Classical Simulation for Quantum Random Circuit Sampling in the Regime of Anti-concentration

---

A Thesis
Presented to
The Division of Mathematical and Natural Sciences
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Xinran Liu

May 2024

Approved for the Division
(Mathematics)

_____

James Pommersheim

Acknowledgements

Thank you, Jamie, for your guidance and support throughout this journey. You are the best thesis adviser one could ever ask for.

Thank you to all my professors I've taken classes with. The profession and passion you brought to your classrooms have, and will continue to inspire me as an insistent leaner.

Thank you to my friends Elle, Guangyi, and Louise. It is our intellectual sparks and giggles that have carried me through the past four years. I remember Reed for you.

Thank you Emily, for encouraging me to always stay true to myself. As I take the stage today—and in all of life's stages—I carry your spirit in my heart.

Last but not least, thank you to mom and dad for always being there during my best and worst times. I feel so loved by you and will love you with all my heart.

Table of Contents

List of Figures

Introduction

Following Turing's conceptualization of the programmable Universal Turing Machine and von Neumann's theoretical model for a practical computer, the development of computer hardware has advanced at a remarkable pace. In 1965, Gordon Moore formalized this exponential growth through what later became known as Moore's Law. It posits that the computational power of computers, relative to cost, doubles approximately every two years.

Although Moore's Law has been applicable for several decades, starting around mid-2010s, it began facing significant challenges due to the physical limitations of silicon-based transistors. Consequently, the computational paradigm is gradually shifting towards quantum computing, first proposed by Feynman for the efficient simulation of (quantum) matter. There are two principles in the field: quantum complexity and quantum error correction.

Quantum complexity is the basis of why we think quantum computing is powerful. It stands for the extravagant complexity of highly entangled quantum states that is impossible for classical data to describe. Innovative quantum algorithms emerge that solve problems that are believed to be hard for classical computers. For example, Shor's algorithm for finding the prime factors of a large composite integer [1] and Grover's algorithm for unstructured search [2]. Thus, theoretically, there is potentially a exponential separation in the computational capabilities of quantum computers compared to classical ones.

However, the realization of a large-scale quantum computer remains largely theoretical. This challenge arises because isolating a physical qubit—so that its state remains stable and undisturbed by noise—is exceedingly difficult. Noise interferes with the functionality of quantum gates within the circuits, making it hard to build a quantum computer that reliably implements a circuit. To address this, researchers have developed quantum error-correcting codes to protect the quantum system. Yet implementing these error corrections on a scalable basis requires a substantial number of qubits, so an experimental realization of reliable quantum computers is still a distant goal.

Noisy Intermediate-Scale Quantum computers (NISQ) technology, a term coined by Preskill in 2018, refers to quantum computers equipped with 50-100 qubits but com-

promised by imperfect noise control. This is an intermediate step that should be achievable in the near term. The question now becomes, is there a computational task that a NISQ computer can perform efficiently, which would require exponentially more time to complete using any classical computer. This event is specifically referred to quantum supremacy, a term coined by John Preskill in 2012.

The lead candidate is the Random Circuit Sampling (RCS), the task of sampling from the output distribution of random quantum circuits. In 2019, the Google *et al.* team announced a successful implementation named "Sycamore," which claims to be an experimental realization of quantum supremacy specifically for the task of random circuit sampling. It uses a processor with programmable superconducting qubits to create quantum states on 53 qubits. Details on how to build such a high-fidelity processor can be found in the original publication [3].

In this thesis, we explore what the RCS experiment actually entails and review the current literature on whether it can underpin a scalable experimental violation demonstrating quantum supremacy. Chapter 1 provides a brief introduction to several relevant concepts in quantum computation that we will rely on later. Chapter 2 gives a mathematical formulation of the RCS experiment and presents a classical algorithm that challenges the complexity-theoretic hardness of noisy random circuit sampling. Chapter 3 studies the conditions for anti-concentration, a major assumption used to claim the simulation is easy. This thesis concludes by suggesting some future directions.

Chapter 1

Quantum Computation

This chapter provides selected details of the linear algebra formalism of quantum computing and its postulates, with the aim of describing the setting in which the experiment of random quantum circuit sampling takes place. A more comprehensive introduction of the theory of quantum computing can be found in Nielsen and Chaung [4].

1.1 State Space

1.1.1 Notation

In the classical setting, information / state is described using a classical bit $x \in \{0, 1\}$. While in the quantum setting, qubit is the quantum analog of the bit. It is a unit vector in $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{C}^2$ where $\alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1$.

The **Dirac notation** is adopted by physicists to represent a qubit in quantum mechanics. The standard notation for a qubit is $|\psi\rangle$, known as *ket*, where the $|\cdot\rangle$ indicate that the object is a vector. For the standard basis $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for $\mathbb{C}^2$, each corresponds to basis states $\{|0\rangle, |1\rangle\}$ in Dirac notation. Thus, a qubit $|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is any linear combination of the computational basis states $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$, with $|\alpha|^2 + |\beta|^2 = 1$.

The conjugate transpose of $|\psi\rangle$, known as *bra*, is the **dual vector** of the state, written $\langle\psi| = |\psi\rangle^\dagger$. In matrix representation, the dual vector is just a row vector. This notation is useful for representing the inner product between vectors, $(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle$.

1.1.2 Inner product space

We start with setting the proper space for describing a physical system.

Proposition 1 (Space Postulate)**.** *The state space of a closed quantum system is a complex inner product space. That is, a vector space over $\mathbb{C}$ equipped with an inner product $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$ such that for all $|v\rangle, |w\rangle \in V$ and $\lambda \in \mathcal{C}$,*

1. *$\langle v|\lambda w\rangle = \lambda \langle v|w\rangle$ (linearity);*
2. *$\langle v|w\rangle = \langle w|v\rangle^*$ (skew-symmetry);*
3. *$\langle v|v\rangle \geq 0$ with equality iff $|v\rangle = 0$ (positive definiteness).*

Note that in the finite-dimensional complex vector spaces that come up in quantum computation, a Hilbert space is equivalent to an inner product space. Throughout this thesis, we will use the two terms interchangeably.

1.1.3 Outer product

A useful way of representing linear operators is the outer product representation.

Definition 1. *Suppose $|v\rangle \in V$ and $|w\rangle \in W$ where V, W are two inner product spaces. Define $|w\rangle\langle v|$ to be the linear operator from V to W defined by*

$$(|w\rangle\langle v|)|v'\rangle \equiv |w\rangle\langle v|v'\rangle = \langle v|v'\rangle|w\rangle$$

The outer product notation has a nice trace property:

Proposition 2. *Let $|\psi\rangle, |\phi\rangle \in V$. Then*

$$\mathrm{Tr}(|\psi\rangle\langle\phi|) = \langle\psi|\phi\rangle \tag{1.1}$$

Proof: Let $\{|i\rangle\}$ be an orthonormal basis of V. Then

$$\mathrm{Tr}(|\psi\rangle\langle\phi|) = \sum_i \langle i|(|\psi\rangle\langle\phi|)|i\rangle = \sum_i \langle i|\psi\rangle\langle i|\phi\rangle^* = \langle\psi|\phi\rangle$$

1.2 Evolution

We next describe how a quantum system change over time. In a closed quantum system, quantum states may undergo only unitary transformations.

Definition 2 (Unitary transformation)**.** *A unitary transformation on a Hilbert space $\mathcal{H}$ is a bijective linear transformation $U : \mathcal{H} \to \mathcal{H}$ that preserves inner products:*

$$< \psi, \phi >=< U\psi, U\phi > \text{ for all } < \psi, \phi > \in \mathcal{H}$$

Proposition 3 (Evolution Postulate)**.** *The evolution of a closed quantum system is described by a unitary transformation on its Hilbert spaces. Suppose a quantum system is in state $|\psi\rangle$ at time t_1, then for all $t_2 \geq t_1$, if the state is $|\psi'\rangle$ at later time t_2, there exists some unitary transform U such that*

$$|\psi'\rangle = U|\psi\rangle \tag{1.2}$$

1.3 Quantum Measurement

Proposition 4 (Measurement Postulate)**.** *Quantum measurements are described by a collection* $\{M_m\}$ *of measurement operators such that* $\sum_m M_m^\dagger M_m = I$*. The index* m *refers to the measurement outcome that may occur in the experiment. If the state of the quantum system is* $|\psi\rangle$ *before the measurement, then the probability that result* m *occurs is given by*

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle \tag{1.3}$$

and the state collapses to

$$\frac{M_m |\psi\rangle}{\langle\psi| M_m^\dagger M_m |\psi\rangle}. \tag{1.4}$$

1.3.1 POVM measurements

Usually for most of the experiments, including the RCS experiment, the quantum system is measured only once, upon conclusion of the experiment. In such instances, the post-measurement state of the system is of little interest, with the main item of interest being the probabilities of the respective measurement outcomes. POVM formalism is a special case for general measurement that simplifies the measurement postulate:

Definition 3. *A POVM is a collection* $\{E_m\}$ *of positive operators defined by*

$$E_m \equiv M_m^\dagger M_m \tag{1.5}$$

such that $\sum_m E_m = I$*. The probability of measuring outcome* m *may now be written as*

$$p(m) = \langle\psi| E_m |\psi\rangle = \mathrm{Tr}(E_m |\psi\rangle \langle\psi|) \tag{1.6}$$

1.4 Density Operator

We next introduce an alternate formalism of quantum mechanics, the density operator. This alternate formulation is mathematically equivalent to the state vector approach, but it provides a more convenient language for describing a quantum system whose state is not known. More precisely, suppose a quantum system is in one of a number of states $|\psi_i\rangle$, where i is an index, with respective probabilities p_i. We call $\{p_i, |\psi_i\rangle\}$ an ensemble of pure states. The density operator for the system is defined by

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i| \tag{1.7}$$

The density operator is characterized by the following theorem:

Theorem 1. *A linear operator* ρ *acting on a Hilbert space* $\mathcal{H}$ *is a density operator if and only if*

1. $\mathrm{Tr}(\rho) = 1$;

2. ρ is a positive operator.

Proof: ($\Rightarrow$) Suppose $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \in \mathcal{H}$. Then

$$\begin{aligned}\mathrm{Tr}(\rho) &= \mathrm{Tr}\left(\sum_i p_i |\psi_i\rangle \langle\psi_i|\right) \\ &= \sum_i p_i \,\mathrm{Tr}(|\psi_i\rangle \langle\psi_i|) \\ &= \sum_i p_i \\ &= 1.\end{aligned}$$

And for any $\phi \in \mathcal{H}$,

$$\begin{aligned}\langle\phi| \psi |\phi\rangle &= \sum_i p_i \langle\phi|\psi_i\rangle \langle\psi_i|\phi\rangle \\ &= \sum_i p_i \|\langle\phi|\psi_i\rangle\|^2 \\ &\geq 0\end{aligned}$$

Thus ρ is a positive operator.

