Classically Simulating Quantum Circuits with Local Depolarizing Noise

Yasuhiro Takahashi
NTT Communication Science Laboratories, NTT Corporation, Atsugi, Japan
yasuhiro.takahashi.rb@hco.ntt.co.jp

Yuki Takeuchi
NTT Communication Science Laboratories, NTT Corporation, Atsugi, Japan
yuki.takeuchi.yt@hco.ntt.co.jp

Seiichiro Tani
NTT Communication Science Laboratories, NTT Corporation, Atsugi, Japan
seiichiro.tani.cs@hco.ntt.co.jp

Abstract

We study the effect of noise on the classical simulatability of quantum circuits defined by computationally tractable (CT) states and efficiently computable sparse (ECS) operations. Examples of such circuits, which we call CT-ECS circuits, are IQP, Clifford Magic, and conjugated Clifford circuits. This means that there exist various CT-ECS circuits such that their output probability distributions are anti-concentrated and not classically simulatable in the noise-free setting (under plausible assumptions). First, we consider a noise model where a depolarizing channel with an arbitrarily small constant rate is applied to each qubit at the end of computation. We show that, under this noise model, if an approximate value of the noise rate is known, any CT-ECS circuit with an anti-concentrated output probability distribution is classically simulatable. This indicates that the presence of small noise drastically affects the classical simulatability of CT-ECS circuits. Then, we consider an extension of the noise model where the noise rate can vary with each qubit, and provide a similar sufficient condition for classically simulating CT-ECS circuits with anti-concentrated output probability distributions.

2012 ACM Subject Classification Theory of computation → Quantum computation theory

Keywords and phrases Classical Simulation, Quantum Circuit, Local Depolarizing Noise

Digital Object Identifier 10.4230/LIPIcs.MFCS.2020.83

Related Version A full version of the paper is available at `https://arxiv.org/abs/2001.08373`.

1 Introduction

1.1 Background and Main Results

A key step toward realizing a large-scale universal quantum computer is to demonstrate quantum computational supremacy [11], i.e., to perform computational tasks that are classically hard. As such a task, many researchers have focused on simulating quantum circuits, or more concretely, sampling the output probability distributions of quantum circuits. They have shown that, under plausible complexity-theoretic assumptions, this task is classically hard for various quantum circuits that seem easier to implement than universal ones. However, these classical hardness results have been obtained in severely restricted settings, such as a noise-free setting with additive approximation [5, 17, 3, 20] and a noise setting with multiplicative approximation [9]: the former requires us to sample the output probability distribution of a quantum circuit with additive error and the latter to sample the output probability distribution of a quantum circuit under a noise model with multiplicative error. Thus, there is great interest in considering the above task in a more reasonable setting.

We study the classical simulatability of quantum circuits in a noise setting with additive approximation, which requires us to sample the output probability distribution of a quantum circuit under a noise model with additive error. This setting is more reasonable than the noise-free setting since the presence of noise is unavoidable in realistic situations. Moreover, our setting is more reasonable than the noise setting with multiplicative approximation in the sense that we adopt a more realistic notion of approximation [1, 5], although noise in this paper is more restrictive than that in [9]. We consider a noise model where a depolarizing channel with an arbitrarily small constant rate $0 < \varepsilon < 1$, which is denoted as D_ε, is applied to each qubit at the end of computation. This channel leaves a qubit unaffected with probability $1 - \varepsilon$ and replaces its state with the completely mixed one with probability ε. We call this model noise model **A**. We also consider its extension where the noise rate can vary with each qubit. More concretely, when a quantum circuit has n qubits, D_{ε_j} is applied to the j-th qubit at the end of computation for any $1 \leq j \leq n$. We call this model noise model **B**. These noise models are simple, but analyzing them is a meaningful step toward studying more general models [10], such as one where noise exists before and after each gate in a quantum circuit. This is because, for example, this general noise model is equivalent to noise model **A** when we focus on instantaneous quantum polynomial-time (IQP) circuits, which are described below, with a particular type of intermediate noise [6].

A representative example of a quantum circuit that is not classically simulatable (in the noise-free setting) is an IQP circuit, which consists of Z-diagonal gates sandwiched by two Hadamard layers. In fact, there exists an IQP circuit such that its output probability distribution is anti-concentrated and not classically samplable in polynomial time with certain constant accuracy in l_1 norm (under plausible assumptions) [5]. On the other hand, Bremner et al. [6] assumed noise model **A** and showed that, if the *exact* value of the noise rate is known, any IQP circuit with an anti-concentrated output probability distribution is classically simulatable in the sense that the resulting probability distribution is classically samplable in polynomial time with arbitrary constant accuracy in l_1 norm. This indicates that, under noise model **A**, if the exact value of the noise rate is known, the presence of small noise drastically affects the classical simulatability of IQP circuits.

In this paper, first, under a weaker assumption on the knowledge of the noise rate, we extend Bremner et al.'s result to a new class of quantum circuits defined by two concepts: computationally tractable (CT) states and efficiently computable sparse (ECS) operations [19]. Examples of such circuits, which we call CT-ECS circuits, are IQP circuits, Clifford Magic circuits [20], and conjugated Clifford circuits [3]. This means that there exist various CT-ECS circuits such that their output probability distributions are anti-concentrated and not classically simulatable in the sense described above for IQP circuits. Constant-depth quantum circuits [18, 4, 2] are also CT-ECS circuits and not classically simulatable, although we do not know whether their output probability distributions are anti-concentrated. We postpone the explanation of CT states and ECS operations until Section 2, but, as depicted in Fig. 1(a), a CT-ECS circuit on n qubits is a polynomial-size quantum circuit $C = VU$ such that $U|0^n\rangle$ is CT and $V^\dagger Z_j V$ is ECS for any $1 \leq j \leq n$, where Z_j is a Pauli-Z operation on the j-th qubit. After performing C, we perform Z-basis measurements on all qubits. The CT-ECS circuit C under noise model **B** is depicted in Fig. 1(b).

