Markov Chains and Hitting Times for Error Accumulation in Quantum Circuits

Ma, Long; Sanders, Jaron

DOI
10.1007/978-3-030-92511-6_3

Publication date
2021

Document Version
Final published version

Published in
Performance Evaluation Methodologies and Tools - 14th EAI International Conference, VALUETOOLS 2021, Proceedings

Markov Chains and Hitting Times for Error Accumulation in Quantum Circuits

Long Ma[1(✉)] and Jaron Sanders[2]

[1] Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, Delft, Netherlands
l.ma-2@tudelft.nl

[2] Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands
jaron.sanders@tue.nl

Abstract. We study a classical model for the accumulation of errors in multi-qubit quantum computations. By modeling the error process in a quantum computation using two coupled Markov chains, we are able to capture a weak form of time-dependency between errors in the past and future. By subsequently using techniques from the field of discrete probability theory, we calculate the probability that error quantities such as the fidelity and trace distance exceed a threshold analytically. The formulae cover fairly generic error distributions, cover multi-qubit scenarios, and are applicable to the randomized benchmarking protocol. To combat the numerical challenge that may occur when evaluating our expressions, we additionally provide an analytical bound on the error probabilities that is of lower numerical complexity. Besides this, we study a model describing continuous errors accumulating in a single qubit. Finally, taking inspiration from the field of operations research, we illustrate how our expressions can be used to decide how many gates one can apply before too many errors accumulate with high probability, and how one can lower the rate of error accumulation in existing circuits through simulated annealing.

Keywords: Markov chains · Error accumulation · Quantum circuits

1 Introduction

The development of a quantum computer is expected to revolutionize computing by being able to solve hard computational problems faster than any classical computer [37]. However, present-day state-of-the-art quantum computers are prone to errors in their calculations due to physical effects such as unwanted qubit–qubit interactions, qubit crosstalk, and state leakage [38]. Minor errors can be corrected, but error correction methods will still be overwhelmed once too

Q. Zhao and L. Xia (Eds.): VALUETOOLS 2021, LNICST 404, pp. 36–55, 2021.
https://doi.org/10.1007/978-3-030-92511-6_3

many errors occur [12,20,31]. Quantum circuits with different numbers of qubits and circuit depths have been designed to implement algorithms more reliably [16], and the susceptibility of a circuit to the accumulation of errors remains an important evaluation criterion. We therefore study now Markov chains that provide a model for the accumulation of errors in quantum circuits. Different types of errors [21] that can occur and are included in our model are e.g. Pauli channels [37], Clifford channels [23,34], depolarizing channels [37], and small rotational errors [7,26]. If the random occurrence of such errors only depends on the last state of the quantum mechanical system, then the probability that error quantities such as the fidelity and trace distance accumulate beyond a threshold can be related to different hitting time distributions of two coupled Markov chains [8]. These hitting time distributions are then calculated analytically using techniques from probability theory and operations research.

Error accumulation models that share similarities with the Markov chains under consideration here can primarily be found in the literature on randomized benchmarking [46]. From the modeling point of view, the dynamical description of error accumulation that we adopt is shared in [3,27,33,43]. These articles however do not explicitly tie the statistics of error accumulation to a hitting time analysis of a coupled Markov chain. Furthermore, while Markovianity assumptions on noise are common [14], the explicit mention of an underlying random walk is restricted to a few papers only [3,17]. From the analysis point of view, research on randomized benchmarking has predominantly focused on generalizing expressions for the expected fidelity over time. For example, the expected decay rates of the fidelity are analyzed for cases of randomized benchmarking with restricted gate sets [9], Gaussian noise with time-correlations [15], gate-dependent noise [43], and leakage errors [45]; and the expected loss rate of a protocol related to randomized benchmarking is calculated in [10,33,39,43,44]. In this article, we focus instead on the probability distributions of both the error and maximum error in the Markov chain model – which capture the statistics in more detail than an expectation – for arbitrary distance measures, and in random as well as nonrandom quantum circuits. Finally, [3,33,43,45] resort to perturbation or approximate analyses (via e.g. Taylor expansions, and independence or decorrelation assumptions) to characterize the fidelity, whereas here we provide the exact, closed-form expressions for the distributions using the theory of Markov chains.