1.4.1 Pure State v.s. Mixed State

Note that probabilistic mixture of quantum states is different from their superposition. We can show this simply by the examples below:

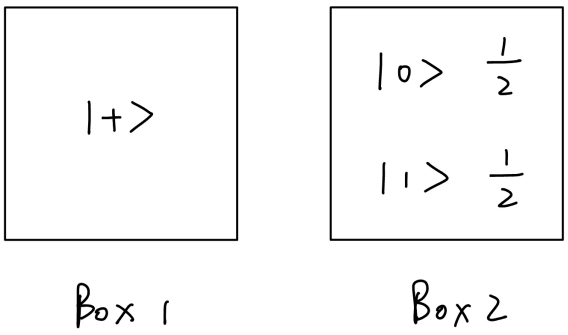


Figure 1.1: Examples of pure and mixed states. Box 1 is a pure state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Box 2 is a mixed state with one half probability of being either $|0\rangle$ or $|1\rangle$.

For Box 1, the quantum system is a superposition of two states $|0\rangle$ and $|1\rangle$ with equal probability amplitudes results in a pure state $|+\rangle$. The density operator of the pure state $|+\rangle$ is just an outer product with itself:

$$\rho = |+\rangle \langle+| = \frac{1}{2}\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}$$

For Box 2, a physical system is prepared to be either in state $|0\rangle$ or $|1\rangle$ with equal probability. The density operator is thus

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note that the density operators for the two different ensembles in Boxes 1 and 2 are the same. This leads to the following theorem.

1.4.2 Schrödinger–HJW theorem

A given density operator does not uniquely determine which ensemble of pure states gives rise to it; in general there are infinitely many different ensembles generating the same density matrix. Those cannot be distinguished by any measurement.

Theorem 2 (Schrödinger–HJW)**.** *Let $\{p_j, |\psi_j\rangle\}$ be an ensemble, then for any unitary matrix U, then ensemble $\{q_i, |\phi_i\rangle\}$ defined by*

$$\sqrt{q_i}\,|\phi_i\rangle = \sum_j U_{ij}\sqrt{p_j}\,|\psi_i\rangle \tag{1.8}$$

will give rise to the same density operator.

1.4.3 Evolution revisited

As we discussed in proposition 2, the evolution of a closed quantum system is described by the unitary operator U where $|\psi\rangle \xrightarrow{U} U|\psi\rangle$. If we use the density operator to represent the quantum system, the evolution postulate enforces the following:

Definition 4 (Unitary Channel)**.** *A unitary channel $\mathcal{U}$ is a linear map that maps density matrices to density matrices by preserving the positivity and trace condition of the density matrix.*

Proposition 5. *Suppose a closed quantum system is an ensemble of pure states in an N-dimensional Hilbert space $\mathcal{H}$ with $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. If the system is then an ensemble of pure states with density operator $\rho' = \sum_i p_i |\psi_i'\rangle\langle\psi_i'|$, there exists some unitary channel $\mathcal{U}$ such that*

$$\rho' = \mathcal{U}\rho \quad \textit{where} \quad \mathcal{U}(\cdot) = U(\cdot)U^\dagger \tag{1.9}$$

Proof. By Proposition 2, if a quantum system is in state $|\psi_i\rangle$ with probability p_i at time t_1, then for all $t_2 \geq t_1$, if the system is in state $|\psi_i'\rangle$ with probability p_i in time t_2, there exists U such that $|\psi_i'\rangle = U|\psi_i\rangle$. Note that each pure state in an ensemble

evolves according to the same unitary operator. Thus,

$$\begin{aligned}\rho' &= \sum_i p_i \left|\psi_i'\right\rangle \left\langle\psi_i'\right| \\ &= \sum_i p_i \big(U \left|\psi_i\right\rangle\big)\big(\left\langle\psi_i\right| U^\dagger\big) \\ &= U\big(\sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i'\right|\big)U^\dagger \\ &= U\rho U^\dagger\end{aligned}$$

□

1.4.4 Measurement revisited

Proposition 6. *Suppose* $\{(\left|\psi_i\right\rangle, p_i\}$ *is a mixed state with density operator* $\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi\right)|$. *If we measure the mixed state with POVM* $\{E_m\}$*, then the probability of measuring outcome* m *now becomes*

$$p(m) = \mathrm{Tr}(E_m \rho) \tag{1.10}$$

Proof.

$$\begin{aligned}\Pr(\text{measuring m}) &= \sum_i \Pr(\text{measure m} \mid \text{state is} \left|\psi_i\right\rangle) \cdot \Pr(\text{state is} \left|\psi_i\right\rangle) \\ &= \sum_i \mathrm{Tr}(E_m \left|\psi_i\right\rangle \left\langle\psi_i\right|) p_i \\ &= \mathrm{Tr}\left(E_m \sum_i \left|\psi_i\right\rangle \left\langle\psi_i\right| p_i\right) \\ &= \mathrm{Tr}(E_m \rho)\end{aligned}$$

□

1.4.5 Pauli Operators

For the last part of density operator, we introduce the Pauli matrices, a set of matrices that will be very useful later in the thesis.

Definition 5 (Pauli Matrices)**.** *The Pauli matrices are four* 2×2 *matrices with the following forms and corresponding notations*

$$\sigma_0 \equiv I = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \quad \sigma_1 \equiv \sigma_x \equiv X = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

$$\sigma_2 \equiv \sigma_y \equiv Y = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}, \quad \sigma_3 \equiv \sigma_z \equiv Z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

Note that he Pauli matrices form a complete basis for representing any 2×2 Hermitian matrix, including the density matrix of a qubit state.

Remark 1. *Any density matrix ρ of a qubit state can be expressed using the Pauli matrices as a basis:*

$$\rho = \frac{1}{2}(I + x\sigma_x + y\sigma_y + z\sigma_z) \tag{1.11}$$

where the coefficients are given by $x = \mathrm{Tr}(\rho\sigma_x), y = \mathrm{Tr}(\rho\sigma_y), z = \mathrm{Tr}(\rho\sigma_z)$.

1.5 Quantum noise

1.5.1 Depolarizing channel

As discussed in the Introduction section, there are noise in the actual quantum circuits. For this thesis, we consider a simple noise model, a constant amount of depolarizing noise, applied to each qubit at each time step.

Definition 6 (Depolarizing noise)**.** *The depolarizing channel is a model of a type of quantum noise.*

$$\mathcal{E}(\rho) = (1-\gamma)\rho + \gamma\frac{I}{2} \tag{1.12}$$

For a single qubit, there is probability γ that the qubit is depolarized, that is, the qubit is replaced by the completely mixed state $I/2$. And there is probability $1-\gamma$ that the qubit is left untouched.

1.6 Quantum Circuit

We conclude chapter one with an introduction to the fundamental model of quantum computation, the language of quantum circuits.

1.6.1 Qubit gate

Recall that a single qubit is a vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with $|a|^2 + |b|^2 = 1$. And operations on a qubit states must preserve this norm, and thus are described by unitary matrices in $\mathbb{U}^{2\times 2}$. We call these operations **single-qubit gates**.

The two-qubit and multi-qubit gates enable quantum entanglement between qubits. One of the most useful two-qubit gates is the controlled-NOT (CNOT) gate, which falls under the category of controlled operations. Its operation takes two input qubits, known as the *control qubit* and *target qubit*, with action defined by $|c\rangle|t\rangle \to |c\rangle|c \oplus t\rangle$.

1.6.2 Circuit

Considering quantum systems of n qudits (each with local Hilbert space dimension q), a quantum circuit is a sequence gates. More specifically, it is specified by the **circuit size** and its **circuit diagram**.

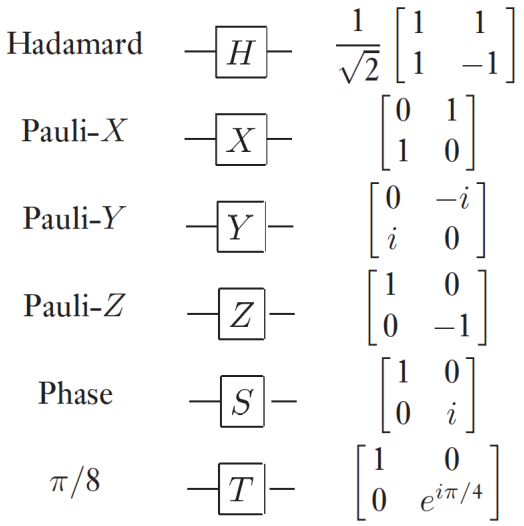


Figure 1.2: Common single qubit gates and their circuit representation.

Definition 7 (circuit size)**.** *The circuit size, denoted s, is the number of gates in the circuit.*

Definition 8 (Circuit diagram)**.** *The circuit diagram determines which qudits each gate acts on. It is chosen according to the circuit architecture, i.e. is a length-s sequence $(A^{(1)}, ..., A^{(s)})$ where each $A^{(i)} \subset \{1, 2, ..., n\}$ indicates which qudits participate in that gate.*

Definition 9 (Circuit depth)**.** *The circuit depth of the circuit diagram, denoted d, is the minimum number of layers of non-overlapping gates needed to implement all s gates in the circuit; that is, the smallest integer such that there exists a sequence $0 = s_0 < s_1 < ... < s_d = s$ where $A^{(t)} \cup A^{(t')} = \emptyset$ whenever $s_j < t < t' < s_{j+1}$.*

There are several common types of circuit architectures. In this paper, we focus on the $1D$ and $2D$ architectures.

Definition 10 (1D architecture)**.** *Assume n is even and $d := 2s/n$ is an integer. The circuit architecture of size s on n qudits is generated by alternating between the two types of layers of $n/2$ non-overlapping nearest-neighbor two-qudit gates on a ring.*

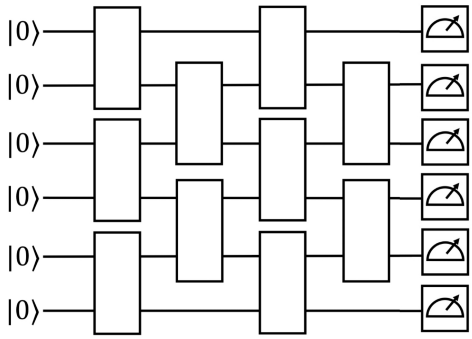


Figure 1.3: The circuit diagram for 1D architecture, with the number of qudit $n = 6$ and the circuit depth $d = 4$. (The gates applied to the boundary are omitted here).

Chapter 2

Classical Simulation of the Random Circuit Sampling

In this chapter, we introduce the random circuit sampling (RCS) experiment in the context of quantum supremacy, and see how we can construct a polynomial-time classical algorithm to simulate RCS under the assumption of anti-concentration.

Notation	Description
n	Number of qubit in the circuit.
p_C	Output distribution of the model circuit C.
p_B	Output distribution of the actual implementation B.
$\tilde{p}(C)$	Output distribution of the model circuit C that includes noise.
$\langle x_j \vert U \vert x_i \rangle$	Inner product between $\vert x_j \rangle$ and $U \vert x_i \rangle$ that represent the transition amplitude from $\vert x_i \rangle to \vert x_j \rangle$.
$p(C, x)$	Probability of measuring output x from the model circuit C with input $\vert 0 \rangle^{\otimes n}$.
s	Pauli path $s \in P_n^{d+1}$.
$f(C, s, x)$	Fourier coefficient of a quantum circuit C with output x over a Pauli path s.
$\vert s \vert$	Hamming weight of a Pauli path s, i.e., the number of non-Identities.
$\bar{q}(C, x)$	Algorithm-computed probability of measuring output x from the circuit C.

