problems that can afflict flow-problem solvers when edges have floating point weights. The precise value of the large constant is unimportant.

Algorithm 4: Alternating Projections outer loop

input : Estimated (diagonalizable) transfer matrix E; initial guess Lindbladian L_0; precision $\beta \geq 0$; maximum search depth T; number of random starts N

output: Estimated Lindbladian L^{est} such that $\exp L^{\text{est}} \approx E$

Compute set of eigenvalue-eigenvector pairs $\{(\mu_j, |r_j\rangle\!\rangle)\}_{j=1}^{d^2}$ for E, and compute set of corresponding left eigenvectors $\{\langle\!\langle l_j|\}_j$ such that $\langle\!\langle l_i|r_j\rangle\!\rangle = \delta_{i,j}$;

for $j = 1$ **to** d^2 **do**

  $\lambda_j \leftarrow$ logarithm of μ_j such that $|\Im(\lambda_j)| \leq \pi$

end

Perform eigenvalue clustering: μ_i, μ_j are taken to be in the same cluster if $|\mu_i - \mu_j| \leq \beta$. Suppose n clusters are detected. The k-th cluster is defined by an index set $C_k = \{i_{k,1}, \ldots, i_{k,w_k}\}$, where $w_k = |C_k|$ corresponds to the rank of the associated subspace so that for each pair $(\mu_j, |r_j\rangle\!\rangle)$, the index j is in exactly one of the index sets ;

For each k, let $\Pi_k = \sum_{j \in C_k} |r_j\rangle\!\rangle\langle\!\langle l_j|$;

Enumerate hermiticity-compatible branches of the matrix logarithm M ;

`// See main text for construction of` M`.`

for $m \in M$ **do**

  $\hat{\lambda}_{i_{k,a}} \leftarrow \lambda_{i_{k,a}} + 2\pi m_{i_{k,a}} i$;

  for $p = 1$ **to** N **do**

    Generate random perturbation $\tilde{L}_0$ of L_0 ;

    `// In our implementation, we alternate between` $\tilde{L}_0 = L_0 + D$ `and` $\tilde{L}_0 = L_0 + H^{\otimes d} D H^{\otimes d}$ `for random diagonal matrix` D `with small norm.`

    $L_{m,p} \leftarrow$ `APCore(`$\tilde{L}_0$, T, $\{\hat{\lambda}_j\}_{j=1}^{d^2}$, $\{\Pi_k\}_{k=1}^{n}$, $\{C_k\}_{k=1}^{n}$`)`;

  end

end

$L_{\text{est}} \leftarrow \text{argmin}_{L_{m,p}} \|e^{L_{m,p}} - E\|$;

return L_{est}

Algorithm 5: Alternating Projections core subroutine

input : Approximate Lindbladian matrix $\tilde{L}_0$; maximum iteration depth T; list of log-eigenvalues $\hat{\lambda}_j$ of tomographic data E; set of n orthogonal projectors $\{\Pi_k\}_{k=1}^n$ satisfying $\sum_{k=1}^n \mathrm{rank}(\Pi_\mathrm{k}) = d^2$; set of index sets $\{C_k\}_{k=1}^n$ corresponding to eigenvalue clustering

output: A candidate Lindbladian $\tilde{L}_T$

Function `APCore`($\tilde{L}_0$, T, $\{\lambda_j\}_{j=1}^{d^2}$, $\{\Pi_k\}_{k=1}^n$, $\{C_k\}_{k=1}^n$)**:**