To be precise: this article first studies a model for discrete Markovian error accumulation in a multi-qubit quantum circuit. We suppose for simplicity that both the quantum gates and errors belong to a finite unitary group $\mathcal{G}_n \subseteq \mathcal{U}(2^n)$, where $\mathcal{U}(2^n)$ is the unitary group for n qubits. The group $\mathcal{G}_n$ can e.g. be the generalized Pauli group (i.e., the discrete Heisenberg–Weyl group), or the Clifford group. By modeling the quantum computation with and without errors as two coupled Markov chains living on the state space consisting of pairs of elements from these groups, we are able to capture a weak form of time-dependency within the process of error accumulation. To see this, critically note that the assumption of a Markov property does not imply that the past and the future in the quantum computation are independent given any information concerning the present [8]. We must also note that while the individual elements of our two-dimensional

Markov chain belong to a group, the two-dimensional Markov chain itself, here, is generally not a random walk on a group. Lastly, our Markov chain model works for an arbitrary number of qubits. These model features are all relevant to the topic of error modeling in quantum computing, and since the Markov property is satisfied in randomized benchmarking, the model has immediate application. The method is generic in the sense that any measure of distance between two pure quantum states may be used to quantify the error, and that it allows for a wide range of error distributions. The method can handle nonuniform, gate-, and time-dependent errors. Concretely, for arbitrary measure of distance and a wide range of error distributions, we will calculate (i) the expected error at time t, (ii) the probability that an error is larger than a threshold δ at time t, and (iii) the probability that the error has *ever* been larger than a threshold δ before time t, and we do so both for random and nonrandom circuits.

In addition to studying a model for discrete Markovian error accumulation in quantum circuits, we also briefly study a random walk model on the three-dimensional sphere [40]. This model is commonly used to describe the average dephasing of a single qubit (or spin) [24]. Using this model, we characterize the distribution and expectation of the trace distance measuring the error that is accumulated over time. These derivations are, essentially, refinements that provide information about the higher-order statistics of the error accumulation in a single qubit.

The approach taken in this article is a hybrid between classical probability theory and quantum information theory. This hybridization allows us to do quite detailed calculations, but not every quantum channel will satisfy the necessary assumptions such as Markovianity of the error distribution. On the other hand, in cases where one introduces their own source of randomness (such as in randomized benchmarking), the assumptions are met naturally. It should furthermore be noted that the numerical complexity of the exact expressions we provide is high for large quantum circuits. The precise difficulty of evaluating our expressions depends on the particulars of the quantum circuit one looks at. For practical purposes, we therefore also provide an explicit bound on the maximum error probability that is of lower numerical complexity. Reference [27] is relevant to mention here, because similar to our observations, these authors also note the generally high computational complexity of error analysis in quantum circuits. The issue is approached in [27] differently and in fact combinatorially by converting circuits into directed graphs, tracing so-called fault-paths through these graphs, and therewith estimating the success rates of circuits.

Finally, we use the expressions that describe how likely it is that errors accumulate to answer two operational questions that will help advance the domain of practical quantum computing [29]. First, we calculate and bound analytically how many quantum gates $t^{\star}_{\delta,\gamma}$ one can apply before an error measure of your choice exceeds a threshold δ with a probability above γ. This information is useful for deciding how often a quantum computer should perform repairs on qubits, and is particularly opportune at this moment since quantum gates fail $O(0.1\text{–}1\%)$ of the time [29]. Related but different ideas can be found in [21, §2.3], where the accumulation of bit-flips and rotations on a repetition code is studied and a time to failure is derived, and in [25, §V], where an upper bound on the number of necessary measurements for a randomized benchmarking protocol