2.1 The RCS experiment

This section will present mathematical formulation of the the task used to demonstrate quantum supremacy: sampling from a pseudo-random quantum circuit. The experiment is proposed by the Martinis-Google group, with a goal to "simulate" a quantum circuit; specifically, a random quantum circuit so that it's potentially hard for classical algorithms. We first lay out the conceptual plan for demonstrating quantum supremacy on the basis of random circuit sampling:

1. Consider a random quantum circuit C on a fix circuit diagram as our model. Initialize n input qubits to $|0\rangle$ and measure the final output.
2. This would give us a probability distribution over $\{0,1\}^n$, denoted p_C.
3. Declare that the computational task is to build a physical implementation B of C that generates output distribution p_B that is similar to p_C.
4. Assert that no classical circuit can have output distribution matching p_C.

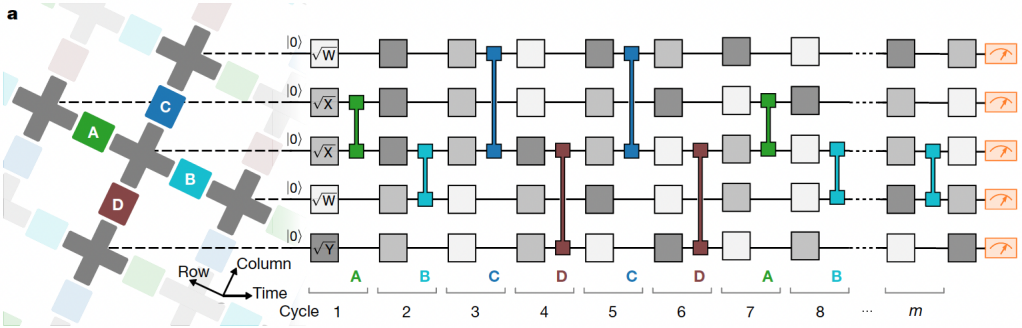


Figure 2.1: A quantum circuit instance used in the Sycamore's RCS experiment.

What the Google team did was to build a quantum computer with $n = 53$ that generates a distribution matching p_C. In this section, we will discuss what this distribution should look like and how we evaluate the performance of the actual implementation. Section 2.2 will present a challenge to the assertion that no classical computer can effectively generate p_C.

2.1.1 Haar measure

We first set the stage for the experiment. Consider the matrix space $\mathbb{U}^{2^n \times 2^n}$ that consists of all unitary $2^n \times 2^n$ matrices. To construct a random quantum circuit is to sample unitary matrices from $\mathbb{U}^{2^n \times 2^n}$. Note that the "random" implicitly requires uniformality. That is, we must sample from a uniform distribution over the space, so that there is no preference for any given unitary matrix. To ensure our unitary matrices are sampled uniformly, we define the **Haar measure** on the space $\mathbb{U}^{2^n \times 2^n}$.

Definition 11 (Haar measure)**.** *A Haar measure is a Borel measure η in a locally compact topological group X such that $\eta(U) > 0$ for every non-empty Borel set U, and $\eta(xU) = \eta(U)$ for every Borel set U.*

Remark 2. *It is shown in [5] that for every locally compact topological group X, there exists a Haar measure η. Given $\mathbb{U}^{2^n \times 2^n}$ is locally compact, we have the left invariance property*

$$\eta(U_0 U) = \eta(U) \tag{2.1}$$

where $U \in \mathbb{U}^{2^n \times 2^n}$. Note that the left invariance property ensures uniformity, i.e., no unitary matrix is more likely to be sampled than any other matrix.

We thus can obtain a random unitary by sampling from the Haar measure.

2.1.2 The Porter-Thomas distribution

After sampling a unitary U from the Haar measure that defines our quantum circuit C, which, upon the initial qubit state $|\psi_0\rangle$, outputs a final state $|\psi\rangle = U|\psi_0\rangle$. We then perform a global measurement of all n qubits in the computational basis (or any other basis obtained from local operations). This will give us a bitstring $x \in \{0,1\}^n$. The probability of observing x is given by

$$p_U(x) = |\langle x| U |\psi_0\rangle|^2 \tag{2.2}$$

One might expect that, for a randomly chosen unitary operator, the expected value of the output distribution might be a uniform distribution. However, as we will show, because of the uniformity of the Haar measure, $p(x)$ is distributed according to the *Porter-Thomas distribution*.

Proposition 7. *Let $\mathcal{P}(p)$ be the probability density function over the continuous variable $p = p(x)$, we have*

$$\mathcal{P}(p) = Ne^{-Np} \tag{2.3}$$

where $N = 2^n$ is the dimension of the Hilbert space.

Proof. Sampling a unitary U from the Haar measure and applying it to some initial state $|\psi_0\rangle$ results in $|\psi\rangle$. Suppose we have an n-qubit system, $|\psi\rangle$ is a random state in the 2^n dimensional Hilbert space $\mathcal{H}$. In the computational basis, we can write this random state as

$$|\psi\rangle = \sum_x (a_x + ib_x)|x\rangle \quad \text{with} \quad \sum_x a_x^2 + b_x^2 = 1$$

where x is the orthonormal basis states for $\mathcal{H}$, and $a_x, b_x \in \mathbb{R}$. The probability of measuring some basis state x_0 is given by $p(x_0) = |\langle x_0|\psi\rangle|^2 = a_{x_0}^2 + b_{x_0}^2$.

The probability density function $\mathcal{P}(p)$ considers the probability of obtaining $p(x_0)$. It can be calculated by considering the Hilbert space $\mathcal{H}_{p,1}$ that contains states satisfying $p(x_0) = a_{x_0}^2 + b_{x_0}^2$ and the normalization condition, and the Hilbert space $\mathcal{H}_1$ that contains states satisfying only the normalization condition. Computing the volumes of these subspaces, and dividing should give us an estimate of $\mathcal{P}(p)$:

$$\mathcal{P}(p) = \frac{\mathrm{Vol}(\mathcal{H}_{p,1})}{\mathrm{Vol}(\mathcal{H}_1)} \to Ne^{-Np} \tag{2.4}$$

A detailed derivation of Porter-Thomas distribution through calculating the integral volumes can be found in [6]. □

Remark 3. *Note that the expected value*

$$\mathbb{E}[p(x)] = \int dp\ p(x) Ne^{-Np(x)} = \frac{Exp(1)}{N} \neq 2^{-n} \tag{2.5}$$

where Exp(1) denotes independent exponential random variables with pdf $f(x) = e^{-x}$. Half of the strings will have probability $\leq \ln 2/N$, and half will have probability $\geq \ln 2/N$. The computational task is then to determine which strings are heavy and which strings are light.

2.1.3 Cross-entropy benchmarking

Suppose we have an actual implementation B that generate some probability distribution p_B, we need a metric to evaluate how close it is from the ideal probability distribution p_C. The Google team chooses to use the linear cross-entropy benchmarking (XEB):

For a given circuit C, we collect several measured bitstrings x_i and compute the probability $p_B(x_i)$ for each. This statistical measure calculates the mean of the simulated probabilities of the measured bitstrings as follows:

$$\mathcal{F}_{\text{XEB}} = 2^n \langle P(x_i) \rangle_i - 1 = 2^n \mathop{\mathbb{E}}_{C} \sum_{x_i} p_B(x_i) p_C(x_i) - 1 \tag{2.6}$$

where n is the number of qubits.

Note that we could not directly compute the ideal probability $p_C(x_i)$ for any large-scale circuit C. Instead, the Google team breaks the circuit up into two patches of qubits and efficiently calculate the probability distributions for each patch using a Schrödinger method. Then "extrapolate" to gain an good estimate of the full circuit probability distribution p_C.

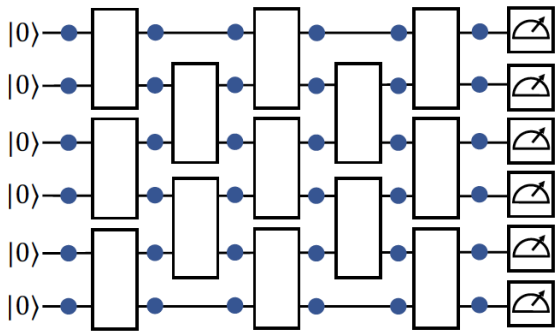
Remark 4. *Values of $\mathcal{F}_{XEB}$ between 0 and 1 correspond to the probability that no error has occurred while running the circuit. When there are no errors in the quantum circuit, the distribution of probabilities is exactly p_C, and sampling from this distribution will produce $\mathcal{F}_{XEB} = 1$. On the other hand, sampling from the uniform distribution will give $\langle P(x_i) \rangle_i = 1/2^n$, thus $\mathcal{F}_{XEB} = 0$.*

The Sycamore processor demonstrated its capabilities by executing ten different random circuit instances, each using 53 qubits across 20 cycles, and collecting 30×10^6 samples in 200 seconds. From this data, the calculated cross-entropy benchmarking (XEB) value was approximately 0.0024.
Quantum supremacy relies on the conjecture that even this very small XEB achieved in the experiment was a classically difficult computational task. This leads us to the next section.

2.2 Simulating noisy RCS

The physical realization of Sycamore and the conjecture above has motivated researchers to challenge the complexity-theoretic hardness of simulating both ideal quantum circuits and circuits with pre-assumed noise model. ([7]). Recent work [8] studies the classical complexity of RCS in the presence of a constant rate of noise per gate, and proposes a polynomial-time classical algorithm to simulate the ideal distribution.

This section presents the main result of the original paper:



(b) Noisy RCS

Figure 2.2: The circuit instance for noisy RCS. Each white box is an independent Haar random 2-qubit gate, with an arbitrarily small constant amount of depolarizing noise is applied to each qubit at each step, which generates a noisy output distribution $\tilde{p}(C)$.

Theorem 3. *Assuming anti-concentration, there exists a classical algorithm that, on input a random circuit C on any fixed architecture, outputs a sample from a distribution that is close to the noisy output distribution $\tilde{p}(C)$ in total variation distance.*

To understand how a classical algorithm can compute the output distribution, we first evaluate this probability distribution more closely.

2.2.1 Feynman path integral

For the probability distribution $p(C, x) := |\langle x| C |0^n\rangle|^2$, we use the concept of the **Feynman path integral** to evaluate this amplitude: In quantum mechanics, the Feynman path integral sums over all possible paths that a particle might take between two points, weighted by the exponential of the action. In a quantum circuit, each "path" can be thought of as a sequence of different configurations of the quantum state as it evolves through the gate sequence.

Definition 12 (Feynman path integral)**.** *Consider a quantum circuit $C = U_d...U_1$ acting on n qubits with circuit depth d, where U_i is a layer of 2-qubit gates. The Feynman path integral in the computational basis is expressed as*

$$\langle x| C |0^n\rangle = \sum_{x_1,...,x_{d-1}\in\{0,1\}^n} \langle x| U_d |x_{d-1}\rangle \langle x_{d-1}| U_{d-1} |x_{d-2}\rangle ... \langle x_1| U_1 |0^n\rangle \tag{2.7}$$

where the LHS is the output amplitude of the quantum circuit, and the RHS is a sum over exponentially $2^{(nd)}$ paths.