- $\hat{D} \leftarrow \mathrm{Diag}(\hat{\lambda}_1, \ldots, \hat{\lambda}_{d^2})$
- **for** $t = 0$ **to** $T - 1$ **do**
  - Compute eigenvalue-eigenvector pairs $(\sigma_1, |\tilde{v}_1\rangle\rangle), \ldots, (\sigma_{d^2}, |\tilde{v}_{d^2}\rangle\rangle)$ for $\tilde{L}_t$
  - Assign each $(\sigma_j, |\tilde{v}_j\rangle\rangle)$ to exactly one subspace projector Π_k by solving minimum-cost maximum-flow problem to minimize $\||\tilde{v}_j\rangle\rangle - \Pi_k|\tilde{v}_j\rangle\rangle\|$ for each pairing. For each k, let this assignment define an index set $C'_k = \{i'_{k,a}\}_{a=1}^{w_k}$, so that j is in C'_k if $|\tilde{v}_j\rangle\rangle$ is assigned to the k-th subspace
  - Initialize $d^2 \times d^2$ matrix K
  - **for** $k = 1$ **to** n **do**
    - Find minimum-cost perfect matching W_k between data log-eigenvalues $\hat{\lambda}_{i_{k,a}}$ where $i_{k,a} \in C_k$ and Lindbladian eigenvalues $\sigma_{i'_{k,b}}$ where $i'_{k,b} \in C'_k$ according to cost function $|\hat{\lambda}_{i_{k,a}} - \sigma_{i'_{k,b}}|$
    - For each $\{\hat{\lambda}_{i_{k,a}}, \sigma_{i'_{k,b}}\} \in W_k$, set the $i_{k,a}$-th column of K to $\Pi_k|\tilde{v}_{i'_{k,b}}\rangle\rangle$
  - **end**
  - $\bar{A} \leftarrow K\hat{D}K^{-1}$
  - Let L_opt be the solution to the convex optimization problem:
    - **minimize** $\|L - \bar{A}\|$
    - **subject to** L Lindbladian
  - **if** $\|e^{L_{opt}} - E\| < \|e^{\tilde{L}_t} - E\|$ **then**
    - $\tilde{L}_{t+1} \leftarrow L_\mathrm{opt}$
  - **else**
    - $\tilde{L}_T \leftarrow \tilde{L}_t$
    - **break**
- **end**
- **return** $\tilde{L}_T$

there is an edge from Π_k to t with capacity the rank of Π_k and weight 0. The capacity constraints on the last set of edges encode the requirement that an eigenspace projector Π_k with rank l needs to accept exactly l eigenvectors. Then, an assignment of the eigenvectors to the eigenspace projectors that minimizes the sum of the weights can be read off from a minimum-cost maximum (s,t)-flow of the graph.

By diagonalizing L^{ideal} as $L^{\text{ideal}} = VDV^{-1}$, assuming $\|L^* - L^{\text{ideal}}\| < \frac{2\pi}{\kappa(V)}$, and together with the simplifying assumption that L^{ideal} is in the principal branch of $\log(E^{\text{ideal}})$, the same argument in the proof of Lemma 5.1 implies that it suffices to consider $|m_j| \leq 1$ for every $j \in \{1, \ldots, d^2\}$. The effect of the randomness introduced in Line 11 of Algorithm 4 is to induce a slightly different eigendecomposition of the initial guess Lindbladian in each call to `APCore` in Line 12. This randomness in the starting point $\tilde{L}_0$ can lead to a different candidate Lindbladian model $\tilde{L}_T$ output for each call. Therefore, we repeat this procedure N times for each branch and accept the candidate model that best fits the tomographic data.

5.7 Future directions

Two avenues suggest themselves for further study. For simplicity, in this work, we have considered individual tomographic snapshots in isolation rather than analyzing sequences of snapshots as a time-series. In particular, for gate sets that include repetitions of the same gate or increased pulse duration of the same type of gate, our Gate Set Flip-Flop algorithm fits separate time-independent Lindbladians for each snapshot, whereas it would make sense to instead treat the snapshots as a time series generated by the same underlying Lindbladian. Indeed, the precursor work [111] provided variants of the algorithm that could be applied to time series data for both time-independent and time-dependent Lindbladians. An obvious next step, then, would be to show how our Alternating Projections and Flip-Flop algorithms could be adapted to deal with time series data with possibly time-dependent Lindbladians. We leave this extension for future work.

Another limitation of our method, as presented, is that it is only practical for application to subsystems of small numbers of qubits. In this work, we have focused on 2-qubit gates. It is not necessarily completely infeasible to apply the method without modification to a 3-qubit gate set, but it would certainly lead to much longer runtimes, and going beyond this is likely not possible due to the exponential overhead in the number of parameters needed to fully characterize an n-qubit process. This limitation is shared by any fully general method for Lindbladian fitting, and indeed full QPT and GST. As noted during the discussion of related works, scalable methods for Lindbladian estimation do exist, but these must rely on strong assumptions about the noise model in order to reduce the number of parameters to be determined. A recent example is the local Pauli-Lindblad model employed in the protocols of [22, 23]. An interesting direction for future work would be to investigate whether the techniques developed in the present work could be extended to larger systems by restricting the class of Lindbladians.