is derived. Second, using techniques from optimization, we design a simulated annealing method that improves existing circuits by swapping out gate pairs to achieve lower rates of error accumulation. There is related literature where the aim is to reduce the circuit depth [2,28,35], but an explicit expression for error accumulation has not yet been leveraged in the same way. Moreover, we also discuss conditions under which this tailor-made method is guaranteed to find the best possible circuit. Both of these excursions illustrate how the availability of an analytical expression for the accumulation of errors allows us to proceed with second-tier optimization methods to facilitate quantum computers in the long-term. We further offer an additional proof-of-concept that simulated annealing algorithms can reduce error accumulation rates in existing quantum circuits when taking error distributions into account: we illustrate that the misclassification probability in a circuit that implements the Deutsch–Jozsa Algorithm for one classical bit [11,13] can be lowered by over 40%. In this proof of concept we have chosen an example error distribution that is gate-dependent and moreover one that is such that *not* applying a gate gives the lowest error rate in this model; applying a single-qubit gate results in a medium error rate; and applying a two-qubit gate gives the largest probability that an error may occur.

This paper is structured as follows. In Sect. 2, we give the model aspects pertaining to the quantum computation (gates, error dynamics, and error measures) and we introduce the coupled Markov chain to describe error accumulation. In Sect. 3, we provide the relation between the probability of error and the hitting time distributions, and we derive the error distributions as well as its bound. We also calculate the higher-order statistics of an error accumulation model for a single qubit that undergoes (continuous) random phase kicks and depolarization. In Sect. 4, we illustrate our theoretical results by comparing to numerical results of a quantum simulator we wrote for this article. In Sect. 5, we discuss the simulated annealing scheme. Finally, we conclude in Sect. 6.

2 Model and Coupled Markov Chain

2.1 Gates and Errors in Quantum Computing

It is generally difficult to describe large quantum systems on a classical computer for the reason that the state space required increases exponentially in size with the number of qubits [36]. However, the stabilizer formalism is an efficient tool to analyze such complex systems [18]. Moreover, the stabilizer formalism covers many paradoxes in quantum mechanics [1], including the Greenberger–Horne–Zeilinger (GHZ) experiment [22], dense quantum coding [5], and quantum teleportation [4]. Specifically, the stabilizer circuits are the smallest class of quantum circuits that consist of the following four gates: $\omega = \mathrm{e}^{\mathrm{i}\pi/4}$, $H = (1/\sqrt{2})\big((1,1);(1,-1)\big)$, $S = \big((1,0);(0,\mathrm{i})\big)$, and $Z_c = \big((1,0,0,0);(0,1,0,0);(0,0,1,0);(0,0,0,-1)\big)$. These four gates are closed under the operations of tensor product and composition [42]. As a consequence of the Gottesman–Knill theorem, stabilizer circuits can be efficiently simulated on a classical computer [19].

Unitary stabilizer circuits are also known as the Clifford circuits; the Clifford group $\mathcal{C}_n$ can be defined as follows. First: let $P \triangleq \{I, X, Y, Z\}$ denote the Pauli matrices, so $I = ((1,0);(0,1))$, $X = ((0,1);(1,0))$, $Y = ((0,-i);(i,0))$, and $Z = ((1,0);(0,-1))$, and let $P_n \triangleq \{\sigma_1 \otimes \cdots \otimes \sigma_n \mid \sigma_i \in P\}$ denote the Pauli matrices on n qubits. The Pauli matrices are commonly used to model errors that can occur due to the interactions of the qubit with its environment [41]. In the case of a single qubit, the matrix I represents that there is no error, the matrix X that there is a bit-flip error, the matrix Z that there is a phase-flip error, and the matrix Y that there are both a bit-flip and a phase-flip error. The multi-qubit case interpretations follow analogously. Second: let $P_n^* = P_n / I^{\otimes n}$. We now define the Clifford group on n qubits by $\mathcal{C}_n \triangleq \{U \in \mathcal{U}(2^n) \mid \sigma \in \pm P_n^* \Rightarrow U\sigma U^\dagger \in \pm P_n^*\}/\mathcal{U}(1)$.

The fact that $\mathcal{C}_n$ is a group can be verified by checking the two necessary properties (see our extended version [32]). The Clifford group on n qubits is finite [30], and we will ignore the global phase throughout this paper for convenience; its size is then $|\mathcal{C}_n| = 2^{n^2+2n} \prod_{i=1}^{n} \left(4^i - 1\right)$. Moreover, for a single qubit, a representation for the Clifford group $\mathcal{C}_1 = \{C_1, C_2, \cdots, C_{24}\}$ can then be enumerated and its elements are for example shown in [46] and [3].