Our first result assumes noise model **A**, which is the case where $\varepsilon_j = \varepsilon$ for any $1 \leq j \leq n$ in Fig. 1(b). We show that, if an *approximate* value of the noise rate is known, any CT-ECS circuit with an anti-concentrated output probability distribution is classically simulatable:

▶ **Theorem 1** (informal). *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p is anti-concentrated, i.e., $\sum_{x\in\{0,1\}^n} p(x)^2 \leq \alpha/2^n$ for some known constant $\alpha \geq 1$. We assume that a depolarizing channel with (possibly unknown)*

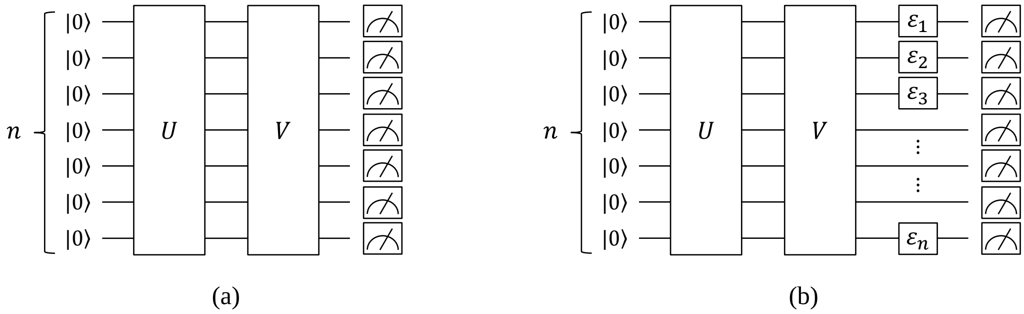

Figure 1 (a): CT-ECS circuit $C = VU$, where $U|0^n\rangle$ is CT and $V^\dagger Z_j V$ is ECS for any $1 \le j \le n$. After performing C, we perform Z-basis measurements on all qubits. (b): CT-ECS circuit $C = VU$ under noise model **B**, where ε_j represents the depolarizing channel D_{ε_j} for any $1 \le j \le n$.

constant rate $0 < \varepsilon < 1$ is applied to each qubit after performing C, which yields the probability distribution $\widetilde{p}_{\mathrm{A}}$. Moreover, we assume that it is possible to choose a constant λ such that

$$1 \le \frac{\varepsilon}{\lambda} \le 1 + c,$$

where c is a certain constant depending on α. Then, $\widetilde{p}_{\mathrm{A}}$ is classically samplable in polynomial time with constant accuracy in l_1 norm.

If ε is known, we can choose $\lambda = \varepsilon$. This case with IQP circuits precisely corresponds to Bremner et al.'s result [6]. As described above, there exist various CT-ECS circuits such that their output probability distributions are anti-concentrated and not classically simulatable in the noise-free setting (under plausible assumptions). Thus, Theorem 1 indicates that, under noise model **A**, if an *approximate* value of the noise rate is known, the presence of small noise drastically affects the classical simulatability of CT-ECS circuits.

Theorem 1 assumes noise model **A** where noise exists only at the end of computation, but, in some cases, it also holds under an input-noise model. For example, Theorem 1 holds for IQP circuits when noise model **A** is replaced with a noise model where D_ε is applied to each qubit only at the *start* of computation, although Bu and Koh's main result implies a similar property of IQP circuits [7]. Moreover, Theorem 2, which is described below and assumes noise model **B**, also holds for IQP circuits when noise model **B** is replaced with an input-noise model where the noise rate can vary with each qubit. Our main result is similar in spirit to Bu and Koh's, which provides classical algorithms for simulating Clifford circuits with nonstabilizer product input states (corresponding to input-noise models). However, we note that, in general, it is difficult to relate the output probability distributions of CT-ECS circuits under output-noise models to those of Clifford circuits under input-noise models.

Our main focus is on a noise setting, but, from a purely theoretical point of view, it is valuable to analyze the classical simulatability of quantum circuits in the noise-free setting. The proof method of Theorem 1 is based on computing the Fourier coefficients of an output probability distribution, and is useful in the noise-free setting. In fact, the proof method shows that, when only $O(\log n)$ qubits are measured, any quantum circuit in a class of CT-ECS circuits on n qubits is classically simulatable. More precisely, its output probability distribution is classically samplable in polynomial time with polynomial accuracy in l_1 norm. This class of CT-ECS circuits is defined by a restricted version of ECS operations, and includes IQP, Clifford Magic, conjugated Clifford, and constant-depth quantum circuits. It is known that the above property or a similar one holds for these quantum circuits, but the proofs provided have depended on each circuit class [18, 4, 12, 3]. Our analysis unifies the previous ones and clarifies a class of quantum circuits for which the above property holds.

Our second result assumes noise model **B**, which is depicted in Fig. 1(b). For classically simulating CT-ECS circuits with anti-concentrated output probability distributions, we provide a sufficient condition, which is similar to Theorem 1:

▶ **Theorem 2** (informal). *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p is anti-concentrated, i.e., $\sum_{x\in\{0,1\}^n} p(x)^2 \leq \alpha/2^n$ for some known constant $\alpha \geq 1$. We assume that a depolarizing channel with (possibly unknown) constant rate $0 < \varepsilon_j < 1$ is applied to the j-th qubit after performing C for any $1 \leq j \leq n$, which yields the probability distribution $\widetilde{p}_{\mathrm{B}}$. Moreover, we assume that it is possible to choose a constant $\lambda_{\min}$ such that*

$$1 \leq \frac{\varepsilon_{\min}}{\lambda_{\min}} \leq 1 + c,$$

where $\varepsilon_{\min} = \min\{\varepsilon_j | 1 \leq j \leq n\}$ and c is a certain constant depending on α, and we assume that it is possible to choose a constant λ_j such that

$$0 \leq \varepsilon_j - \lambda_j \leq c\lambda_{\min}$$

for any $1 \leq j \leq n$ with $\varepsilon_j \neq \varepsilon_{\min}$. Then, $\widetilde{p}_{\mathrm{B}}$ is classically samplable in polynomial time with constant accuracy in l_1 norm.

To the best of our knowledge, this is the first analysis of the classical simulatability of quantum circuits under noise model **B**. Theorem 2 indicates that, under this noise model, if approximate values of the minimum noise rate and the other noise rates are known, the presence of small noise drastically affects the classical simulatability of CT-ECS circuits.

1.2 Overview of Techniques

To prove Theorem 1, we generalize Bremner et al.'s proof for IQP circuits [6]. The first key point is to provide a general method for approximating the Fourier coefficients of the output probability distribution p. The probability distribution $\widetilde{p}_{\mathrm{A}}$, which we want to approximate, can be simply represented by the noise rate ε and the Fourier coefficients $\widehat{p}(s)$ of p for all $s \in \{0,1\}^n$ [15]. We show that, for any CT-ECS circuit on n qubits, there exists a polynomial-time classical algorithm for approximating each of the low-degree Fourier coefficients, i.e., $\widehat{p}(s)$ for any $s = s_1 \cdots s_n \in \{0,1\}^n$ with $\sum_{j=1}^n s_j = O(1)$. Bremner et al. [6] showed that such an algorithm exists for IQP circuits through a direct calculation of the Fourier coefficients for them. In contrast, we first provide a general relation between a quantum circuit and the Fourier coefficients of its output probability distribution. We then approximate each of the low-degree Fourier coefficients by combining this general relation with Nest's classical algorithm for approximating the inner product described by a CT state and an ECS operation [19]. This general relation seems to be a new tool to investigate the output probability distribution of a quantum circuit and thus may be of independent interest.