Bibliography

[1] Sergey Bravyi, David Gosset, and Yinchen Liu. Classical simulation of peaked shallow quantum circuits. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 561–572, 2024.

[2] Adam Bene Watts, David Gosset, Yinchen Liu, and Mehdi Soleimanifar. Quantum advantage from measurement-induced entanglement in random shallow circuits. *PRX Quantum*, 6(1):010356, 2025.

[3] Yinchen Liu, Max McGinley, Thomas Schuster, and David Gosset. Conditional dependence and Scrooge ensembles in shallow random quantum circuits. *arXiv preprint arXiv:2608.12255*, 2026.

[4] Yinchen Liu, James R Seddon, Tamara Kohler, Emilio Onorati, and Toby S Cubitt. Robust Lindbladian estimation for quantum dynamics. *arXiv preprint arXiv:2507.07912*, 2025.

[5] John C Napp, Rolando L La Placa, Alexander M Dalzell, Fernando GSL Brandao, and Aram W Harrow. Efficient classical simulation of random shallow 2d quantum circuits. *Physical Review X*, 12(2):021021, 2022.

[6] David Peral-García, Juan Cruz-Benito, and Francisco José García-Peñalvo. Systematic literature review: Quantum machine learning and its applications. *Computer Science Review*, 51:100619, 2024.

[7] Ryan Babbush, Robbie King, Sergio Boixo, William Huggins, Tanuj Khattar, Guang Hao Low, Jarrod R McClean, Thomas O'Brien, and Nicholas C Rubin. Grand challenge of quantum applications. *PRX Quantum*, 7(2):020101, 2026.

[8] Madelyn Cain, Qian Xu, Robbie King, Lewis RB Picard, Harry Levine, Manuel Endres, John Preskill, Hsin-Yuan Huang, and Dolev Bluvstein. Shor's algorithm is possible with as few as 10,000 reconfigurable atomic qubits. *arXiv preprint arXiv:2603.28627*, 2026.

[9] Alec Eickbusch, Matt McEwen, Volodymyr Sivak, Alexandre Bourassa, Juan Atalaya, Jahan Claes, Dvir Kafri, Craig Gidney, Christopher W Warren, Jonathan Gross, et al. Demonstration of dynamic surface codes. *Nature Physics*, pages 1–8, 2025.

[10] A Paetznick, BW Reichardt, MP da Silva, C Ryan-Anderson, D Aasen, JM Bello-Rivas, JP Campora, R Chao, A Chernoguzov, W van Dam, et al. Improved quantum processor logical error rates via correction and detection. *Nature*, 654(8118):349–355, 2026.

[11] Atom Computing et al. Quantum error correction with the toric code. *arXiv preprint arXiv:2606.04079*, 2026.

[12] Edwin Tham, Michael L Goldman, Shantanu Debnath, Ashay N Patel, Jyothi Saraladevi, Jason Nguyen, Erik Nielsen, Neal Pisenti, Kenneth Wright, John Gamble, et al. Breakeven demonstration of quantum low-density parity-check codes. *arXiv preprint arXiv:2606.06455*, 2026.

[13] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O'Brien. Quantum error mitigation. *Reviews of Modern Physics*, 95:045005, December 2023.

[14] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):4213, 2014.

[15] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *nature*, 549(7671):242–246, 2017.

[16] Ying Li and Simon C Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Physical Review X*, 7(2):021050, 2017.

[17] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Science and Technology*, 3(3):030503, 2018.

[18] Tushar Pandey, Jason Saroni, Abdullah Kazi, and Kartik Sharma. Efficient variational quantum eigensolver methodologies on quantum processors. *arXiv preprint arXiv:2407.16107*, 2024.

[19] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[20] Sergey Bravyi, Sarah Sheldon, Abhinav Kandala, David C Mckay, and Jay M Gambetta. Mitigating measurement errors in multiqubit experiments. *Physical Review A*, 103(4):042605, 2021.