2.2 Dynamics of Error Accumulation

Suppose that we had a faultless, perfect quantum computer. Then a faultless quantum mechanical state ρ_t at time t could be calculated under a gate sequence $\mathcal{U}_\tau = \{U_1, \ldots, U_\tau\}$ from the initial state $\rho_0 \triangleq |\psi_0\rangle \langle\psi_0|$. Here $\tau < \infty$ denotes the sequence length, and $t \in \{0, 1, \cdots, \tau\}$ enumerates the intermediate steps. On the other hand, with an imperfect quantum computer, a possibly faulty quantum mechanical state σ_t at time t would be calculated under both $\mathcal{U}_t$ and some (unknown) noise sequence $\mathcal{E}_t = \{\Lambda_1, \ldots, \Lambda_t\}$ starting from an initial state $\sigma_0 \triangleq |\Psi_0\rangle \langle\Psi_0|$ possibly different from ρ_0. We define the set of all pure states for n qubits as $\mathcal{S}^n$ and consider the situation that $|\psi_0\rangle, |\Psi_0\rangle \in \mathcal{S}^n$.

To be precise, define for the faultless quantum computation

$$\rho_t \triangleq |\psi_t\rangle \langle\psi_t| = U_t |\psi_{t-1}\rangle \langle\psi_{t-1}| U_t^\dagger$$

for times $t = 1, 2, \ldots, \tau$. Let $X_t \triangleq U_t U_{t-1} \cdots U_1$ be shorthand notation such that $\rho_t = X_t \rho_0 X_t^\dagger$. For the possibly faulty quantum computation, define

$$\sigma_t \triangleq |\Psi_t\rangle \langle\Psi_t| = \Lambda_t U_t |\Psi_{t-1}\rangle \langle\Psi_{t-1}| U_t^\dagger \Lambda_t^\dagger$$

for times $t = 1, 2, \ldots, \tau$, respectively. Introduce also the shorthand notation $Y_t \triangleq \Lambda_t U_t \Lambda_{t-1} U_{t-1} \cdots \Lambda_1 U_1$ such that $\sigma_t = Y_t \sigma_0 Y_t^\dagger$. The analysis in this paper can immediately be extended to the case where errors (also) precede the gate. The error accumulation process is also illustrated in Fig. 1.

2.3 Distance Measures for Quantum Errors

The error can be quantified by any measure of distance between the faultless quantum-mechanical state ρ_t and the possibly faulty quantum-mechanical

a) Faultless computation:

$$\rho_0 \xrightarrow{U_1} \rho_1 \xrightarrow{U_2} \cdots \xrightarrow{U_{\tau-1}} \rho_{\tau-1} \xrightarrow{U_\tau} \rho_\tau$$

b) Potentially faulty computation:

$$\sigma_0 \xrightarrow{\Lambda_1 U_1} \sigma_1 \xrightarrow{\Lambda_2 U_2} \cdots \xrightarrow{\Lambda_{\tau-1} U_{\tau-1}} \sigma_{\tau-1} \xrightarrow{\Lambda_\tau U_\tau} \sigma_\tau$$

Fig. 1. Schematic depiction of the coupled quantum mechanical states ρ_t and σ_t for times $t = 0, 1, \cdots, \tau$. a) Faultless computation. The state ρ_t is calculated based on a gate sequence $\mathcal{U}_t = \{U_1, \ldots, U_t\}$ from the initial state ρ_0. b) Potentially faulty computation. The state σ_t is calculated using *the same* gate sequence $\mathcal{U}_t = \{U_1, \ldots, U_t\}$ and an additional error sequence $\mathcal{E}_t = \{\Lambda_1, \ldots, \Lambda_t\}$. The final state σ_τ can depart from the faultless state ρ_τ because of errors.