The second key point is to approximate $\widetilde{p}_{\mathrm{A}}$ using an approximate value λ of ε. We define a *function* q that seems to be close to $\widetilde{p}_{\mathrm{A}}$ on the basis of its representation with ε and $\widehat{p}(s)$ for all $s \in \{0,1\}^n$. In contrast to Bremner et al.'s setting, we do not know ε. Thus, we define q using λ and an appropriate (polynomial) number of the low-degree Fourier coefficients. This number depends on λ, the constant α associated with the anti-concentration assumption, and the desired approximation accuracy. We then evaluate the approximation accuracy of q. Here, we need to care about the error caused by the difference between λ and ε. We upper-bound this error using the anti-concentration assumption.

To prove Theorem 2, we represent the effect of noise under noise model **B** as the effect of noise under noise model **A** with rate $\varepsilon_{\min}$ and the remaining effects. We do this by transforming the representation of the probability distribution $\widetilde{p}_{\mathrm{B}}$, which we want to approximate, with several basic properties of noise operators on real-valued functions over $\{0,1\}^n$ [15]. The obtained representation means that, to sample $\widetilde{p}_{\mathrm{B}}$, it suffices to sample the probability distribution $\widetilde{p}_{\mathrm{A}}$ (resulting from p) under noise model **A** with rate $\varepsilon_{\min}$ and then to classically simulate noise corresponding to the remaining effects. By Theorem 1, $\widetilde{p}_{\mathrm{A}}$ is classically simulatable. Moreover, we can simulate noise corresponding to the remaining effects using the approximate values of $\varepsilon_{\min}$ and the other noise rates.

2 Preliminaries

2.1 Quantum Circuits and Their Output Probability Distributions

Pauli matrices X, Y, Z, and I are

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The Hadamard operation H is $(X+Z)/\sqrt{2}$. For any real number θ, the rotation operations $R_x(\theta)$ and $R_z(\theta)$ are $\cos(\theta/2)I - i\sin(\theta/2)X$ and $\cos(\theta/2)I - i\sin(\theta/2)Z$, respectively, where $R_z(\pi/4)$ and $R_z(\pi/2)$ are denoted as T and S, respectively. It holds that the inverse of $R_x(\theta)$ is $R_x(-\theta)$ and the inverse of $R_z(\theta)$ is $R_z(-\theta)$. The controlled-Z operation CZ and the controlled-controlled-Z operation CCZ are $|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes Z$ and $|00\rangle\langle 00|\otimes I + |01\rangle\langle 01|\otimes I + |10\rangle\langle 10|\otimes I + |11\rangle\langle 11|\otimes Z$, respectively, where the states $|0\rangle$ and $|1\rangle$ are ± 1-eigenstates of Z, respectively. The operations H, S, and CZ are called Clifford operations.

A quantum circuit consists of elementary gates, each of which is in the gate set $\mathcal{G}$. Here, $\mathcal{G} = \{R_x(\theta), R_z(\theta), CZ \,|\, \theta = \pm 2\pi/2^t \text{ with integer } t \geq 1\}$. Each $R_x(\theta)$ has its inverse in $\mathcal{G}$ and so does $R_z(\theta)$. Moreover, each of the one-qubit operations X, Y, Z, I, and H can be decomposed into a constant number of $R_x(\theta)$'s and $R_z(\theta)$'s (up to an unimportant global phase). Similarly, CCZ can be decomposed into a constant number of $R_x(\theta)$'s, $R_z(\theta)$'s, and CZ's. Since the gate set $\{H, T, CZ\}$ is approximately universal for quantum computation [13], so is $\mathcal{G}$. The complexity measures of a quantum circuit are its size and depth. The size of a quantum circuit is the number of elementary gates in the circuit. To define the depth, we regard the circuit as a set of layers $1, \ldots, d$ consisting of elementary gates, where gates in the same layer act on pairwise disjoint sets of qubits and any gate in layer j is applied before any gate in layer $j+1$. The depth is defined as the smallest possible value of d [8].

We deal with a uniform family of polynomial-size quantum circuits $\{C_n\}_{n\geq 1}$. Each C_n has n input qubits initialized to $|0^n\rangle$. After performing C_n, we perform Z-basis measurements on all qubits. The output probability distribution p of C_n over $\{0,1\}^n$ is defined as $p(x) = |\langle x|C_n|0^n\rangle|^2$. A symbol denoting a quantum circuit also denotes its matrix representation in the Z basis. The uniformity means that the function $1^n \mapsto \overline{C_n}$ is computable by a polynomial-time classical Turing machine, where $\overline{C_n}$ is the classical description of C_n [14].

2.2 Fourier Expansions and Effects of Noise

Let $f : \{0,1\}^n \to \mathbb{R}$ be an arbitrary (real-valued) function. Then, f can be uniquely represented as an $\mathbb{R}$-linear combination of 2^n basis functions

$$f(x) = \sum_{s\in\{0,1\}^n} \widehat{f}(s)(-1)^{s\cdot x},$$

which is called the Fourier expansion of f [15]. Here, $\widehat{f}(s)$ is called the Fourier coefficient of f and the symbol "$\cdot$" represents the inner product of two n-bit strings, i.e., $s \cdot x = \sum_{j=1}^{n} s_j x_j$ for any $s = s_1 \cdots s_n$, $x = x_1 \cdots x_n \in \{0,1\}^n$. It holds that

$$\widehat{f}(s) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)(-1)^{s \cdot x}$$

for any $s \in \{0,1\}^n$. The l_k norm of f is defined as

$$||f||_k = \left(\sum_{x \in \{0,1\}^n} |f(x)|^k \right)^{1/k},$$

where $k = 1, 2$. It is known that $||f||_1^2 \le 2^n ||f||_2^2 = 2^{2n} ||\widehat{f}||_2^2$ [15].