Since we have a random circuit, each term in the sum contributes equally to the total amplitude. For a poly-time algorithm, we can only simulate poly(n) random paths,

too small fraction for strong correlation with the ideal distribution.

However, if we switch to the Pauli operators, the paths are no longer uniform. If fact, as we will show, due to noise, the contribution of a path decays exponentially with the number of non-identities Pauli matrices in the path.

2.2.2 Pauli path integral

The first step is to switch from vector basis to operator basis. That is, instead of considering a quantum circuit as applying unitary matrices to vectors, we think of it as applying unitary channels density matrices.

Next, instead of computational basis operators, we consider normalized Pauli operators

$$P_n := \{X/\sqrt{2}, Y/\sqrt{2}, Z/\sqrt{2}\}^{\otimes n}$$

as an operator basis and decompose density matrices into a linear combination of Pauli operators.

Remark 5. *The quantity $\langle x_j| U |x_i\rangle$ is equivalent to $\mathrm{Tr}\big(s_j U s_i U^\dagger\big)$, where the first term is the transition amplitude from x_i to x_j for x_i, x_j in the vector basis, while the second term is the transition amplitude from s_i to s_j for $s_i, s_j \in P_n$.*

We thus can define the Pauli path integral by rewriting the Feynman path integral in the Pauli basis.

Definition 13 (Pauli path integral)**.** *Let $C = U_d ... U_1$ be a quantum circuit acting on n qubits, where U_i is a layer of 2-qubit gates and d is circuit depth. let $p(C, x) := |\langle x| C |0^n\rangle|^2$ be the output probability distribution. The Pauli path integral is written as*

$$\begin{aligned} p(C,x) &= \sum_{s_0,\ldots,s_d \in P_n} \mathrm{Tr}(|x\rangle\langle x| s_d)\, \mathrm{Tr}\Big(s_d U_d s_{d-1} U_d^\dagger\Big) ... \mathrm{Tr}\Big(s_1 U_1 s_0 U_1^\dagger\Big)\, \mathrm{Tr}(s_0 |0^n\rangle\langle 0^n|) \\ &= \sum_{s_0,\ldots,s_d \in P_n} \langle\langle x|s_d\rangle\rangle \langle\langle s_d|\mathcal{U}_d|s_{d-1}\rangle\rangle \cdots \langle\langle s_1|\mathcal{U}_1|s_0\rangle\rangle \langle\langle s_0|0^n\rangle\rangle \end{aligned}$$

with each term of the sum corresponds to a Pauli path $s \in P_n^{d+1}$.

Definition 14 (Fourier coefficient)**.** *Each term of the sum in the Pauli path integral corresponds to a Pauli path $s \in P_n^{d+1}$. Define each term as the **Fourier coefficient** of a quantum circuit C with output x over Pauli path s as*

$$f(C, s, x) := \mathrm{Tr}(|x\rangle\langle x| s_d)\, \mathrm{Tr}\Big(s_d U_d s_{d-1} U_d^\dagger\Big) ... \mathrm{Tr}\Big(s_1 U_1 s_0 U_1^\dagger\Big)\, \mathrm{Tr}(s_0 |0^n\rangle\langle 0^n|) \tag{2.8}$$

thus the output probability is the sum

$$p(C, x) = \sum_{s_0,\ldots,s_d \in P_n} f(C, s, x) \tag{2.9}$$

We next show the **orthogonality property** of Fourier coefficients which will be useful later. Recall that all the unitaries in the circuit are drawn from the Haar random 2-qubit gate set. The randomness of the gate set gives us the following lemma.

Lemma 1 (Gate-set orthogonality)**.** *Let $U \in \mathbb{U}(4)$ be a Haar random 2-qubit gate, and $p, q, r, s \in P_2$. Then*

$$\mathop{\mathbb{E}}_{U\sim\mathbb{U}(4)} \langle\langle p|\mathcal{U}|q\rangle\rangle\langle\langle r|\mathcal{U}|s\rangle\rangle = 0 \quad \textit{if} \quad p \neq r \ \textit{or}\ q \neq s. \tag{2.10}$$

The idea is that if we average over two copies of a random unitary, when averaging two copies of the transitions over the Haar random unitaries U, the only terms that do not average to zero are those where the input and output Pauli operators are the same across both copies of the unitary.

The orthogonality of Fourier coefficients thus follows:

Lemma 2 (Orthogonality of Fourier coefficients)**.** *Let C be a random circuit drawn from the Haar random 2-qubit gates. For any Pauli path $s \neq s \in P_n^{d+1}$ and for any output $x \in \{0,1\}^n$, we have*

$$\mathop{\mathbb{E}}_{C\sim\mathcal{D}} [f(C, s, x) f(C, s', x)] = 0 \tag{2.11}$$

Proof. For two different Pauli paths $s \neq s'$, there exists a 2-qubit gate U at some step in the circuit C that get two different inputs. Thus, $f(C, s, x)$ and $f(C, s', x)$, which are the products of a sequence of transition amplitude, each contains an element $\langle\langle p_1|\mathcal{U}|q_1\rangle\rangle$ and $\langle\langle p_2|\mathcal{U}|q_2\rangle\rangle$ such that $p_1 \neq p_2$. The previous lemma implies that

$$\mathop{\mathbb{E}}_{U} \langle\langle p_1|\mathcal{U}|q_1\rangle\rangle\langle\langle p_2|\mathcal{U}|q_2\rangle\rangle = 0$$

Thus the overall expectation of product to two different Pauli paths equals 0. □

2.2.3 Non-uniformality of Pauli path under noise

We now introduce noise to our circuit. Here we consider the depolarizing channel introduced in Section 1.5, a common model for simulations of random circuits in the presence of noise. We restate the definition below:

Definition 15 (Depolarizing Noise)**.** *In a noisy quantum circuit, let*

$$\mathcal{E}(\rho) := (1-\gamma)\rho + \gamma \frac{I}{2} \operatorname{Tr}(\rho) \tag{2.12}$$

be the single-qubit depolarizing noise with strength γ.

Remark 6. *Since Pauli matrices all have trace 0, we have $\mathcal{E}(I) = I$ and $\mathcal{E}(P) = (1-\gamma)P$ where $P \in \{X, Y, Z\}$.*

As defined in Section 2.2, all n qubits are subject to noise at each step in the circuit. We can rewrite the Pauli Path integral to include such noise:

Definition 16 (Pauli path integral for noisy quantum circuits)**.** *For a quantum circuit $C = U_d...U_1$, let $\tilde{C}$ denote the noisy quantum circuit where each qubit in C is subject to depolarizing noise with strength γ in each layer. The output probability distribution is then*

$$\tilde{p}(C, x) := \langle\langle x|\mathcal{E}^{\otimes n}\mathcal{U}_d\mathcal{E}^{\otimes n}...\mathcal{U}_1\mathcal{E}^{\otimes n}|0^n\rangle\rangle \tag{2.13}$$

The path integral for $\tilde{C}$ is again the sum over all possible Pauli paths, and is defined as

$$\tilde{p}(C, x) := \sum_{s\in P_n^{d+1}} \tilde{f}(C, s, x) \tag{2.14}$$

where $\tilde{f}$ is the Fourier coefficient subject to depolarizing noise

$$\tilde{f}(C, s, x) := \langle\langle x|\mathcal{E}^{\otimes n}|s_d\rangle\rangle\langle\langle s_d|\mathcal{U}_d\mathcal{E}^{\otimes n}|s_{d-1}\rangle\rangle \cdots \langle\langle s_1|\mathcal{U}_1\mathcal{E}^{\otimes n}||s_0\rangle\rangle\langle\langle s_0|0^n\rangle\rangle \tag{2.15}$$

Note that the noise does not have any effect on I operator, while incurring a $1-\gamma$ reduction on any other Pauli operator. Given such performance of depolarizing noise on Pauli operators, we introduce another property of a Pauli path, the Hamming weight.

Definition 17 (Hamming weight)**.** *Let $|s|$ be the Hamming weight of a Pauli path s, i.e., the number of non-identity Pauli in s. The definition of the depolarizing noise implies that*

$$\tilde{f}(C, s, x) = (1-\gamma)^{|s|} f(C, s, x) \tag{2.16}$$

Remark 7. *Due to the depolarizing noise γ, the Pauli paths are highly nonuniform, the contribution decays exponentially with the number of non-identities.*

2.2.4 Algorithm overview

Now we are ready to present the algorithm to simulate a noisy random circuit sampling. Our goal is the compute a function $\bar{q}(C, x)$ that is only a small L_1 distance away from the output distribution of the noisy circuit $\tilde{p}(C, x)$:

$$\Delta := ||\tilde{p} - \bar{q}||_1 := \sum_{x\in\{0,1\}^n} |\tilde{p}(C, x) - \bar{q}(C, x)| \tag{2.17}$$

We will choose a sufficient truncating parameter ℓ. Then, given input quantum circuit C and expected output $x \in \{0, 1\}^n$, the algorithm outputs an approximation of $\tilde{p}(C, x)$ by summing its low-degree Fourier coefficients:

2.2.5 Bounding the truncating error

It turns out that by choosing $\ell = O(\log 1/\epsilon)$, we can bound the total variance distance. The proof requires the assumption of anti-concentration.

1: $q \leftarrow 0$
2: **for** all legal Pauli path s with $|s| \leq \ell$ **do**
3: calculate $f(C, s, x)$
4: $q \leftarrow q + (1-\gamma)^{|s|} f(C, s, x)$
5: **end for**
6: **return** q

Definition 18 (Anti-concentration)**.** *We say a distribution over quantum circuits $\mathcal{D}$ satisfies anti-concentration if*

$$\mathop{\mathbb{E}}_{C\sim\mathcal{D}} 2^n \sum_{x\in\{0,1\}^n} p(C,x)^2 = O(1) \tag{2.18}$$

This says the the second moment of the ideal probability distribution, averaging over random quantum circuit, is small. To relate the second moment property with the error of the simulation, we define the Fourier weight of a circuit with respect to the hamming weight.

Definition 19 (Fourier weight)**.** *The Fourier weight of a random circuit C at degree k is defined as*

$$W_k = 2^{2n} \mathop{\mathbb{E}}_{C} \sum_{s\in P_n^{d+1}:|s|=k} f(C,s,0^n)^2 \tag{2.19}$$

The assumption of anti-concentration implies that the total Fourier weight is upper bounded by a constant.

Lemma 3 (Total Fourier weight)**.** *Let $\mathcal{D}$ be a distribution over quantum circuits that satisfies anti-concentration. The Fourier weights $\{W_k\}$ satisfy*

1. $W_0 = 1$,
2. $W_k = 0, \forall 0 < k \leq d$,
3. $\sum_{k\geq d+1} W_k = O(1)$.

where d is the circuit depth.

Proof. 1. The base case W_0 corresponds to the unique all-identity paths $s = I^{d+1}$, thus

$$W_0 = 2^{2n} f(C,s,0^n)^2 = 2^{2n} \cdot (\frac{1}{2^n})^2 = 1$$

2. For a Pauli path s with $|s| = k \in (0, d]$, we know there exists a 2-qubit gate U in the circuit C that brings an identity to a non-identity or vice versa. Thus, $f(C, s, x)$ contains transition amplitude $\langle\langle p|\mathcal{U}|q\rangle\rangle$ where either $p = I^{\otimes 2}/2$ and $q \neq I^{\otimes 2}/2$ or vice versa. We then have $\langle\langle p|\mathcal{U}|q\rangle\rangle = 0$ and thus $W_k = 0$.