state σ_t for steps $t = 0, 1, \ldots, \tau$. For example, we can use the fidelity $F_t \triangleq \mathrm{Tr}\sqrt{{\rho_t}^{1/2}\sigma_t{\rho_t}^{1/2}}$ [37], or the Schatten d-norm [6] defined by

$$D_t \triangleq \|\sigma_t - \rho_t\|_d = \tfrac{1}{2}\mathrm{Tr}\left[\left\{(\sigma_t - \rho_t)^\dagger(\sigma_t - \rho_t)\right\}^{\frac{d}{2}}\right]^{\frac{1}{d}}$$

for any $d \in [1, \infty)$. The Schatten d–norm reduces to the trace distance for $d = 1$, the Frobenius norm for $d = 2$, and the spectral norm for $d = \infty$. In the case of one qubit, the trace distance between quantum-mechanical states ρ_t and σ_t equals half of the Euclidean distance between ρ_t and σ_t when representing them on the Bloch sphere [37]. It is well known that the trace distance is invariant under unitary transformations [37]; a fact that we leverage in Sect. 3.

In this paper, we are going to analyze the statistical properties of some arbitrary distance measure (one may choose) between the quantum mechanical states ρ_t and σ_t for times $t = 0, 1, \ldots, \tau$. For illustration, we will state the results in terms of the Schatten d–norm, and so are after its expectation $\mathbb{E}[D_t]$, as well as the probabilities $\mathbb{P}[D_t \leq \delta]$, $\mathbb{P}[\max_{0 \leq s \leq t} D_s \leq \delta]$. Throughout this paper, the operator $\mathbb{P}$ and thus also $\mathbb{E}$ are with respect to a sufficiently rich probability space $(\Omega, \mathbb{P}, \mathcal{F})$ that each time can describe the Markov chain being considered.

3 Error Accumulation

3.1 Discrete, Random Error Accumulation (Multi-Qubit Case)

Following the model described in Sect. 2 and illustrated in Fig. 1 and Fig. 2a, we define the gate pairs $Z_t \triangleq (X_t, Y_t)$ for $t = 0, 1, 2, \ldots, \tau$, and suppose that $Z_0 = z_0$ with probability one where $z_0 = (x_0, y_0)$ is deterministic and given a priori. Note in particular that if the initial state is prepared without error, then $\rho_0 = \sigma_0$ and consequently $z_0 = (I^{\otimes n}, I^{\otimes n})$. If on the other hand the initial state is prepared incorrectly as $y_0 |\psi_0\rangle$ instead of $|\psi_0\rangle$, then $z_0 = (I^{\otimes n}, y_0)$.

The Case of Random Circuits. We consider first the scenario that each next gate is selected randomly and independently from everything but the last system state. This assumption is satisfied in the randomized benchmarking protocol [3,9,14,15,17,27,33,43–46]. The probabilities $\mathbb{P}_{z_0}[D_t > \delta]$ and $\mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s \leq \delta]$ can then be calculated once the initial states $|\psi_0\rangle$, $|\Psi_0\rangle$ and the *transition matrix* are known. Here, the subscript z_0 reminds us of the initial state the Markov chain is started from.

Let the transition matrix of the Markov chain $\{Z_t\}_{t\geq 0}$ be denoted element-wise by $P_{z,w} \triangleq \mathbb{P}[Z_{t+1} = w | Z_t = z]$ for $z = (x, y), w = (u, v) \in \mathcal{G}_n^2$. The transition matrix satisfies $P \in [0,1]^{|\mathcal{G}_n|^2\times|\mathcal{G}_n|^2}$ and the elements of each of its rows sum to one. Let $P^{(t)}_{z_0,w} \triangleq \mathbb{P}[Z_t = w | Z_0 = z_0] = (P^t)_{z_0,w}$ stand in for the probability that the process is at state w at time t starting from $Z_0 = z_0$. Note that the second equality follows from the Markov property [8].