Let C be an arbitrary quantum circuit on n qubits and p be its output probability distribution over $\{0,1\}^n$. We consider C under noise model **A** where a depolarizing channel D_ε with rate $0 < \varepsilon < 1$ is applied to each qubit after performing C. Here, $D_\varepsilon(\rho) = (1-\varepsilon)\rho + \varepsilon \frac{I}{2}$ for any density operator ρ of a qubit. We perform Z-basis measurements on all qubits and let $\widetilde{p}_{\mathrm{A}}$ be the resulting probability distribution over $\{0,1\}^n$. As shown in [6], we can sample $\widetilde{p}_{\mathrm{A}}$ by sampling an n-bit string according to p and then flipping each bit of the string with probability $\varepsilon/2$. This implies the following Fourier expansion of $\widetilde{p}_{\mathrm{A}}$ [15]:

$$\widetilde{p}_{\mathrm{A}}(x) = \sum_{s \in \{0,1\}^n} (1-\varepsilon)^{|s|} \widehat{p}(s)(-1)^{s \cdot x},$$

where $|s| = \sum_{j=1}^{n} s_j$ for any $s = s_1 \cdots s_n \in \{0,1\}^n$. We also consider C under noise model **B** where D_{ε_j} with rate $0 < \varepsilon_j < 1$ is applied to the j-th qubit after performing C for any $1 \le j \le n$. We perform Z-basis measurements on all qubits and let $\widetilde{p}_{\mathrm{B}}$ be the resulting probability distribution over $\{0,1\}^n$. As with $\widetilde{p}_{\mathrm{A}}$, we can sample $\widetilde{p}_{\mathrm{B}}$ by sampling an n-bit string according to p and then flipping its j-th bit with probability $\varepsilon_j/2$ for any $1 \le j \le n$. The Fourier expansion of $\widetilde{p}_{\mathrm{B}}$ is as follows [15]:

$$\widetilde{p}_{\mathrm{B}}(x) = \sum_{s \in \{0,1\}^n} \left[\prod_{j=1}^{n} (1-\varepsilon_j)^{s_j} \right] \widehat{p}(s)(-1)^{s \cdot x}.$$

2.3 CT States and ECS Operations

We introduce CT states and (a restricted version of) ECS operations [19]. Let $|\varphi\rangle$ be an arbitrary (pure) quantum state on n qubits and p be the probability distribution over $\{0,1\}^n$ defined as $p(x) = |\langle x|\varphi\rangle|^2$. Then, $|\varphi\rangle$ is CT if p is classically samplable in polynomial time and, for any $x \in \{0,1\}^n$, $\langle x|\varphi\rangle$ is classically computable in polynomial time.[1] An example of a CT state is a product state. Let U be an arbitrary quantum operation on n qubits that is both unitary and Hermitian. The operation U is sparse if there exists a polynomial s in n such that, for any $x \in \{0,1\}^n$, $U|x\rangle$ is a linear combination of at most $s(n)$ computational basis states. When U is a sparse operation (associated with s), for any $1 \le j \le s(n)$, we define two functions, $\beta_j : \{0,1\}^n \to \mathbb{C}$ and $\gamma_j : \{0,1\}^n \to \{0,1\}^n$, as follows: for any $x \in \{0,1\}^n$, if the

[1] For simplicity, we require perfect accuracy in sampling probability distributions and computing values, although irrational numbers may be involved. Precisely speaking, it suffices to require exponential accuracy. This is also applied to similar situations in this paper, such as the definition of ECS operations.

j-th non-zero entry exists in the column indexed by x when traversing this column from top to bottom, $\beta_j(x)$ is this entry and $\gamma_j(x)$ is the row index associated with $\beta_j(x)$. If the j-th non-zero entry does not exist in this column, $\beta_j(x) = 0$ and $\gamma_j(x) = 0^n$. The sparse operation U is ECS if, for any $x \in \{0,1\}^n$ and $1 \le j \le s(n)$, $\beta_j(x)$ and $\gamma_j(x)$ are classically computable in polynomial time. In particular, an ECS operation with $s(n) = O(1)$ is called ECS_1, and an ECS operation with $s(n) = 1$ is called efficiently computable basis-preserving. An efficiently computable basis-preserving operation preserves the class of CT states as follows:

▶ **Theorem 3** (Lemma 2 of [19]). *Let $|\varphi\rangle$ be an arbitrary CT state on n qubits and U be an arbitrary efficiently computable basis-preserving operation on n qubits. Then, $U|\varphi\rangle$ is CT.*

The following theorem is a rephrased version of the one in [19]:

▶ **Theorem 4** (Theorem 3 of [19]). *Let U be an arbitrary quantum operation on n qubits such that $U|0^n\rangle$ is CT, and O be an arbitrary observable with $||O|| \le 1$, where $||O||$ is the maximum of the absolute values of the eigenvalues of O. Let V be an arbitrary quantum operation on n qubits such that $V^\dagger O V$ is ECS, and f be an arbitrary polynomial in n. Then, there exists a polynomial-time randomized algorithm which outputs a real number r such that*

$$\Pr\left[\left|\langle 0^n|U^\dagger V^\dagger O V U|0^n\rangle - r\right| \le \frac{1}{f(n)}\right] \ge 1 - \frac{1}{\exp(n)}.$$

3 Target Quantum Circuits and Associated Fourier Coefficients

We focus on a new class of quantum circuits defined by CT states and ECS operations:

▶ **Definition 5.** *A quantum circuit C on n qubits initialized to $|0^n\rangle$ is CT-ECS if C consists of two blocks $C = VU$ such that U and V are polynomial-size quantum circuits, $U|0^n\rangle$ is CT, and $V^\dagger Z_j V$ is ECS for any $1 \le j \le n$, where Z_j is a Pauli-Z operation on the j-th qubit. In particular, C is CT-ECS$_1$ if C is CT-ECS and the operation $V^\dagger Z_j V$ associated with C is ECS$_1$ for any $1 \le j \le n$.*

To give examples of CT-ECS circuits, we consider the following quantum circuits on n qubits:

- An IQP circuit is of the form $H^{\otimes n} D H^{\otimes n}$, where D is a polynomial-size quantum circuit consisting of Z, CZ, and CCZ gates [5].
- A Clifford Magic circuit is of the form $ET^{\otimes n}H^{\otimes n}$, where E is a polynomial-size Clifford circuit, which consists of H, S, and CZ gates [20].
- A conjugated Clifford circuit is of the form $R_x(-\theta)^{\otimes n} R_z(-\phi)^{\otimes n} E R_z(\phi)^{\otimes n} R_x(\theta)^{\otimes n}$ for arbitrary real numbers ϕ, θ, where E is a polynomial-size Clifford circuit [3].
- A constant-depth quantum circuit is a polynomial-size quantum circuit F whose depth is constant [18, 4, 2].