[21] Piotr Czarnik, Andrew Arrasmith, Patrick J Coles, and Lukasz Cincio. Error mitigation with Clifford quantum-circuit data. *Quantum*, 5:592, 2021.

[22] Ewout van den Berg, Zlatko K. Minev, Abhinav Kandala, and Kristan Temme. Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors. *Nature Physics*, 19(8):1116–1121, May 2023.

[23] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023.

[24] Peter W Shor. Fault-tolerant quantum computation. In *Proceedings of 37th conference on foundations of computer science*, pages 56–65. IEEE, 1996.

[25] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 176–188, 1997.

[26] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation. *Science*, 279(5349):342–345, 1998.

[27] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, 2017.

[28] Daniel Gottesman. The Heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[29] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Physical Review A—Atomic, Molecular, and Optical Physics*, 70(5):052328, 2004.

[30] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2010.

[31] Yasuhiro Kondo, Ryuhei Mori, and Ramis Movassagh. Quantum supremacy and hardness of estimating output probabilities of quantum circuits. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1296–1307. IEEE, 2022.

[32] 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.

[33] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani. A polynomial-time classical algorithm for noisy random circuit sampling. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 945–957, 2023.

[34] Hrant Gharibyan, Siddharth Hariprakash, Mohammed Zuhair Mullath, and Vincent P Su. A practical guide to using Pauli path simulators for utility-scale quantum experiments. *arXiv preprint arXiv:2507.10771*, 2025.

[35] Craig Gidney. Stim: a fast stabilizer circuit simulator. *Quantum*, 5:497, July 2021.

[36] Graeme Smith and Debbie Leung. Typical entanglement of stabilizer states. *Physical Review A—Atomic, Molecular, and Optical Physics*, 74(6):062314, 2006.

[37] Oscar C. O. Dahlsten and Martin B. Plenio. Entanglement probability distribution of bi-partite randomised stabilizer states. *Quantum Info. Comput.*, 6(6):527–538, September 2006.

[38] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347, 2005.

[39] Patrick Hayden, Debbie Leung, Peter W Shor, and Andreas Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250:371–391, 2004.

[40] Emanuel Knill. Approximation by quantum circuits. *arXiv preprint quant-ph/9508006*, 1995.

[41] Joseph Emerson, Yaakov S. Weinstein, Marcos Saraceno, Seth Lloyd, and David G. Cory. Pseudo-random unitary operators for quantum information processing. *Science*, 302(5653):2098–2100, 2003.

[42] Aram W Harrow and Richard A Low. Random quantum circuits are approximate 2-designs. *Communications in Mathematical Physics*, 291:257–302, 2009.

[43] Fernando GSL Brandao, Aram W Harrow, and Michał Horodecki. Local random quantum circuits are approximate polynomial-designs. *Communications in Mathematical Physics*, 346:397–434, 2016.

[44] Aram W Harrow and Saeed Mehraban. Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates. *Communications in Mathematical Physics*, 401(2):1531–1626, 2023.

[45] Nicholas Hunter-Jones. Unitary designs from statistical mechanics in random quantum circuits. *arXiv preprint arXiv:1905.12053*, 2019.

[46] Jonas Haferkamp and Nicholas Hunter-Jones. Improved spectral gaps for random quantum circuits: Large local dimensions and all-to-all interactions. *Physical Review A*, 104(2):022417, 2021.

[47] Jonas Haferkamp. Random quantum circuits are approximate unitary t-designs in depth $O(nt^{5+o(1)})$. *Quantum*, 6:795, 2022.

[48] Shivan Mittal and Nicholas Hunter-Jones. Local random quantum circuits form approximate designs on arbitrary architectures. *arXiv preprint arXiv:2310.19355*, 2023.

[49] Toshihiro Yada, Ryotaro Suzuki, Yosuke Mitsuhashi, and Nobuyuki Yoshioka. Non-haar random circuits form unitary designs as fast as Haar random circuits. *Physical Review Letters*, 136(3):030401, 2026.

[50] John Watrous. *The theory of quantum information*. Cambridge University Press, 2018.

[51] Scott Aaronson and Lijie Chen. Complexity-theoretic foundations of quantum supremacy experiments. In *Proceedings of the 32nd Computational Complexity Conference*, CCC '17, Dagstuhl, DEU, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[52] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, 2018.