We are now after the probability that the distance D_t is larger than a threshold δ. We define thereto the set of *δ-bad gate pairs* by

$$\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta} \triangleq \left\{(x, y) \in \mathcal{G}_n^2 \middle| \|x\rho_0x^\dagger - y\sigma_0y^\dagger\|_d > \delta\right\} \tag{1}$$

for $|\psi_0\rangle, |\Psi_0\rangle \in \mathcal{S}^n, \delta \geq 0$, as well as the *hitting time* of any set $\mathcal{A} \subseteq \mathcal{G}_n^2$ by

$$T_{\mathcal{A}} \triangleq \inf\{t \geq 0 | Z_t \in \mathcal{A}\} \tag{2}$$

with the convention that $\inf \phi = \infty$. Note that $T_{\mathcal{A}} \in \mathbb{N}_0 \cup \{\infty\}$ and that it is random. With Definitions (1), (2), we have the convenient representation

$$\mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s \leq \delta] = 1 - \mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s > \delta] = 1 - \mathbb{P}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}} \leq t] \tag{3}$$

for this homogeneous Markov chain. As a consequence of (3), the analysis comes down to an analysis of the hitting time distribution for this coupled Markov chain (Fig. 2b).

Results. Define the matrix $B^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta} \in [0,1]^{|\mathcal{G}_n|^2\times|\mathcal{G}_n|^2}$ element-wise by

$$\left(B^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}\right)_{z,w} \triangleq \begin{cases} P_{z,w} & \text{if} \quad w \notin \mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}, \\ 0 & \text{otherwise.} \end{cases}$$

Let the initial state vector be denoted by e_{z_0}, a $|\mathcal{G}_n|^2 \times 1$ vector with just the z_0-th element 1 and the others 0. Also let $1_{\mathcal{A}}$ denote the $|\mathcal{G}_n|^2 \times 1$ vector with ones in every coordinate corresponding to an element in the set $\mathcal{A}$. Let the transpose of an arbitrary matrix A be denoted by A^{T} and defined element-wise $(A^{\mathrm{T}})_{i,j} = A_{j,i}$. Finally, we define a $|\mathcal{G}_n|^2 \times 1$ vector $d^{|\Psi_0\rangle}_{|\psi_0\rangle} = \left(\|x\rho_0x^\dagger - y\sigma_0y^\dagger\|_d\right)_{(x,y)\in\mathcal{G}_n^2}$ enumerating all possible Schatten d-norm distances. We now state our first result, and defer to [32] for its proof:

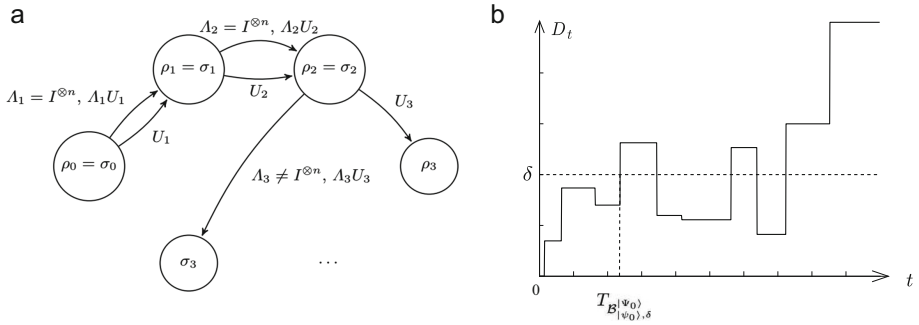

Fig. 2. a) Coupled chain describing the quantum circuit with errors. In this depiction, we start from *the same* initial state for simplicity. Here an error $\Lambda_3 \neq I^{\otimes n}$ occurs as the third gate is applied. Note that the coupled chain ρ_t, σ_t separates. b) Schematic diagram of the hitting time $T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}$.

Proposition 1 (Error accumulation in random circuits). *For any $z_0 \in \mathcal{G}_n^2$, $\delta \geq 0$, $t = 0, 1, \ldots, \tau < \infty$: the expected error is given by $\mathbb{E}_{z_0}[D_t] = e_{z_0}^{\mathrm{T}} P^t d^{|\Psi_0\rangle}_{|\psi_0\rangle}$. Similarly, the probability of error is given by*

$$\mathbb{P}_{z_0}[D_t > \delta] = e_{z_0}^{\mathrm{T}} P^t 1_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}, \tag{4}$$

and is nonincreasing in δ. Furthermore; if $z_0 \notin \mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}$, the probability of maximum error is given by

$$\mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s > \delta] = \sum_{s=1}^{t} e_{z_0}^{\mathrm{T}} \big(B^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}\big)^{s-1} \big(P - B^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}\big) 1_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}, \tag{5}$$

and otherwise it equals one. Lastly, (5) is nonincreasing in δ, and nondecreasing in t.