In the common definition of an IQP circuit, D consists of more general Z-diagonal gates. However, for simplicity, we adopt the above definition. The resulting class includes a quantum circuit of the form $H^{\otimes n} D H^{\otimes n}$ such that its output probability distribution is anti-concentrated and not classically simulatable (under plausible assumptions).

We show that the above circuits are CT-ECS (in fact, CT-ECS$_1$). The proof can be found in the full version of the paper [16].

▶ **Lemma 6.** *Let C be one of the following quantum circuits on n qubits: an IQP, a Clifford Magic, a conjugated Clifford, or a constant-depth quantum circuit. Then, C is CT-ECS$_1$.*

Lemma 6 implies that there exist various CT-ECS circuits such that their output probability distributions are anti-concentrated and not classically simulatable (under plausible assumptions), although we do not know whether the output probability distributions of constant-depth quantum circuits are anti-concentrated.

To approximate the Fourier coefficients associated with CT-ECS circuits, we provide a general relation between a quantum circuit and the Fourier coefficients of its output probability distribution. The proof can be found in the full version of the paper [16].

▶ **Lemma 7.** *Let C be an arbitrary quantum circuit on n qubits initialized to $|0^n\rangle$ and p be its output probability distribution over $\{0,1\}^n$. Then,*

$$\widehat{p}(s) = \frac{1}{2^n}\langle 0^n|C^\dagger Z^s C|0^n\rangle$$

for any $s = s_1 \cdots s_n \in \{0,1\}^n$, where $Z^s = \bigotimes_{j=1}^n Z_j^{s_j}$, i.e., the tensor product of a Pauli-Z operation on the j-th qubit with $s_j = 1$ for any $1 \le j \le n$.

Using Theorem 4 and Lemma 7, we show that the low-degree Fourier coefficients of the output probability distribution of a CT-ECS circuit can be approximated classically in polynomial time. The proof can be found in the full version of the paper [16].

▶ **Lemma 8.** *Let C be an arbitrary CT-ECS circuit on n qubits and p be its output probability distribution over $\{0,1\}^n$. Let f be an arbitrary polynomial in n and s be an arbitrary element of $\{0,1\}^n$ with $|s| = O(1)$. Then, there exists a polynomial-time randomized algorithm which outputs a real number $\widehat{p}'(s)$ such that*

$$\Pr\left[|\widehat{p}(s) - \widehat{p}'(s)| \le \frac{1}{2^n f(n)}\right] \ge 1 - \frac{1}{\exp(n)}.$$

For any probability distribution p over $\{0,1\}^n$, it holds that $\widehat{p}(0^n) = 1/2^n$. Thus, when we consider a classical algorithm for approximating $\widehat{p}(s)$, we only consider an algorithm that outputs $1/2^n$ when $s = 0^n$. This simplifies the analysis in the following sections.

4 CT-ECS Circuits under Noise Model A

In this section, we prove Theorem 1. Its precise statement is as follows:

▶ **Theorem 1.** *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p satisfies $\sum_{x\in\{0,1\}^n} p(x)^2 \le \alpha/2^n$ for some known constant $\alpha \ge 1$. Let $0 < \delta < 1$ be an arbitrary constant. We assume that a depolarizing channel with (possibly unknown) constant rate $0 < \varepsilon < 1$ is applied to each qubit after performing C, which yields the probability distribution $\widetilde{p}_{\mathrm{A}}$, and it is possible to choose a constant λ such that*

$$1 \le \frac{\varepsilon}{\lambda} \le 1 + \frac{1}{\frac{10\sqrt{\alpha}}{\delta}\log\frac{10\sqrt{\alpha}}{\delta}}.$$

Then, there exists an $n^{O\left(\frac{1}{\lambda}\log\frac{\alpha}{\delta}\right)}$-time randomized algorithm which outputs (a classical description of) a probability distribution $\widetilde{q}_{\mathrm{A}}$ over $\{0,1\}^n$ such that

$$\Pr\left[||\widetilde{p}_{\mathrm{A}} - \widetilde{q}_{\mathrm{A}}||_1 \le \delta\right] \ge 1 - \frac{1}{\exp(n)}$$

and $\widetilde{q}_{\mathrm{A}}$ is classically samplable in time $n^{O\left(\frac{1}{\lambda}\log\frac{\alpha}{\delta}\right)}$.

Throughout the paper, the base of the logarithm is 2.

4.1 Proof of Theorem 1

We show that $\widetilde{p}_{\mathrm{A}}$ can be approximated by a function whose Fourier coefficients can be obtained classically in polynomial time. The proof can be found in the full version of the paper [16].

▶ **Lemma 9.** *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p satisfies $\sum_{x\in\{0,1\}^n} p(x)^2 \leq \alpha/2^n$ for some known constant $\alpha \geq 1$. Let $0<\delta<1$ be an arbitrary constant. We assume that a depolarizing channel with constant rate $0<\varepsilon<1$ is applied to each qubit after performing C, which yields the probability distribution $\widetilde{p}_{\mathrm{A}}$, and it is possible to choose a constant λ such that*

$$1 \leq \frac{\varepsilon}{\lambda} \leq 1 + \frac{1}{\frac{10\sqrt{\alpha}}{\delta}\log\frac{10\sqrt{\alpha}}{\delta}}.$$

Then, there exists an $n^{O\left(\frac{1}{\lambda}\log\frac{\alpha}{\delta}\right)}$-time randomized algorithm which outputs the Fourier coefficients of a function q over $\{0,1\}^n$ such that

$$\Pr\left[||\widetilde{p}_{\mathrm{A}} - q||_1 \leq \frac{\delta}{3}\right] \geq 1 - \frac{1}{\exp(n)}.$$

The remaining problem is to sample a probability distribution close to the function q. To do this, we use a classical sampling algorithm proposed by Bremner et al. [6], although the following analysis is slightly simpler than theirs. We can represent q as

$$q(x) = \sum_{s\in\{0,1\}^n,|s|\leq c} \widehat{q}(s)(-1)^{s\cdot x},$$

where $c = \left\lceil \frac{1}{\lambda}\log\frac{10\sqrt{\alpha}}{\delta}\right\rceil$, $\widehat{q}(s) = (1-\lambda)^{|s|}\widehat{p}'(s)$, $\widehat{p}'(s)$ is the approximate value of $\widehat{p}(s)$ for all $s\in\{0,1\}^n \setminus \{0^n\}$ with $|s|\leq c$, $\widehat{p}'(0^n) = 1/2^n$, and $\widehat{p}'(s) = 0$ for all $s\in\{0,1\}^n$ with $|s|>c$. It holds that $\sum_{x\in\{0,1\}^n} q(x) = 2^n\widehat{q}(0^n) = 2^n\widehat{p}'(0^n) = 1$. We define $S_\emptyset = 1$ and

$$S_y = \sum_{x\in\{0,1\}^n, x_1\cdots x_k = y} q(x)$$

for any $1\leq k\leq n$ and $y\in\{0,1\}^k$, where $\emptyset$ denotes the empty string. The sampling algorithm is described as follows:

1. Set $y \leftarrow \emptyset$.
2. Perform the following procedure n times:
   a. If $S_{yz} < 0$ for some $z\in\{0,1\}$, set $y \leftarrow y\bar{z}$, where $\bar{z} = 1-z$.
   b. Otherwise, set $y \leftarrow y0$ with probability S_{y0}/S_y and $y \leftarrow y1$ with probability $1 - S_{y0}/S_y$.
3. Output $y \in \{0,1\}^n$.

A direct calculation shows that

$$S_y = 2^{n-k} \sum_{s\in\{0,1\}^n,|s|\leq c, s_{k+1}\cdots s_n = 0^{n-k}} \widehat{q}(s)(-1)^{s_1\cdots s_k\cdot y}$$

for any $1\leq k\leq n$ and $y\in\{0,1\}^k$. Since S_y is classically computable in time $n^{O\left(\frac{1}{\lambda}\log\frac{\alpha}{\delta}\right)}$ (when we have $\widehat{q}(s)$ for all $s\in\{0,1\}^n$ with $|s|\leq c$), the runtime of the sampling algorithm is $n^{O\left(\frac{1}{\lambda}\log\frac{\alpha}{\delta}\right)}$. This sampling algorithm defines a real-valued function, denoted as $\mathrm{Alg}(q)$, that maps $y\in\{0,1\}^n$ to the probability that the sampling algorithm outputs y.

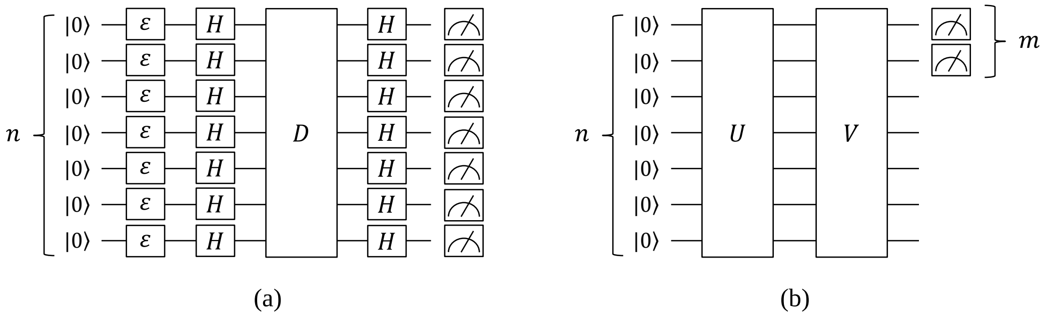

Figure 2 (a): IQP circuit $C = H^{\otimes n} D H^{\otimes n}$ under the input-noise model, where D consists of Z, CZ, and CCZ gates. The depolarizing channel $\mathcal{D}_\varepsilon$, which is represented as ε, is applied to each qubit initialized to $|0\rangle$. (b): CT-ECS circuit $C = VU$ in the noise-free setting, where $m = O(\log n)$.

Bremner et al. [6] defined $\mathrm{Fix}(q) = \left(\sum_{x\in\{0,1\}^n} q(x)\right) \mathrm{Alg}(q)$ and showed that

$$||q - \mathrm{Fix}(q)||_1 = 2 \sum_{x\in\{0,1\}^n, q(x)<0} |q(x)|$$

and $\mathrm{Alg}(q)$ is a probability distribution. Since $\sum_{x\in\{0,1\}^n} q(x) = 1$, we can deal with the situation where $\mathrm{Fix}(q) = \mathrm{Alg}(q)$. Therefore, Bremner et al.'s analysis of $\mathrm{Fix}(q)$ directly works for $\mathrm{Alg}(q)$. This allows us to show the following lemma, which is a special case of the property of $\mathrm{Alg}(q)$ shown by Bremner et al. [6]. The proof can be found in the full version of the paper [16].

▶ **Lemma 10.** *We assume that there exists a constant $0 < \delta < 1$ such that $||\widetilde{p}_{\mathrm{A}} - q||_1 \leq \delta/3$. Then, $||\widetilde{p}_{\mathrm{A}} - \mathrm{Alg}(q)||_1 \leq \delta$.*

The above lemmas immediately imply Theorem 1. Details can be found in the full version of the paper [16].

4.2 Input-Noise Model and Noise-Free Setting

We first deal with an input-noise model where $\mathcal{D}_\varepsilon$ is applied to each qubit (initialized to $|0\rangle$) only at the start of computation, and consider IQP circuits under this input-noise model as depicted in Fig. 2(a). We show that, when noise model **A** is replaced with this input-noise model, Theorem 1 holds for IQP circuits. Let $C = H^{\otimes n} D H^{\otimes n}$ be an arbitrary IQP circuit on n qubits, where D consists of Z, CZ, and CCZ gates. Let $\widetilde{p}_{\mathrm{in}}$ be the resulting probability distribution over $\{0,1\}^n$. The input state $|0^n\rangle$ affected by noise is represented as

$$\left[\left(1 - \frac{\varepsilon}{2}\right)|0\rangle\langle 0| + \frac{\varepsilon}{2}|1\rangle\langle 1|\right]^{\otimes n} = \sum_{y\in\{0,1\}^n} \left(1 - \frac{\varepsilon}{2}\right)^{n-|y|} \left(\frac{\varepsilon}{2}\right)^{|y|} X^y |0^n\rangle\langle 0^n| X^y,$$

where $X^y = H^{\otimes n} Z^y H^{\otimes n}$. A direct calculation shows that $\widetilde{p}_{\mathrm{in}} = \widetilde{p}_{\mathrm{A}}$, where $\widetilde{p}_{\mathrm{A}}$ is the output probability distribution of C under noise model **A** (with rate ε). This is because $H^{\otimes n} D H^{\otimes n} X^y = X^y H^{\otimes n} D H^{\otimes n}$ for any $y \in \{0,1\}^n$. Thus, when the output probability distribution of C is anti-concentrated, $\widetilde{p}_{\mathrm{A}}$ is classically simulatable by Theorem 1, and so is $\widetilde{p}_{\mathrm{in}}$. A similar proof shows that, when noise model **B** is replaced with the input-noise model, Theorem 2 holds for IQP circuits.