3. To bound the total Fourier weights, we start with anti-concentration.

$$\begin{aligned} O(1) &= \mathbb{E}_{C\sim\mathcal{D}} 2^n \sum_{x\in\{0,1\}^n} p(C,x)^2 \\ &= \mathbb{E}_{C\sim\mathcal{D}} 2^n \sum_{x\in\{0,1\}^n} \left(\sum_{s\in\mathbb{F}_2^{d+1}} f(C,s,x) \right)^2 \\ &= \mathbb{E}_{C\sim\mathcal{D}} 2^n \sum_{x\in\{0,1\}^n} \sum_{s,s'\in\mathbb{F}_2^{d+1}} f(C,s,x) f(C,s',x) \end{aligned}$$

By Lemma 2, we know that the products for Fourier coefficients of two different Pauli paths $s \neq s'$ equals 0, then

$$\begin{aligned} O(1) &= \mathbb{E}_{C\sim\mathcal{D}} 2^n \sum_{x\in\{0,1\}^n} \sum_{s\in\mathbb{F}_2^{d+1}} f(C,s,x)^2 \\ &= 2^{2n} \mathbb{E}_{C\sim\mathcal{D}} \sum_{s\in\mathbb{F}_2^{d+1}} f(C,s,0^n)^2 \\ &= 2^{2n} \mathbb{E}_{C\sim\mathcal{D}} \sum_{k\geq 0} \sum_{s\in\mathbb{F}_2^{d+1}:|s|=k} f(C,s,0^n)^2 \\ &= 1 + \sum_{k\geq d+1} W_k. \end{aligned}$$

□

Next, we show that the expected value of the square of the total variation distance is upper bounded by an exponential decay of the Fourier weights.

$$\begin{aligned} \mathop{\mathbb{E}}_{C}[\Delta^2] &= 2^n \mathop{\mathbb{E}}_{C} \left(\sum_{x\in\{0,1\}^n} |\tilde{p}(C,x) - \bar{q}(C,x)| \right)^2 \\ &\leq 2^n \mathop{\mathbb{E}}_{C} \sum_{x\in\{0,1\}^n} (\tilde{p}(C,x) - \bar{q}(C,x))^2 \qquad \text{by Cauchy-Schwarz} \end{aligned}$$

Given the algorithm, the difference between $\tilde{p}(C,x)$ and $\bar{q}(C,x)$ is the sum of Fourier coefficients with hamming weights higher than the truncating parameter. Thus

$$\begin{aligned} \mathop{\mathbb{E}}_{C}[\Delta^2] &\leq 2^n \mathop{\mathbb{E}}_{C} \sum_{x\in\{0,1\}^n} \left(\sum_{s:|s|\geq\ell} (1-\gamma)^{|s|} f(C,s,x) \right)^2 \\ &= 2^n \mathop{\mathbb{E}}_{C} \sum_{x\in\{0,1\}^n} \sum_{s:|s|\geq\ell} (1-\gamma)^{2|s|} f(C,s,x)^2 \qquad \text{orthogonality of } f(C,s,x) \\ &= 2^{2n} \mathop{\mathbb{E}}_{C} \sum_{s:|s|\geq\ell} (1-\gamma)^{2|s|} f(C,s,0^n)^2 \\ &= \sum_{k\geq\ell} (1-\gamma)^{2k} W_k \qquad \text{definition of Fourier weight} \end{aligned}$$

Recall that the Fourier weights satisfy $\sum_{k\geq d+1} W_k = O(1)$, we can derive a upper bound for the total variance distance square as follows:

$$\begin{aligned}
\mathop{\mathbb{E}}_{C}[\Delta^2] &\leq \sum_{k>\ell}(1-\gamma)^{2k}W_k \\
&\leq \sum_{k\geq\ell}(1-\gamma)^{2k}W_k \\
&\leq O(1)\cdot e^{-2\gamma\ell} \qquad\qquad (1-x)^y \approx e^{-xy} \text{ when } x \text{ is small}
\end{aligned}$$

We can thus conclude that by setting $\ell = O(\log 1/\epsilon)$, we have $\Delta \leq \epsilon$ with high probability.

Chapter 3

Anti-concentration in Quantum Architecture

As shown in the previous chapter, the anti-concentration assumption of the circuit is used to bound the total variance of the simulated output distribution. A natural question to ask is, in what cases can we make this assumption. In this chapter, we focus on the original research studying the number of gates required to achieve anti-concentration [9]. We present a detailed walk through for calculating upper bound for collision probability in $1D$ circuit architecture, which appears as Theorem 5 in [9]. The main theorem of this chapter is the following:

Theorem 4. *Consider a fixed one-dimensional random quantum circuit architecture, with the two-qudit gates each drawn from the Haar measure.* $O(n\log(n))$ *circuit gates are sufficient for the circuit to be anti-concentrated.*

We prove this theorem by considering the quantum circuits as a stochastic process, and then introducing the domain wall notation associating with each trajectory.

Notation	Description
n	Number of qubit in the circuit.
s	Circuit size.
d	Circuit depth.
Z	Collision probability.
Z_H	Collision probability with infinite circuit size.
$\vec{\gamma}^{(t)}$	Circuit configuration at time t.
γ	Circuit trajectory $\gamma = (\vec{\gamma}^{(0)}, ..., \vec{\gamma}^{(s)})$.
P_u	Configuration space in unbiased walk.
Λ_u	Initial configuration space in unbiased walk.
P_b	Configuration space in unbiased walk.
Λ_b	Initial configuration space in biased walk.

3.1 Preliminaries

3.1.1 Random Quantum Circuit

We first specify the quantum circuit. Recall the definitions of circuit architecture and circuit size in Section 1.6.

Definition 20 (circuit size)**.** *The circuit size, denoted s, is the number of gates in the circuit.*

Definition 21 (1D architecture)**.** *Assume n is even and $d := 2s/n$ is an integer. The circuit architecture of size s on n qudits is generated by alternating between the two types of layers of $n/2$ non-overlapping nearest-neighbor two-qudit gates on a ring.*

To create a random quantum circuit of size s, we first fix the circuit architecture to be $1D$ architecture, then draw s unitary gates $U^{(t)}$ that determine the circuit instance randomly from the Haar measure. A **quantum circuit instance** is generated by additionally specifying n random $q \times q$ single-qudit matrices that made up the first layer of the circuit. They prepare the initial input to a randomized state.

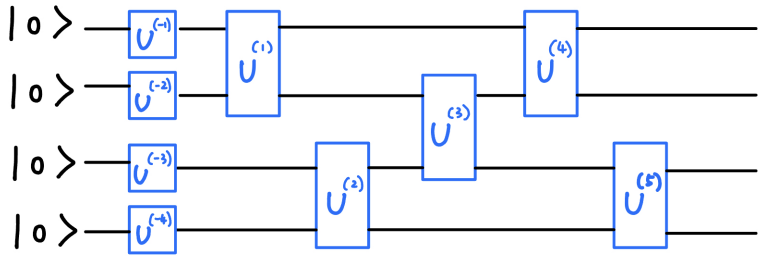


Figure 3.1: The circuit instance for 1D architecture, with the number of qudit $n = 4$ and the circuit size $s = 5$. (The gates applied to the boundary are omitted here).

We denote the global $q^n \times q^n$ unitary operator implemented by the circuit by U, where

$$U := U^{(s)}_{A^{(s)}} ... U^{(1)}_{A^{(1)}} U^{(-1)}_{\{-1\}} ... U^{(-n)}_{\{-n\}} \tag{3.1}$$

3.1.2 Anti-concentration and Collision Probability

Now we proceed to study the anti-concentration property of this circuit, which can be defined based on the **collision probability**. As the name suggested, the collision probability is the probability that two independently drawn outcomes will agree.

Definition 22 (Collision Probability)**.** *For a RQC architecture at a specified qubit number n and circuit size s, the collision probability averaged over the randomly chosen circuit instance U is defined by*

$$Z := \mathop{\mathbb{E}}_{U} \left[\sum_{x \in [q]^n} p_U(x)^2 \right] \tag{3.2}$$

Remark 8. *By symmetry, each of the q^n terms in the sum contributes equally under expectation as long as at least one Haar-random gate acts on each qudit, we can rewrite the expectation by*

$$Z := q^n \mathop{\mathbb{E}}_{U} \Big[p_U(0^n)^2 \Big] \tag{3.3}$$

Anti-concentration captures the property of a circuit that the probability mass is well spread out over all the outcomes. Note that in the most extreme case, if p_U is the uniform distribution, the collision probability is $Z = q^{-n}$, the minimal possible value.

Definition 23 (α-anti-concentrated)**.** *We say that a random quantum circuit architecture is α-anti-concentrated for $0 < \alpha \leq 1$ at circuit size $s(n)$ if there exists n_0 such that whenever $n \geq n_0$,*

$$Z := \mathop{\mathbb{E}}_{U} \left[\sum_{x \in [q]^n} p_U(x)^2 \right] \leq \frac{1}{\alpha} \cdot \frac{1}{q^n} \tag{3.4}$$

That is, the collision probability of the circuit is only a constant factor larger than the minimum value. This is compatible with the general definition of anti-concentration in Definition 12, which requires $(\alpha q^n)^{-1} \in O(1)$.

3.2 RQC as a stochastic process

In this section, we introduce the concept of stochastic process to study how to bound the collision probability. This is based on content in Appendix B of [9].

3.2.1 Re-expressing Z

Given the expectation property under action by a Haar-random unitary, we can analyze the collision probability of RQCs, originally an integral expectation over continuously varying unitary matrices drawn from the Haar measure, as performing the Haar expectation over each local unitary individually.

Note that $p_U(0^n) = \langle 0^n | U | 0^n \rangle \langle 0^n | U^\dagger | 0^n \rangle$, then we have

$$Z := q^n \mathop{\mathbb{E}}_{U} \left[(\langle 0^n | U | 0^n \rangle \langle 0^n | U^\dagger | 0^n \rangle)^2 \right] \tag{3.5}$$

Here we treat the square as having two identical n-qudit circuits on top of each other.

$$Z = q^n \mathop{\mathbb{E}}_{U} \left[\langle 0^n |^{\otimes 2} U^{\otimes 2} | 0^n \rangle^{\otimes 2} \langle 0^n |^{\otimes 2} U^{\dagger^{\otimes 2}} | 0^n \rangle^{\otimes 2} \right] \tag{3.6}$$

By Proposition 1, $\langle \psi | \phi \rangle = \mathrm{Tr}(|\psi\rangle \langle \phi|)$. We now represent our system under density operator notation. We thus have

$$Z = q^n \, \mathrm{Tr} \left[\mathop{\mathbb{E}}_{U} [U^{\otimes 2} | 0^n \rangle \langle 0^n |^{\otimes 2} U^{\dagger \otimes 2}] | 0^n \rangle \langle 0^n |^{\otimes 2} \right] \tag{3.7}$$

where $|0^n\rangle\langle 0^n|^{\otimes 2}$ is two copies of the circuit input state.