The probability in (5) is a more stringent error measure than (4) is. The event $\{\max_{0\leq s\leq t} D_s < \delta\}$ implies after all that the error D_t has always been below the threshold δ up to and including at time t. The expected error $\mathbb{E}_{z_0}[D_t]$ and probability $\mathbb{P}_{z_0}[D_t > \delta]$ only concern the error *at* time t. Additionally, (5) allows us to calculate the maximum number of gates that can be performed. That is, $\mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s > \delta] \leq \gamma$ as long as $t \leq t^\star_{\delta,\gamma} \triangleq \max\big\{t \in \mathbb{N}_0 \big| \mathbb{P}_{z_0}[\max_{0\leq s\leq t} D_s > \delta] \leq \gamma\big\}$. In words: at most $t^\star_{\delta,\gamma}$ gates can be applied before an accumulated error of size at least δ occurred with probability at least γ.

For general $\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}$, the explicit calculation of (5) can be numerically intensive. It is however possible to provide a lower bound of lower numerical complexity via the expected hitting time of the set $\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}$.

Lemma 1 (Lower bound for random circuits). *For any set $\mathcal{A} \subseteq \mathcal{G}_n^2$, the expected hitting times of a homogeneous Markov chain are the solutions to the*

linear system of equations $\mathbb{E}_z[T_{\mathcal{A}}] = 0$ *for* $z \in \mathcal{A}$, $\mathbb{E}_z[T_{\mathcal{A}}] = 1 + \sum_{w \notin \mathcal{A}} P_{z,w}\mathbb{E}_w[T_{\mathcal{A}}]$ *for* $z \notin \mathcal{A}$. *Furthermore; for any* $z_0 \in \mathcal{G}_n^2$, $\delta \geq 0$, $t = 0, 1, \ldots, \tau < \infty$*:*

$$\mathbb{P}_{z_0}[\max_{0 \leq s \leq t} D_s > \delta] \geq 0 \vee \Big(1 - \frac{\mathbb{E}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}]}{t+1}\Big). \tag{6}$$

Here $a \vee b \triangleq \max\{a, b\}$.

A proof of (6) can be found in our extended version [32]. As a consequence of Lemma 1, $\mathbb{P}_{z_0}[\max_{0 \leq s \leq t} D_s > \delta] \geq \gamma$ when $t \geq \mathbb{E}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}]/(1-\gamma) - 1$, and in particular $\mathbb{P}_{z_0}[\max_{0 \leq s \leq t} D_s > 0] > 0$ when $t \geq \mathbb{E}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,0}}]$. The values in the right-hand sides are thus upper bounds to the number of gates $t^{\star}_{\delta,\gamma}$ one can apply before δ error has occurred with probability γ:

$$t^{\star}_{\delta,\gamma} \leq \mathbb{E}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,0}}] \wedge \Big(\frac{\mathbb{E}_{z_0}[T_{\mathcal{B}^{|\Psi_0\rangle}_{|\psi_0\rangle,\delta}}]}{1-\gamma} - 1\Big)$$

for $\delta \geq 0, \gamma \in [0, 1]$. Here, $a \wedge b \triangleq \min\{a, b\}$.

Limitations of the Method: Types of Quantum Noise Channels. The approach taken in this article is a hybrid between classical probability theory and quantum information theory. The results of this article are therefore not applicable to all quantum channels, and it is important that we signal you the limitations.