[53] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Average-case complexity versus approximate simulation of commuting quantum computations. *Physical review letters*, 117(8):080501, 2016.

[54] Adam Bouland, Bill Fefferman, Zeph Landau, and Yunchao Liu. Noise and the frontier of quantum supremacy. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1308–1317. IEEE, 2022.

[55] Adam Bouland, Ishaun Datta, Bill Fefferman, and Felipe Hernández. Exponential improvements to the average-case hardness of bosonsampling. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 912–933. IEEE, 2025.

[56] Boaz Barak, Chi-Ning Chou, and Xun Gao. Spoofing Linear Cross-Entropy Benchmarking in Shallow Quantum Circuits. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[57] Alexander M Dalzell, Nicholas Hunter-Jones, and Fernando GSL Brandão. Random quantum circuits anticoncentrate in log depth. *PRX Quantum*, 3(1):010333, 2022.

[58] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. Random unitaries in extremely low depth. *Science*, 389(6755):92–96, 2025.

[59] Nicholas LaRacuente and Felix Leditzky. Approximate unitary k-designs from shallow, low-communication circuits. *Communications in Mathematical Physics*, 407(3):51, 2026.

[60] Markus Heinrich, Jonas Haferkamp, Ingo Roth, and Jonas Helsen. Anti-concentration is (almost) all you need. *arXiv preprint arXiv:2510.23719*, 2025.

[61] Hanchen Liu, Tianci Zhou, and Xiao Chen. Measurement-induced entanglement transition in a two-dimensional shallow circuit. *Physical Review B*, 106(14):144311, 2022.

[62] Yimu Bao, Maxwell Block, and Ehud Altman. Finite-time teleportation phase transition in random quantum circuits. *Physical Review Letters*, 132(3):030401, 2024.

[63] Google Quantum AI and Collaborators. Measurement-induced entanglement and teleportation on a noisy quantum processor. *Nature*, 622(7983):481–486, 2023.

[64] Max McGinley, Wen Wei Ho, and Daniel Malz. Measurement-induced entanglement and complexity in random constant-depth 2d quantum circuits. *Physical Review X*, 15(2):021059, 2025.

[65] Zak Webb. The Clifford group forms a unitary 3-design. *arXiv preprint arXiv:1510.02769*, 2015.

[66] Huangjun Zhu. Multiqubit Clifford groups are unitary 3-designs. *Physical Review A*, 96(6):062336, 2017.

[67] David P DiVincenzo, Debbie W Leung, and Barbara M Terhal. Quantum data hiding. *IEEE Transactions on Information Theory*, 48(3):580–598, 2002.

[68] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A—Atomic, Molecular, and Optical Physics*, 80(1):012304, 2009.

[69] Scott Aaronson and Yuxuan Zhang. On verifiable quantum advantage with peaked circuit sampling. *arXiv preprint arXiv:2404.14493*, 2024.

[70] Yuxuan Zhang. Complexity and hardness of random peaked circuits. *arXiv preprint arXiv:2510.00132*, 2025.

[71] Hrant Gharibyan, Mohammed Zuhair Mullath, Nicholas E Sherman, Vincent P Su, Hayk Tepanyan, and Yuxuan Zhang. Heuristic quantum advantage with peaked circuits. *arXiv preprint arXiv:2510.25838*, 2025.

[72] David Kremer and Nicolas Dupuis. Efficient classical simulation of heuristic peaked quantum circuits. *arXiv preprint arXiv:2604.21908*, 2026.

[73] Barbara M Terhal and David P DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. *arXiv preprint quant-ph/0205133*, 2002.

[74] 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.

[75] Sergey Bravyi, David Gosset, and Ramis Movassagh. Classical algorithms for quantum mean values. *Nature Physics*, 17(3):337–341, 2021.

[76] Nolan J. Coble and Matthew Coudron. Quasi-polynomial time approximation of output probabilities of geometrically-local, shallow quantum circuits. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 598–609. IEEE, 2021.