Then, for a fixed quantum circuit $U = U^{(s)}_{A^{(s)}}...U^{(1)}_{A^{(1)}}...U^{(-n)}_{(-n)}$, since each unitary is independently chosen according to the Haar measure, we can evaluate the expectation separately. For each gate, denote

$$M^{(t)}[\rho] := \mathop{\mathbb{E}}_{U^{(t)}} \left[{U^{(t)}_{A^{(t)}}}^{\otimes 2} \rho {U^{(t)}_{A^{(t)}}}^{\dagger \otimes 2} \right] \tag{3.8}$$

Then we have a composite representation

$$\mathop{\mathbb{E}}_{U} \left[U^{\otimes 2} |0^n\rangle\langle 0^n|^{\otimes 2} U^{\dagger \otimes 2} \right] = M^{(s)} \circ M^{(s-1)} \circ ... \circ M^{(1)} \circ M^{(-1)} \circ ... \circ M^{(-n)}[|0^n\rangle\langle 0^n|^{\otimes 2}] \tag{3.9}$$

3.2.2 Evaluation of the action of $M^{(t)}$ for all t

We first look at the n $q \times q$ unitaries acting on the input state at time $t \in \{-n, ..., -1\}$. The expression of $M^{(t)}[\rho]$ has the general form: Denote each qudit in two copies of system with the $q^2 \times q^2$ density operator σ. With unitary U chosen from the Haar measure over the set of $q \times q$ unitaries, we define the operator

$$M[\sigma] := \mathop{\mathbb{E}}_{U} \left[U^{\otimes 2} \sigma {U^{\dagger}}^{\otimes 2} \right] \tag{3.10}$$

Then, we see that for any unitary V and any σ, given the left invariance of the Haar measure under substitution $U \to VU$, we can have the commutativity property

$$\begin{aligned} M[\sigma]V^{\otimes 2} &= \mathop{\mathbb{E}}_{U} \left[U^{\otimes 2} \sigma (V^{\dagger} U)^{\dagger \otimes 2} \right] \\ &= \mathop{\mathbb{E}}_{U} \left[(VU)^{\otimes 2} \sigma U^{\dagger \otimes 2} \right] \\ &= V^{\otimes 2} M[\sigma] \end{aligned}$$

Claim 1. *By Schur-Weyl duality [10], any operator on k copies of a system that commutes with $V^{\otimes k}$ for any unitary V must be a linear combination of permutation operators over the k systems.*

Since we have $k = 2$, the only permutation operators are the identity operation I and the swap operation S, where $S(|\psi\rangle \otimes |\phi\rangle) = |\phi\rangle \otimes |\psi\rangle$. Thus, we can write $M[\sigma] = \alpha I + \beta S$.

To solve for α and β, we evaluate the trace of $\mathrm{Tr}[M[\sigma]]$ and $\mathrm{Tr}[M[\sigma]S]$:

$$\begin{aligned}\mathrm{Tr}[M[\sigma]] &= \mathrm{Tr}\left[\mathop{\mathbb{E}}_{U}\left[U^{\otimes 2}\sigma U^{\dagger\otimes 2}\right]\right] \\ &= \mathop{\mathbb{E}}_{U}\mathrm{Tr}\left[U^{\otimes 2}\sigma U^{\dagger\otimes 2}\right] \\ &= \mathop{\mathbb{E}}_{U}\mathrm{Tr}\left[U^{\otimes 2}U^{\dagger\otimes 2}\sigma\right] \\ &= \mathrm{Tr}[\sigma]\end{aligned}$$

At the same time, $\mathrm{Tr}[M[\sigma]] = \mathrm{Tr}[\alpha I + \beta S] = \alpha\,\mathrm{Tr}[I] + \beta\,\mathrm{Tr}[S]$. The diagonal of I is all 1's, so the trace is simply q^2. The diagonal elements of S, $S_{(i,j)(j,i)}$ is 1 only when $i = j$, so the trace equals q. We thus have

$$\mathrm{Tr}[M[\sigma]] = \mathrm{Tr}[\sigma] = \alpha q^2 + \beta q \tag{3.11}$$

Similarly,

$$\mathrm{Tr}[M[\sigma]S] = \mathrm{Tr}[\sigma S] = \alpha q + \beta q^2 \tag{3.12}$$

These help determine α and β in terms of $\mathrm{Tr}(\sigma)$ and $\mathrm{Tr}(\sigma S)$. We can write

$$M[\sigma] = \frac{\mathrm{Tr}(\sigma) - q^{-1}\,\mathrm{Tr}(\sigma S)}{q^2-1} I + \frac{\mathrm{Tr}(\sigma S) - q^{-1}\,\mathrm{Tr}(\sigma)}{q^2-1} S \tag{3.13}$$

Given the explicit formula for operator $M[\sigma]$, we next evaluate the performance of $M^{(t)}$ at each time t. Denote the value of the input state on qudit j by $|0\rangle\langle 0|^{\otimes 2}_{\{j\}}$, and the input state on the other $n-1$ qudits by $\rho_{[n]\setminus\{j\}}$. Then the action of $M^{(-j)}$ on the circuit is

$$\begin{aligned}M^{(-j)}\left[\rho_{[n]\setminus\{j\}} \otimes |0\rangle\langle 0|^{\otimes 2}_{\{j\}}\right] &= \rho_{[n]\setminus\{j\}} \otimes M\left[|0\rangle\langle 0|^{\otimes 2}_{\{j\}}\right] \\ &= \rho_{[n]\setminus\{j\}} \otimes \left(\frac{1-q^{-1}}{q^2-1} I + \frac{1-q^{-1}}{q^2-1} S\right) \\ &= \rho_{[n]\setminus\{j\}} \otimes \frac{1}{q(q+1)}(I+S)_{\{j\}}\end{aligned}$$

To express the evolution of the circuit under a series of actions more closely, we introduce the term **circuit configuration**.

Definition 24 (Circuit configuration)**.** *A circuit configuration at time t is a length-n vector that records the value of each qudit at the time, denoted $\vec{\gamma}^{(t)} \in \{I, S\}^n$. And γ_j denotes the jth qudit of the configuration.*

Hence for the first n single-qudit unitaries, $t \in \{-1, ..., -n\}$

$$M^{(-1)} \circ ... \circ M^{(-n)} \left[|0^n\rangle \langle 0^n|^{\otimes 2} \right] = \bigotimes_{j=1}^{n} \left(\frac{1}{q(q+1)} (I+S)_{\{j\}} \right) = \frac{1}{q^n (q+1)^n} \sum_{\vec{\gamma} \in \{I,S\}^n} \bigotimes_{j=1}^{n} \gamma_j \tag{3.14}$$

where each $\vec{\gamma} \in \{I, S\}^n$ is a configuration.

Remark 9. *The expected value of two copies of the state after application of all the single-qudit unitaries is a uniform sum over all configurations consisting of identity and swap at the n sites. Denote the configurations at this time by* $\vec{\gamma^{(0)}}$.

When the operator γ_j is the identity operator I, it leaves the state unchanged. When the operator γ_j is the swap operation S, applied to a system in the state $|0^n\rangle$, it swaps the states of two subsystems. However, since both subsystems are in the same state $|0\rangle$ the swap has no effect, leaving the state unchanged.
Thus,

$$\mathrm{Tr} \left(\bigotimes_{j=1}^{n} \gamma_j \right) |0^n\rangle\langle 0^n|^{\otimes 2} = 1$$

for all $\tilde{\gamma} \in \{I, S\}^n$.

For the following two-qubit gates at time $t \in \{1, ..., s\}$, we have $q^2 \times q^2$ 2-qudit unitaries acting on the qudit pair $A^{(t)}$. The explicit formula of $M^{(t)}$ can be obtained by rewriting Equation 3.13: substituting $q \to q^2$, replacing $I \to I \otimes I$, the identity operation on two copies of two qudits, and similarly, $S \to S \otimes S$. We thus obtain

$$\begin{aligned} M^{(t)}[\rho_{A^{(t)}}] = &\frac{\mathrm{Tr}(\rho_{A^{(t)}}) - q^{-2}\,\mathrm{Tr}(\rho_{A^{(t)}}(S \otimes S))}{q^4 - 1} (I \otimes I)_{A^{(t)}} \\ &+ \frac{\mathrm{Tr}(\rho_{A^{(t)}}(S \otimes S)) - q^{-2}\,\mathrm{Tr}(\rho_{A^{(t)}})}{q^4 - 1} (S \otimes S)_{A^{(t)}} \end{aligned}$$

As we've derived in Equation 3.14, after the single-qudit gates, the input state to $M^{(t)}$ will be a sum of tensor products of operators I and S. Thus, the possible values of the pair $A^{(t)}$ are $I \otimes I, I \otimes S, S \otimes I$, and $S \otimes S$. We only need to evaluate the performance of $M^{(t)}$ on these inputs:

$$M^{(t)}[I \otimes I] = \frac{q^4 - q^{-2} \times q^2}{q^4 - 1} (I \otimes I) + \frac{q^2 - q^{-2} \times q^4}{q^4 - 1} (S \otimes S) = I \otimes I \tag{3.15}$$

Similarly we have,

$$M^{(t)}[S \otimes S] = S \otimes S \tag{3.16}$$

And for inputs with different values,

$$M^{(t)}[I \otimes S] = M^{(t)}[S \otimes I] = \frac{q}{q^2 + 1} (I \otimes I) + \frac{q}{q^2 + 1} (S \otimes S) \tag{3.17}$$

Remark 10. *The three equations above define how the circuit will evolve under the two-qudit unitaries. If the two input values agree, the output should be just as the input. Otherwise, if the two input values disagree, either one of them must be flipped to make the two output values agree, while the coefficient is reduced by a factor of* q/q^2+1.

To generalize, denote input and output configurations of $M^{(t)}$ as $\vec{\gamma}, \vec{\nu}$. Let $M^{(t)}_{\vec{\nu}\vec{\gamma}}$ be the coefficient such that

$$M^{(t)}\left[\bigotimes_{j=1}^{n}\gamma_j\right] = \sum_{\vec{\nu}\in\{I,S\}^n} M^{(t)}_{\vec{\nu}\vec{\gamma}} \bigotimes_{j=1}^{n}\nu_j \tag{3.18}$$

Suppose $U^{(t)}$ acts on qudits $A^{(t)} = \{a, b\} \subset [n]$. Then, we have

$$M^{(t)}_{\vec{\nu}\vec{\gamma}} = \begin{cases} 1 & \text{if } \gamma_a = \gamma_b \text{ and } \vec{\gamma} = \vec{\nu} \\ \frac{q}{q^2+1} & \text{if } \gamma_a \neq \gamma_b \text{ and } \nu_a = \nu_b \text{ and } \gamma_c = \nu_c \, \forall c \in [n] \setminus \{a, b\} \\ 0 & \text{otherwise} \end{cases}$$

Knowing the performance of $M^{(t)}$ ($t \in \{1, ..., s\}$) on all possible input configurations $\gamma^{(0)}$, we can determine the evolution of the entire circuit, usually referred to as **circuit trajectory**.