We then consider CT-ECS circuits in the noise-free setting as depicted in Fig. 2(b), and show that, when only $O(\log n)$ qubits are measured, any quantum circuit in a class of CT-ECS circuits on n qubits is classically simulatable. Let $C = VU$ be an arbitrary CT-ECS circuit on n qubits such that only $m = O(\log n)$ qubits are measured, and let p be its output probability distribution over $\{0,1\}^m$. We can represent p as

$$p(x) = \sum_{s \in \{0,1\}^m} \widehat{p}(s)(-1)^{s \cdot x}.$$

As with Lemma 7,

$$\widehat{p}(s) = \frac{1}{2^m} \langle 0^n | U^\dagger V^\dagger (Z^s \otimes I_{n-m}) V U | 0^n \rangle$$

for any $s \in \{0,1\}^m$, where I_{n-m} is the identity on the qubits that are not measured. By an argument similar to the proof of Lemma 8 [16], we want to approximate $\widehat{p}(s)$ for all $s \in \{0,1\}^m$. However, in the present case, $V^\dagger (Z^s \otimes I_{n-m}) V$ is the product of at most m ECS operations, which is not always ECS.

We assume that $V^\dagger (Z^s \otimes I_{n-m}) V$ is ECS for all $s \in \{0,1\}^m$. In this case, all the Fourier coefficients can be approximated classically in polynomial time with polynomial accuracy. Then, we define a function q over $\{0,1\}^m$ as

$$q(x) = \sum_{s \in \{0,1\}^m} \widehat{p}'(s)(-1)^{s \cdot x},$$

where $\widehat{p}'(s)$ is the approximate value of $\widehat{p}(s)$ for all $s \in \{0,1\}^m \setminus \{0^m\}$ and $\widehat{p}'(0^m) = 1/2^m$. Since $||p - q||_1$ is upper-bounded by some inverse polynomial in n, the classical sampling algorithm in Section 4.1 implies that p is classically simulatable, i.e, can be approximated by a classically samplable probability distribution, which is $\mathrm{Alg}(q)$, with polynomial accuracy in l_1 norm. For example, we consider a CT-ECS$_1$ circuit, which is a CT-ECS circuit such that the associated ECS operation $V^\dagger Z_j V$ is ECS$_1$. In this case, $V^\dagger (Z^s \otimes I_{n-m}) V$ is ECS for all $s \in \{0,1\}^m$, since it is the product of at most m ECS$_1$ (not ECS) operations. Thus, when only $O(\log n)$ qubits are measured, any CT-ECS$_1$ circuit on n qubits is classically simulatable in the noise-free setting.

5 CT-ECS Circuits under Noise Model B

In this section, we prove Theorem 2. Its precise statement is as follows:

▶ **Theorem 2.** *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p satisfies $\sum_{x \in \{0,1\}^n} p(x)^2 \leq \alpha/2^n$ for some known constant $\alpha \geq 1$. Let $0 < \delta < 1$ be an arbitrary constant. We assume that a depolarizing channel with (possibly unknown) constant rate $0 < \varepsilon_j < 1$ is applied to the j-th qubit after performing C for any $1 \leq j \leq n$, which yields the probability distribution $\widetilde{p}_{\mathrm{B}}$. Moreover, we assume that it is possible to choose a constant $\lambda_{\min}$ such that*

$$1 \leq \frac{\varepsilon_{\min}}{\lambda_{\min}} \leq 1 + \frac{1}{\frac{10\sqrt{\alpha}}{\delta} \log \frac{10\sqrt{\alpha}}{\delta}},$$

where $\varepsilon_{\min} = \min\{\varepsilon_j | 1 \leq j \leq n\}$, and it is possible to choose a constant λ_j such that

$$0 \leq \varepsilon_j - \lambda_j \leq \frac{\lambda_{\min}}{\frac{10\sqrt{\alpha}}{\delta} \log \frac{10\sqrt{\alpha}}{\delta}}$$

for any $1 \leq j \leq n$ with $\varepsilon_j \neq \varepsilon_{\min}$, where all numbers j with $\varepsilon_j \neq \varepsilon_{\min}$ are known. Then, there exists an $n^{O\left(\frac{1}{\lambda_{\min}} \log \frac{\alpha}{\delta}\right)}$-time randomized algorithm which outputs (a classical description of) *a probability distribution* $\widetilde{q}_{\mathrm{B}}$ *over* $\{0,1\}^n$ *such that*

$$\Pr\left[||\widetilde{p}_{\mathrm{B}} - \widetilde{q}_{\mathrm{B}}||_1 \le \left(1 + \frac{1}{1-\lambda_{\min}}\right)\delta\right] \ge 1 - \frac{1}{\exp(n)}$$

and $\widetilde{q}_{\mathrm{B}}$ *is classically samplable in time* $n^{O\left(\frac{1}{\lambda_{\min}}\log\frac{\alpha}{\delta}\right)}$.

Although the approximation accuracy of $\widetilde{q}_{\mathrm{B}}$ depends on $\lambda_{\min}$, a typical situation might be when $\varepsilon_{\min}$ is not so large. For example, when $\varepsilon_{\min} \le 1/2$, it holds that $||\widetilde{p}_{\mathrm{B}} - \widetilde{q}_{\mathrm{B}}||_1 \le 3\delta$.

We represent the effect of noise under noise model **B** as the effect under noise model **A** with rate $\varepsilon_{\min}$ and the remaining effects. To do this, for any $1 \le j \le n$ and $0 \le \delta_j \le 1$, we consider the noise operator $T^j_{\delta_j}$ [15] on real-valued functions over $\{0,1\}^n$ defined as

$$T^j_{\delta_j} f(x) = \mathrm{E}_{y_j \sim N_{\delta_j}(x_j)}[f(x_1, \cdots, x_{j-1}, y_j, x_{j+1}, \ldots, x_n)].$$

Here, $y_j \sim N_{\delta_j}(x_j)$ means that y_j is drawn as follows: y_j is x_j with probability $1 - \delta_j/2$ and $1 - x_j$ with probability $\delta_j/2$. Let p be an arbitrary probability distribution over $\{0,1\}^n$. It holds that $T^1_{\delta_1} \cdots T^n_{\delta_n} p(x)$ is equal to the probability of obtaining x by sampling an n-bit string according to p and then flipping its j-th bit with probability $\delta_j/2$ for any $1 \le j \le n$. Thus, the representation of $\widetilde{p}_{\mathrm{B}}$ described in Section 2.2 is obtained as $T^1_{\varepsilon_1} \cdots T^n_{\varepsilon_n} p(x)$. It is known that $T^1_{\delta_1} \cdots T^n_{\delta_n}$ is a contraction operator, i.e., $||T^1_{\delta_1} \cdots T^n_{\delta_n} f||_1 \le ||f||_1$ for any real-valued function f over $\{0,1\}^n$ [15].