[77] Suchetan Dontha, Shi Jie Samuel Tan, Stephen Smith, Sangheon Choi, and Matthew Coudron. Approximating output probabilities of shallow quantum circuits which are geometrically-local in any fixed dimension. In François Le Gall and Tomoyuki Morimae, editors, *17th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2022, July 11-15, 2022, Urbana Champaign, Illinois, USA*, volume 232 of *LIPIcs*, pages 9:1–9:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[78] Svante Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24(3):234–248, 2004.

[79] Anurag Anshu and Tony Metger. Concentration bounds for quantum states and limitations on the QAOA from polynomial approximations. *Quantum*, 7:999, 2023.

[80] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, April 2017.

[81] Johnnie Gray. quimb: A python package for quantum information and many-body calculations. *Journal of Open Source Software*, 3(29):819, 2018.

[82] Sergey Bravyi and Dmitri Maslov. Hadamard-free circuits expose the structure of the Clifford group. *IEEE Transactions on Information Theory*, 67(7):4546–4563, 2021.

[83] Dmitri Maslov and Willers Yang. Cnot circuits need little help to implement arbitrary Hadamard-free Clifford transformations they generate. *npj Quantum Information*, 9(1):96, 2023.

[84] Yinchen Liu. Moments of random quantum circuits and applications in random circuit sampling. Master's thesis, University of Waterloo, 2021.

[85] Sergey Bravyi, David Fattal, and Daniel Gottesman. Ghz extraction yield for multipartite stabilizer states. *Journal of Mathematical Physics*, 47(6), 2006.

[86] Maarten Van den Nest, Jeroen Dehaene, and Bart De Moor. Graphical description of the action of local Clifford transformations on graph states. *Physical Review A*, 69(2):022316, 2004.

[87] Peter Selinger. Generators and relations for n-qubit Clifford operators. *Logical Methods in Computer Science*, 11, 2015.

[88] Richard Jozsa, Daniel Robb, and William K. Wootters. Lower bound for accessible information in quantum mechanics. *Phys. Rev. A*, 49:668–677, February 1994.

[89] Sergey Bravyi and David Gosset. Improved classical simulation of quantum circuits dominated by Clifford gates. *Physical review letters*, 116(25):250501, 2016.

[90] Antonio Anna Mele. Introduction to Haar measure tools in quantum information: A beginner's tutorial. *Quantum*, 8:1340, 2024.

[91] David Gosset, Daniel Grier, Alex Kerzner, and Luke Schaeffer. Fast simulation of planar Clifford circuits. *Quantum*, 8:1251, 2024.

[92] Isaac L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11–12):2455–2467, November 1997.

[93] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: The two-bit quantum gate. *Physical Review Letters*, 78:390–393, January 1997.

[94] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Jonathan Mizrahi, Jonathan D. Sterk, and Peter Maunz. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit, 2013.

[95] Daniel Greenbaum. Introduction to quantum gate set tomography, 2015.

[96] Erik Nielsen, John King Gamble, Kenneth Rudinger, Travis Scholten, Kevin Young, and Robin Blume-Kohout. Gate Set Tomography. *Quantum*, 5:557, October 2021.

[97] Erik Nielsen, Stefan Seritan, Timothy Proctor, Kenneth Rudinger, Kevin Young, Antonio Russo, Robin Blume-Kohout, Robert Payton Kelly, John King Gamble, and Lucas Saldyt. pygstio/pygsti: Version 0.9.13.1, March 2025.

[98] Haggai Landa, Dekel Meirom, Naoki Kanazawa, Mattias Fitzpatrick, and Christopher J Wood. Experimental Bayesian estimation of quantum state preparation, measurement, and gate errors in multiqubit devices. *Physical Review Research*, 4(1):013199, 2022.

[99] Shahnawaz Ahmed, Fernando Quijandría, and Anton Frisk Kockum. Gradient-descent quantum process tomography by learning Kraus operators. *Physical Review Letters*, 130(15):150402, 2023.

[100] Yigal Ilin and Itai Arad. Learning a quantum channel from its steady-state. *New Journal of Physics*, 26(7):073003, July 2024. Publisher: IOP Publishing.

[101] Pablo Viñas and Alejandro Bermudez. Microscopic parametrizations for gate set tomography under coloured noise. *npj Quantum Information*, 11(1):23, 2025.

[102] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. OUP Oxford, 2002.