Definition 25 (circuit trajectory)**.** *The circuit trajectory is a length-*$s+1$ *sequence of configurations, denoted* $\gamma = (\vec{\gamma}^{(0)}, ..., \vec{\gamma}^{(s)})$.

Finally, we rewrite Equation 3.7 as follows

$$\begin{aligned} Z &= q^n \operatorname{Tr}\left[[M^{(s)} \circ M^{(s-1)} \circ ... \circ M^{(1)} \circ M^{(-1)} \circ ... \circ M^{(-n)}] \left|0^n\right\rangle \left\langle 0^n\right|^{\otimes 2}\right] \\ &= \frac{1}{(q+1)^n} \sum_{\gamma\in\{I,S\}^{n\times(s+1)}} \prod_{t=1}^{s} M^{(t)}_{\vec{\gamma}^{(t)}\vec{\gamma}^{(t-1)}} \\ &=: \frac{1}{(q+1)^n} \sum_{\gamma} \text{weight}(\gamma) \end{aligned}$$

where the weight of a trajectory γ is the product of $M^{(t)}_{\vec{\gamma}^{(t)}\vec{\gamma}^{(t-1)}}$ for each step in the trajectory. We have now arrived at the concise formula for Z. In the next subsection, we will re-express Z in terms of random walks through the configuration space $\{I, S\}^n$.

3.2.3 Unbiased Walk

Consider the s-gates circuit in terms of a length-s unbiased random walk through configuration space $\{I, S\}^n$. Denote the configuration space with P_u. At time $t = 0$, a initial configuration $\vec{\gamma}^{(0)}$ is chosen uniformly at random, where we denote this uniform distribution of initial configuration space by Λ_u.

The configuration $\vec{\gamma}^{(t+1)}$ is derived from the configuration $\vec{\gamma}^{(t)}$ as follows: Suppose $A^{(t)} = \{a, b\}$, that is, the gate acts on qudits a and b at time t. If the ath and bth bits of $\vec{\gamma}^{(t)}$ agree, then the configuration is left unchanged at time $t+1$. Otherwise, if they disagree, either the value at a or b is flipped, each with probability $\frac{1}{2}$ to form $\vec{\gamma}^{(t+1)}$. Also the weight is reduced by $\frac{q}{q^2+1}$ each time a bit is flipped.

Note that the given P_u, Λ_u, the probability of a certain initial configuration occurring is 2^{-n}, and the probability of certain trajectory occurring is $2^{-n}(1/2)^{\text{number of bit flips}}$. Each trajectory contributes uniformly to Z, thus, we have

$$\begin{aligned}
Z &= \frac{1}{(q+1)^n} \sum_{\gamma} \left(\frac{q}{q^2+1} \right)^{\text{number of bit flips in } \gamma} \\
&= \frac{1}{(q+1)^n} \mathbb{E}_{P_u, \Lambda_u} \left[\frac{2^n}{\frac{1}{2}^{\text{number of bit flips}}} \left(\frac{q}{q^2+1} \right)^{\text{number of bit flips in } \gamma} \right] \\
&= \frac{1}{(q+1)^n} \mathbb{E}_{P_u, \Lambda_u} \left[2^n \left(\frac{2q}{q^2+1} \right)^{\text{number of bit flips in } \gamma} \right] \\
&= \frac{2^n}{(q+1)^n} \mathbb{E}_{P_u, \Lambda_u} \left[\left(\frac{2q}{q^2+1} \right)^{(\text{number of bit flips})} \right]
\end{aligned}$$

3.2.4 Biased Walk

We now introduce a biased random walk, denoted P_b. The initial distribution Λ_b is no longer uniform over $\{I, S\}^n$. Let the probability of choosing $\vec{\gamma}^{(0)} = \vec{\nu}$ is proportional to $q^{|-\vec{\nu}|}$. That is, the more number of S entries, the less likely. Specifically,

$$\Lambda_b(\vec{\nu}) = \frac{q^n}{(q+1)^n} q^{-|\vec{\nu}|}$$

Note that the factor $\frac{q^n}{(q+1)^n}$ normalizes the probability so that

$$\sum_{x=0}^{n} \frac{q^n}{(q+1)^n} q^{-x} \binom{n}{x} = 1$$

where x is the possible Hamming weight of a configuration.

P_b follows the same rules as P_u except that when two bits involved in a gate disagree, it chooses to flip the S to I with higher probability $q^2/(q^2+1)$ and flip the I to S with lower probability $1/(q^2+1)$. Overall, P_b is biased toward the I direction.

Proposition 8. *In biased walk, the weight of a particular walk is no longer dependent on the number of bit flips occur during that walk, but only on the Hamming weight of its final endpoint:*

$$Z = \frac{1}{q^n} \mathbb{E}_{P_b, \Lambda_b} \left[q^{|\vec{\gamma}^{(s)}|} \right]$$

Proof. We can show this by rewriting the formulation of Z in terms of the biased walk.
For a trajectory γ, denote the Hamming weight of the initial configuration $\left|\vec{\gamma}^{(0)}\right|$ by w_0, and denote that of the final configuration $\left|\vec{\gamma}^{(s)}\right|$ by w. Suppose among the n number of flips in γ, there are n_{S} flips that changes I to S, and n_{I} flips that changes S to I. Then,

$$n = n_{\mathrm{S}} + n_{\mathrm{I}}$$

$$w = w_0 + n_{\mathrm{S}} - n_{\mathrm{I}}$$

And the probability of a certain trajectory occurring is

$$\mathrm{Prob}(\gamma) = \frac{q^n}{(q+1)^n} q^{-w_0} \left(\frac{1}{q^2+1}\right)^{n_{\mathrm{S}}} \left(\frac{q^2}{q^2+1}\right)^{n_{\mathrm{I}}}$$

$$\begin{aligned}
Z &= \frac{1}{(q+1)^n} \sum_{\gamma} \left(\frac{q}{q^2+1}\right)^{\text{number of bit flips in } \gamma} \\
&= \frac{1}{(q+1)^n} \left(\frac{(q+1)^n}{q^n} q^{w_0} (q^2+1)^{n_{\mathrm{S}}} (\frac{q^2+1}{q^2})^{n_{\mathrm{I}}}\right) \mathbb{E}_{P_{b,\Lambda_b}} \left[(\frac{q}{q^2+1})^{n_{\mathrm{S}}+n_{\mathrm{I}}}\right] \\
&= \frac{1}{q^n} q^{w-n_{\mathrm{S}}+n_{\mathrm{I}}} q^{-2n_{\mathrm{I}}} \mathbb{E}_{P_{b,\Lambda_b}} \left[q^{n_{\mathrm{S}}+n_{\mathrm{I}}}\right] \\
&= \frac{1}{q^n} \mathbb{E}_{P_{b,\Lambda_b}} \left[q^w\right]
\end{aligned}$$

□

3.2.5 Evaluating circuit at infinite size

Now we have a good sense of the trajectory evolution, we can analyze our circuit in the infinite size limit. By Remark 10, each time a two-qudit gate acts on a disagreeing pair of input values, it will flip one of them to have a agreeing pair of output. Thus, each trajectory is forced to keep flipping values until all values agree, that is, it reaches a fixed point I^n or S^n.

The following proposition shows the probabilities of a trajectory reaching either I^n or S^n based on the hamming weight of its initial configuration. A detailed proof using recursion relation can be found in the Appendix of [9]. We will only be using it in this paper.

Proposition 9. *If a trajectory γ with input configuration $\vec{\gamma}^{(0)}$ has $\left|\vec{\gamma}^{(0)}\right| = x$, and we allow the biased walk to evolve until it reaches one of the fixed points I^n or S^n, then the probability that the trajectory ends at I^n is*

$$P_I(x) = \frac{1}{1-q^{-2n}}(1-q^{-2n+2x}) \tag{3.19}$$

and the probability that it ends at S^n is

$$P_S(x) = \frac{q^{-2n+2x}}{1-q^{-2n}}(1-q^{-2x}) \tag{3.20}$$

Given the two expression above, the expected trajectory weight of biased walk in the infinite circuit size is

$$\mathbb{E}_{P_b,\vec{\gamma}^{(0)}}\left[q^{|\vec{\gamma}^{(s)}|}\right] = P_I(|\vec{\gamma}^{(0)}|) + q^n P_S(|\vec{\gamma}^{(0)}| \tag{3.21}$$

The collision probability of at infinite size limit, denoted Z_H, can be calculated as follows

$$\begin{aligned}
Z_H &= \frac{1}{q^n}\mathbb{E}_{P_b,\Lambda_b}\left[q^{|\vec{\gamma}^{(s)}|}\right] \\
&= \frac{1}{q^n}\sum_{\vec{\gamma}^{(0)}}\Lambda_b(\vec{\gamma}^{(0)})\mathbb{E}_{P_b,\vec{\gamma}^{(0)}}\left[q^{|\vec{\gamma}^{(s)}|}\right] \\
&= \frac{1}{q^n}\left(\frac{q^n}{(q+1)^n}q^{-|\vec{\gamma}^{(0)}|}\right)(P_I(|\vec{\gamma}^{(0)}|) + q^n P_S(|\vec{\gamma}^{(0)}|) \\
&= \frac{1}{(q+1)^n(1-q^{-2n})}\sum_{x=0}^{n}\binom{n}{x}q^{-x}(1-q^{-2n+2x}+q^{-n+2x}-q^{-n}) \\
&= \frac{2}{q^n+1}
\end{aligned}$$

Remark 11. *When our circuit size s is finite, Z is much larger than Z_H because many trajectories have not yet reached a fixed point, and thus contribute larger weights to Z than they have in infinite circuit. To upper bound Z, we want to understand how quickly these trajectories approach the fixed points, and consequently Z approaches Z_H.*

3.3 Upper Bound for 1D architecture

In this section, we show an upper bound on the collision probability in $1D$ architecture. For $1D$ architecture, we assume periodic boundary conditions and have qudits arranged locally in the circuit. Thus, we can think of a configuration as a composition of domains containing consecutive sites that share the same values. For gate that acts on two qudits in different domains, i.e. have disagreeing values, it flips one of the value and causes the domain boundary to move. We will implement this concept in the next subsection, and then proof the upper bound for Z.

3.3.1 Domain Wall notation

Definition 26 (Domain Wall)**.** *Let $DW(\vec{\gamma}) := \{e \in \{1,2,...,n-1\} : \gamma_e \neq \gamma_{e+1}\}$ be the set of* domain wall positions *for a configuration $\vec{\gamma}$, where γ_0 is identified with γ_n for periodic boundary conditions.*

Remark 12. *For each set of domain wall locations, there are two configuration $\vec{\gamma} \in \{I, S\}^n$ that map to it, since we can either choose $\gamma_0 = I$ or $\gamma_0 = S$ and the values of all other sites follow.*

Then, define $G = (g^{(0)}, ..., g^{(s)})$ to be the *domain wall trajectory*, i.e. a sequence of sets of domain wall locations associated with a configuration trajectory $\gamma = (\vec{\gamma}^{(0)}, ..., \vec{\gamma}^{(s)})$, where $g^{(t)} = DW(\vec{\gamma}^{(t)})$.

Recall that the circuit trajectory follows rules discussed in Remark 10. We translate these rules for the domain wall trajectory.