Lemma 9 with these basic facts on the noise operator implies the following lemma on the representation of $\widetilde{p}_{\mathrm{B}}$ and its approximability. The proof can be found in the full version of the paper [16].

▶ **Lemma 11.** *Let C be an arbitrary CT-ECS circuit on n qubits such that its output probability distribution p satisfies $\sum_{x\in\{0,1\}^n} p(x)^2 \le \alpha/2^n$ for some known constant $\alpha \ge 1$. Let $0 < \delta < 1$ be an arbitrary constant. We assume that a depolarizing channel with constant rate $0 < \varepsilon_j < 1$ is applied to the j-th qubit after performing C for any $1 \le j \le n$, which yields the probability distribution $\widetilde{p}_{\mathrm{B}}$, and it is possible to choose a constant $\lambda_{\min}$ such that*

$$1 \le \frac{\varepsilon_{\min}}{\lambda_{\min}} \le 1 + \frac{1}{\frac{10\sqrt{\alpha}}{\delta}\log\frac{10\sqrt{\alpha}}{\delta}},$$

where $\varepsilon_{\min} = \min\{\varepsilon_j | 1 \le j \le n\}$. Then, the following items hold:

(i) $\widetilde{p}_{\mathrm{B}}(x) = T^1_{\delta_1} \cdots T^n_{\delta_n} \widetilde{p}_{\mathrm{A}}(x)$, *where*

$$\delta_j = \frac{\varepsilon_j - \varepsilon_{\min}}{1 - \varepsilon_{\min}}, \quad \widetilde{p}_{\mathrm{A}}(x) = \sum_{s\in\{0,1\}^n} (1-\varepsilon_{\min})^{|s|}\widehat{p}(s)(-1)^{s\cdot x}.$$

(ii) *There exists an $n^{O\left(\frac{1}{\lambda_{\min}}\log\frac{\alpha}{\delta}\right)}$-time randomized algorithm which outputs the Fourier coefficients of a function q over $\{0,1\}^n$ such that*

$$\Pr\left[||\widetilde{p}_{\mathrm{A}} - q||_1 \le \frac{\delta}{3}\right] \ge 1 - \frac{1}{\exp(n)}.$$

(iii) *Let $0 \le \delta'_j \le 1$ be an arbitrary constant for any $1 \le j \le n$. We assume that there exists a constant $\beta \ge 0$ such that $|\delta_j - \delta'_j| \le \beta$ for any $1 \le j \le n$. Then,*

$$\Pr\left[||T^1_{\delta_1} \cdots T^n_{\delta_n} q - T^1_{\delta'_1} \cdots T^n_{\delta'_n} q||_1 \le c\beta\sqrt{\alpha+1}\right] \ge 1 - \frac{1}{\exp(n)},$$

where $c = \left\lceil \frac{1}{\lambda_{\min}} \log\frac{10\sqrt{\alpha}}{\delta}\right\rceil$.

Lemma 11 implies Theorem 2. Details can be found in the full version of the paper [16].

References

1 S. Aaronson and A. Arkhipov. The computational complexity of linear optics. *Theory of Computing*, 9(4):143–252, 2013.

2 J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert. Architectures for quantum simulation showing a quantum speedup. *Physical Review X*, 8(2):021010, 2018.

3 A. Bouland, J. F. Fitzsimons, and D. E. Koh. Complexity classification of conjugated Clifford circuits. In *Proceedings of the 33rd Computational Complexity Conference (CCC)*, volume 102 of *Leibniz International Proceedings in Informatics*, pages 21:1–21:25, 2018.

4 M. J. Bremner, R. Jozsa, and D. J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A*, 467(2126):459–472, 2011.

5 M. J. Bremner, A. Montanaro, and D. J. Shepherd. Average-case complexity versus approximate simulation of commuting quantum computations. *Physical Review Letters*, 117(8):080501, 2016.

6 M. J. Bremner, A. Montanaro, and D. J. Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, 2017.

7 K. Bu and D. E. Koh. Efficient classical simulation of Clifford circuits with nonstabilizer input states. *Physical Review Letters*, 123(17):170502, 2019.

8 S. Fenner, F. Green, S. Homer, and Y. Zhang. Bounds on the power of constant-depth quantum circuits. In *Proceedings of Fundamentals of Computation Theory (FCT)*, volume 3623 of *Lecture Notes in Computer Science*, pages 44–55, 2005.

9 K. Fujii and S. Tamate. Computational quantum-classical boundary of noisy commuting quantum circuits. *Scientific Reports*, 6:25598, 2016.

10 X. Gao and L. Duan. Efficient classical simulation of noisy quantum computation, 2018. arXiv:1810.03176.

11 A. W. Harrow and A. Montanaro. Quantum computational supremacy. *Nature*, 549:203–209, 2017.

12 D. E. Koh. Further extensions of Clifford circuits and their classical simulation complexities. *Quantum Information and Computation*, 17(3&4):262–282, 2017.

13 M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

14 H. Nishimura and M. Ozawa. Computational complexity of uniform quantum circuit families and quantum Turing machines. *Theoretical Computer Science*, 276(1–2):147–181, 2002.

15 R. O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

16 Y. Takahashi, Y. Takeuchi, and S. Tani. Classically simulating quantum circuits with local depolarizing noise, 2020. arXiv:2001.08373.

17 Y. Takeuchi and Y. Takahashi. Ancilla-driven instantaneous quantum polynomial time circuit for quantum supremacy. *Physical Review A*, 94(6):062336, 2016.

18 B. M. Terhal and D. P. DiVincenzo. Adaptive quantum computation, constant-depth quantum circuits and Arthur-Merlin games. *Quantum Information and Computation*, 4(2):134–145, 2004.

19 M. van den Nest. Simulating quantum computers with probabilistic methods. *Quantum Information and Computation*, 11(9&10):784–812, 2011.

20 M. Yoganathan, R. Jozsa, and S. Strelchuk. Quantum advantage of unitary Clifford circuits with magic state inputs. *Proceedings of the Royal Society A*, 475(2225), 2019.