[103] Andrew M Childs, Isaac L Chuang, and Debbie W Leung. Realization of quantum process tomography in nmr. *Physical Review A*, 64(1):012314, 2001.

[104] G Mauro D'Ariano, Martina De Laurentis, Matteo G A Paris, Alberto Porzio, and Salvatore Solimeno. Quantum tomography as a tool for the characterization of optical devices. *Journal of Optics B: Quantum and Semiclassical Optics*, 4(3):S127–S132, April 2002.

[105] Nicolas Boulant, Timothy F Havel, Marco A Pravia, and David G Cory. Robust method for estimating the Lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points. *Physical Review A*, 67(4):042322, 2003.

[106] Michael Marc Wolf, Jens Eisert, Toby S Cubitt, and J Ignacio Cirac. Assessing non-Markovian quantum dynamics. *Physical review letters*, 101(15):150402, 2008.

[107] Toby S Cubitt, Jens Eisert, and Michael M Wolf. The complexity of relating quantum channels to master equations. *Communications in Mathematical Physics*, 310(2):383–418, 2012.

[108] Haimeng Zhang, Bibek Pokharel, E.M. Levenson-Falk, and Daniel Lidar. Predicting non-Markovian superconducting-qubit dynamics from tomographic reconstruction. *Physical Review Applied*, 17(5), May 2022.

[109] Gabriel O. Samach, Ami Greene, Johannes Borregaard, Matthias Christandl, Joseph Barreto, David K. Kim, Christopher M. McNally, Alexander Melville, Bethany M. Niedzielski, Youngkyu Sung, Danna Rosenberg, Mollie E. Schwartz, Jonilyn L. Yoder, Terry P. Orlando, Joel I-Jan Wang, Simon Gustavsson, Morten Kjaergaard, and William D. Oliver. Lindblad tomography of a superconducting quantum processor. *Physical Review Applied*, 18(6), December 2022.

[110] Dmitrii Dobrynin, Lorenzo Cardarelli, Markus Müller, and Alejandro Bermudez. Compressed-sensing lindbladian quantum tomography with trapped ions. *Quantum science and technology*, 10(4):045041, 2025.

[111] Emilio Onorati, Tamara Kohler, and Toby S. Cubitt. Fitting quantum noise models to tomography data. *Quantum*, 7:1197, December 2023.

[112] Emilio Onorati, Tamara Kohler, and Toby S Cubitt. Fitting time-dependent Markovian dynamics to noisy quantum channels. *arXiv preprint arXiv:2303.08936*, 2023.

[113] Lorenzo Pastori, Tobias Olsacher, Christian Kokail, and Peter Zoller. Characterization and verification of trotterized digital quantum simulation via Hamiltonian and Liouvillian learning. *PRX Quantum*, 3(3):030324, 2022.

[114] Liran Shirizly, Grégoire Misguich, and Haggai Landa. Dissipative dynamics of graph-state stabilizers with superconducting qubits. *Physical Review Letters*, 132(1):010601, 2024.

[115] IBM quantum computing. https://www.ibm.com/quantum, 2024. [Online; accessed: 3rd May 2024].

[116] Daniel Stilck França, Liubov A. Markovich, V. V. Dobrovitski, Albert H. Werner, and Johannes Borregaard. Efficient and robust estimation of many-qubit Hamiltonians. *Nature Communications*, 15(1):311, 2024.

[117] Tobias Olsacher, Tristan Kraft, Christian Kokail, Barbara Kraus, and Peter Zoller. Hamiltonian and Liouvillian learning in weakly-dissipative quantum many-body systems. *Quantum Science and Technology*, 10(1):015065, 2025.

[118] Giovanni Cemin, Marcel Cech, Erik Weiss, Stanislaw Soltan, Daniel Braun, Igor Lesanovsky, and Federico Carollo. Machine learning of reduced quantum channels on noisy intermediate-scale quantum devices. *Physical Review A*, 110(5):052418, 2024.

[119] Sohail Reddy, Stefanie Günther, and Yujin Cho. Data-driven characterization of latent dynamics on quantum testbeds. *AVS Quantum Science*, 6(3), 2024.

[120] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[121] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.