Definition 27 (non-trivial DW)**.** *A non-zero contribution (non-trivial) domain wall trajectory G must obey the following rules:*

1. *when there is a domain wall at position e and a gate acts on qudits $\{e, e+1\}$, the domain wall must move to position $e-1$ or $e+1$ and may annihilate with another domain wall if there is already a domain wall at the new position.*
2. *Each time there is such a movement, the weight of G is reduced by the constant factor $\frac{q}{q^2+1}$.*

Let $\mathcal{G}$ be the set of all domain wall trajectories that are non-trivial. Knowing each G corresponds to two circuit trajectories (Remark 11), we can now express Z by

$$Z = \frac{2}{(q+1)^n} \sum_{G \in \mathcal{G}} \text{weight}(G) \tag{3.22}$$

where the weight of G is given by

$$\text{weight}(G) := \Pi_{t=1}^{s} M^{(t)}_{g^{(t-1)} g^{(t)}} \tag{3.23}$$

with

$$M^{(t)}_{g^{(t-1)} g^{(t)}} := \begin{cases} \frac{q}{q^2+1} & \text{if } \min(A^{(t)}) \in g^{(t-1)} \\ 1 & \text{otherwise} \end{cases} \tag{3.24}$$

3.3.2 Domain Wall Composition

Say we have two domain wall trajectories G, G', consider the combined domain wall trajectory simply as

$$G \sqcup G' := (g^{(0)} \sqcup g'^{(0)}, ..., g^{(s)} \sqcup g'^{(s)})$$

where $\sqcup$ is the disjoint union and is defined only under the assumption $g^{(t)} \cap g'^{(t)} = \emptyset \forall t$.

Remark 13. *if $H = G \sqcup G'$, then weight(H) = weight$(G) \cdot$ weight(G').*

Note that for each domain wall position, it either goes up, down, or annihilates. Thus, we can decompose a domain wall trajectory G particularly into $G = G_U \sqcup G_0$ where G_U is the trajectory for the set of positions that do not annihilate, while G_0 is the trajectory for the set of positions for which $|G_0^{(s)}| = 0$, i.e. all the domain walls have annihilated at the end of the trajectory.

3.3.3 Calculation of the upper bound

We associate each domain wall trajectory $G = (g^{(0)}, ..., g^{(s)})$ with an integer $k = |g^{(s)}|$, the number of domain walls that remain unannihilated at the end of the trajectory.

Let $\mathcal{G}_k \subset \mathcal{G}$ be the associated set of length-s domain wall trajectories that ends with k domain walls. Let $\mathcal{G}_{U,k} \subset \mathcal{G}_k$ be the subset of domain wall trajectories that start with these k unannihilated domain walls. As shown above, there exists an unique decomposition $H \in \mathcal{G}_k$ such that $H = G \sqcup G'$, where $G \in \mathcal{G}_{U,k}$ and $G' \in \mathcal{G}_0$.

We proceed to look at the total weight for each set separately. We first try to upper bound the total weight for $\mathcal{G}_{U,k}$. Given a fixed k,there are $\binom{n}{k}$ number of $G \in \mathcal{G}_{U,k}$ that start and end with k domain walls, i.e. $|g^{(0)}| = k, |g^{(s)}| = k$.

During each layer of gates, each domain wall must move either left or right, except for the first layer if the domain walls all begin at an even position. Thus, the total weight of all trajectories in $\mathcal{G}_{U,k}$ is upper bounded by

$$\sum_{G \in \mathcal{G}_{U,k}} \text{weight}(G) \leq \binom{n}{k} \left(\frac{2q}{(q^2+1)} \right)^{k(d-1)} \tag{3.25}$$

where $\left(\frac{2q}{(q^2+1)}\right)^{k(d-1)}$ is the weight reduction the k gates experienced in the $d-1$ layers of the circuit.

To upper bound the total weight of $\mathcal{G}_0$, we follow the analysis in Section 3.2.5, which shows that the sum over all trajectories that reach a fixed point in infinite circuit size is $Z_H(q+1)^n$. Then, the sum of the weights of all domain wall trajectories in $\mathcal{G}_0$ approaches $Z_H(q+1)^n/2$ from below as depth increases.

$$\sum_{G \in \mathcal{G}_0} \text{weight}(G) \leq Z_H(q+1)^n/2 \tag{3.26}$$

With each decomposition bounded, now we can calculate the upper bound for the whole weight. To sum over all non-trivial domain wall trajectories $G \in \mathcal{G}$ is equivalent to sum over all possible number of domain wall k at the end of the trajectory, and then sum over all $G \in \mathcal{G}_k$.

Since at each layer in 1D architecture, there are $n/2$ 2-qudit gates acts consecutively on all n qudits, and each gate would either result in both I's or S's, thus the number of domain wall is at most $n/2$. Also note that we have periodic boundary conditions, i.e. $\gamma_0 = \gamma_n$, thus the number of domain wall positions must always be even. Denote $k_0 = k/2$, then

$$Z = \frac{2}{(q+1)^n} \sum_{G \in \mathcal{G}} \text{weight}(G) = \frac{2}{(q+1)^n} \sum_{k_0=0}^{n/4} \sum_{G \in \mathcal{G}_{2k_0}} \text{weight}(G) \tag{3.27}$$

Theorem 5 (Upper bound in 1D)**.** *For the 1D architecture, let*

$$a := \log\left(\frac{q^2+1}{2q}\right)$$

$$s^* := \frac{1}{2a} n \log(n) + O(n)$$

Then,

$$Z \leq Z_H(1 + e^{-\frac{2a}{n}(s-s^*)}) \tag{3.28}$$

whenever $s \geq s^*$.

Proof. Knowing $\text{weight}(H) = \text{weight}(G) \cdot \text{weight}(G')$, we have

$$\begin{aligned} Z &= \frac{2}{(q+1)^n} \sum_{k_0=0}^{n/4} \sum_{G \in \mathcal{G}_{2k_0}} \text{weight}(G) \\ &= \frac{2}{(q+1)^n} \sum_{k_0=0}^{n/4} \sum_{G \in \mathcal{G}_{U,2k_0}} \sum_{\substack{G' \in \mathcal{G}_0 \\ G \cap G' = \emptyset}} \text{weight}(G) \cdot \text{weight}(G') \end{aligned}$$

Apply the two upper bounds we derived,

$$\begin{aligned} Z &\leq \left(\sum_{k_0=0}^{n/4} \sum_{G \in \mathcal{G}_{U,2k_0}} \text{weight}(G) \right) \left(\frac{2}{(q+1)^n} \sum_{G' \in \mathcal{G}_0} \text{weight}(G') \right) \\ &\leq \left(\sum_{k_0=0}^{n/4} \binom{n}{2k_0} \left(\frac{2q}{q^2+1} \right)^{2k_0(d-1)} \right) (Z_H) \end{aligned}$$

Use variable a to represent the constant factor $a := \log\left(\frac{q^2+1}{2q}\right)$,

$$\begin{aligned} &= Z_H \sum_{k_0=0}^{n/4} \binom{n}{2k_0} (e^{-a})^{2k_0(d-1)} \\ &\leq Z_H(1 + e^{-a(d-1)})^n \\ &\leq Z_H(1 + (e-1)ne^{-a(d-1)}) \\ &= Z_H(1 + \exp(\log(n) - da + \log(e-1) + a)) \\ &\leq Z_H(1 + \exp(-a(d-d^*))) \end{aligned}$$

Set $d^* := \frac{1}{a}\log(n) + O(1)$, then for all $d \geq d^*$, the value in the parenthesis remains approximately a constant. □

Thus, the collision probability Z is upper bounded by a constant multiple of Z_H when the circuit depth d is in log-depth.

Conclusion

4.1 Summary and Future Prospects

In this thesis, we focus on the anti-concentration assumption within the framework of random circuit sampling. It is important to understand the conditions for this assumption, as it is a necessary requirement to determine whether a quantum circuit is difficult to simulate classically.

We start with defining anti-concentration by the expected collision probability, which captures the concept that the output probability mass gradually spreads out over all possible measurement outcomes as the circuit depth increases.

We then bound the collision probability with stochastic analysis in $1D$ architecture. We show that on an n qudit circuit where all random unitaries are each drawn from the Haar measure, at least $n\log(n)$ gates, and thus $\log(n)$ circuit depth, are needed for this condition to be met.

Having seen how anti-concentration is achieved in $1D$ architecture, a natural extension is to study this property under the $2D$ **architecture**. What circuit depth would be sufficient for Equation 3.28 to hold.

Definition 28 (2D architecture)**.** *Assume n is even. The circuit diagram on $n \times n$ qudits is generated by alternating between the four types of layers of $n^2/2$ non-overlapping nearest-neighbor two-qudit gates on a array.*

Note that the $2D$-grid architecture is a natural implementation of a circuit when we have to physically lay out the qubits. Current quantum processor like Sycamore uses this layout.

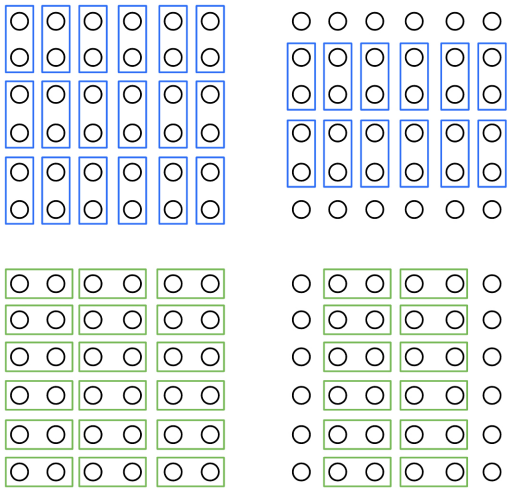


Figure 4.2: The four layers of 2-qudit gates applied to the circuits consecutively. Each rectangle represents a 2-qubit gate.

Appendix A

Relevant Proofs

A.1 Proof that if $b, c > 0$ and $cb \leq 1$, then

$$(1+c)^b \leq 1 + cb(e-1)$$

Proof:

$$\begin{aligned}
(1+c)^b &= \sum_{k=0}^{b} \binom{b}{k} c^k \\
&= 1 + cb \sum_{k=1}^{b} \binom{b}{k} \frac{c^{k-1}}{b} \\
&\leq 1 + cb \sum_{k=1}^{b} \binom{b}{k} b^{-k} \\
&\leq 1 + cb((1+b^{-1})^b - 1) \\
&\leq 1 + cb(e-1)
\end{aligned}$$

References

[1] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[2] Lov K Grover. Quantum computers can search rapidly by using almost any transformation. *Physical Review Letters*, 80(19):4329, 1998.

[3] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[4] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[5] Paul R Halmos. *Measure theory*, volume 18. Springer, 2013.

[6] Sean Mullane. Sampling random quantum circuits: a pedestrian's guide. *arXiv preprint arXiv:2007.07872*, 2020.

[7] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, 2019.

[8] 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.

[9] Alexander M Dalzell, Nicholas Hunter-Jones, and Fernando GSL Brandao. Random quantum circuits anticoncentrate in log depth. *PRX Quantum*, 3(1):010333, 2022.

[10] Roe Goodman and Nolan R Wallach. *Representations and invariants of the classical groups*. Cambridge University Press, 2000.