[122] Brendan O'Donoghue, Eric Chu, Neal Parikh, and Stephen Boyd. SCS: Splitting conic solver, version 3.2.4. https://github.com/cvxgrp/scs, November 2023.

[123] George C Knee, Eliot Bolduc, Jonathan Leach, and Erik M Gauger. Quantum process tomography via completely positive and trace-preserving projection. *Physical Review A*, 98(6):062336, 2018.

[124] Kyle Gulshen, Joshua Combes, Matthew P. Harrigan, Peter J. Karalekas, Marcus P. da Silva, M. Sohaib Alam, Amy Brown, Shane Caldwell, Lauren Capelluto, Gavin Crooks, Daniel Girshovich, Blake R. Johnson, Eric C. Peterson, Anthony Polloreno, Nicholas C. Rubin, Colm A. Ryan, Alexa Staley, Nikolas A. Tezak, and Joseph Valery. Forest Benchmarking: QCVV using PyQuil, 2019.

[125] Timothy Proctor, Melissa Revelle, Erik Nielsen, Kenneth Rudinger, Daniel Lobser, Peter Maunz, Robin Blume-Kohout, and Kevin Young. Detecting and tracking drift in quantum information processors. *Nature Communications*, 11(1), October 2020.

[126] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3), March 2008.

[127] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60(3):1888, 1999.

[128] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 2002.

[129] Dieter Kraft. Algorithm 733: TOMP–fortran modules for optimal control calculations. *ACM Transactions on Mathematical Software*, 20:262–281, 1994.

[130] Steven G. Johnson. The NLopt nonlinear-optimization package. https://github.com/stevengj/nlopt, 2007.

[131] Qiskit. https://www.ibm.com/quantum/qiskit, 2024. [Online; accessed: 3rd May 2024].

[132] Thomas Alexander, Naoki Kanazawa, Daniel J Egger, Lauren Capelluto, Christopher J Wood, Ali Javadi-Abhari, and David C McKay. Qiskit pulse: programming quantum computers through the cloud with pulses. *Quantum Science and Technology*, 5(4):044006, August 2020.

[133] J. Burnett, L. Faoro, I. Wisby, V. L. Gurtovoi, A. V. Chernykh, G. M. Mikhailov, V. A. Tulin, R. Shaikhaidarov, V. Antonov, P. J. Meeson, A. Ya. Tzalenchuk, and T. Lindström. Evidence for interacting two-level systems from the 1/f noise of a superconducting resonator. *Nature Communications*, 5(1), June 2014.

[134] S. E. de Graaf, L. Faoro, L. B. Ioffe, S. Mahashabde, J. J. Burnett, T. Lindström, S. E. Kubatkin, A. V. Danilov, and A. Ya. Tzalenchuk. Two-level systems in superconducting quantum devices due to trapped quasiparticles. *Science Advances*, 6(51), December 2020.

[135] Chad T. Rigetti and Michel H. Devoret. Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. *Physical Review B*, 81:134507, 2010.

[136] Jerry M. Chow, A. D. Córcoles, Jay M. Gambetta, Chad Rigetti, B. R. Johnson, John A. Smolin, J. R. Rozen, George A. Keefe, Mary B. Rothwell, Mark B. Ketchen, and M. Steffen. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Physical Review Letters*, 107(8), August 2011.

[137] Daniel A. Lidar. *Review of Decoherence-Free Subspaces, Noiseless Subsystems, and Dynamical Decoupling*, pages 295–354. John Wiley and Sons, Ltd, 2014.

[138] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

[139] Alireza Shabani and Daniel A Lidar. Completely positive post-Markovian master equation via a measurement approach. *Physical Review A*, 71(2):020101, 2005.

[140] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76(5):722–725, January 1996.

[141] E. Knill. Fault-tolerant postselected quantum computation: Threshold analysis, 2004.

[142] O. Kern, G. Alber, and D. L. Shepelyansky. Quantum error correction of coherent errors by randomization. *The European Physical Journal D*, 32(1):153–156, January 2005.

[143] Michael R. Geller and Zhongyuan Zhou. Efficient error models for fault-tolerant architectures and the Pauli twirling approximation. *Physical Review A*, 88(1), July 2013.

[144] Joel J. Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Physical Review A*, 94(5), November 2016.