As an illustrative example, consider the elementary circuit of depth $\tau = 1$ with $n = 1$ qubit, in which the one gate is restricted to the Clifford group $\{C_1, \ldots, C_{24}\}$, say. For such an elementary circuit, this article describes a classical stochastic process that chooses one of twenty-four quantum noise channel $\mathcal{F}^{(1)}, \ldots \mathcal{F}^{(24)}$ say according to some arbitrary classical probability distribution $\{p_i(\rho)\}$, i.e.,

$$\rho_0 \to \rho_1 = \mathcal{F}(\rho_0) = \begin{cases} \mathcal{F}^{(1)}(\rho_0) = C_1\rho_0 C_1^{\dagger} & \text{w.p. } p_1(\rho_0), \\ \mathcal{F}^{(2)}(\rho_0) = C_2\rho_0 C_2^{\dagger} & \text{w.p. } p_2(\rho_0), \\ \ldots & \\ \mathcal{F}^{(24)}(\rho_0) = C_{24}\rho_0 C_{24}^{\dagger} & \text{w.p. } p_{24}(\rho_0). \end{cases} \tag{7}$$

Here, the classical probability distribution $\{p_i(\rho)\}$ may be chosen arbitrarily, and depend on the initial quantum state ρ_0 as indicated. For this elementary quantum circuit of depth $\tau = 1$ with $n = 1$ qubit, (7) characterizes the set of stochastic processes covered by our results in its entirety.

For example, Proposition 1 cannot be applied to the deterministic process

$$\rho_0 \to \rho_1 = \Big\{\mathcal{E}^{(1)}(\rho_0) = (1-p)\rho_0 + pY\rho_0 Y^{\dagger} \text{ w.p. } 1,$$

nor to the deterministic process

$$\rho_0 \to \rho_1 = \Big\{\mathcal{E}^{(2)}(\rho_0) = (1-p)\rho_0 + \tfrac{p}{2}U\rho_0 U^{\dagger} + \tfrac{p}{2}U^{\dagger}\rho_0 U \text{ w.p. } 1.$$

Here, $p \in (0,1)$ can be chosen arbitrarily and $U = \mathrm{e}^{-\mathrm{i}\pi Y/4}$ is a Clifford gate. The reason is that $\big(\mathcal{F}^{(1)} \neq \mathcal{F}^{(2)} \neq \cdots \neq \mathcal{F}^{(24)}\big) \neq \big(\mathcal{E}^{(1)} = \mathcal{E}^{(2)}\big)$ by the unitary freedom in the operator-sum representation [37, Thm. 8.2]. A meticulous reader will now note that the example quantum channels $\mathcal{E}^{(1)}, \mathcal{E}^{(2)}$ are however *averages* of two particular stochastic processes $\mathcal{F}$. That is: if $p_I = 1-p, p_Y = p$, then $\mathcal{E}^{(1)}(\rho) = \mathbb{E}[\mathcal{F}(\rho)]$; or if $p_I = 1-p, p_U = p_{U^\dagger} = \frac{p}{2}$, then $\mathcal{E}^{(2)}(\rho) = \mathbb{E}[\mathcal{F}(\rho)]$.

The Case of Nonrandom Circuits. Suppose that the gate sequence $\mathcal{U}_\tau = \{U_1, ..., U_\tau\}$ is fixed *a priori* and that it is not generated randomly. Because the gate sequence is nonrandom, we have now that the faultless state $\rho_t = X_t \rho_0 X_t^\dagger$ is deterministic for times $t = 0, 1, \ldots, \tau$. On the other hand the potentially faulty state $\sigma_t = Y_t \rho_0 Y_t^\dagger$ is still (possibly) random.

We can now use a lower dimensional Markov chain to represent the system. To be precise: we will now describe the process $\{Y_t\}_{t\geq 0}$ (and consequently $\{\sigma_t\}_{t\geq 0}$) as an *inhomogeneous Markov chain*. Its transition matrices will now be time-dependent and given element-wise by $Q_{y,v}(t) = \mathbb{P}[Y_{t+1} = v | Y_t = y]$ for $y, v \in \mathcal{G}_n, t \in \{0, 1, \ldots, \tau - 1\}$. Letting $Q_{y,v}^{(t)} \triangleq \mathbb{P}[Y_t = v | Y_0 = y]$ stand in for the probability that the process $\{Y_t\}_{t\geq 0}$ is at state v at time t starting from y, we have by the Markov property [8] that $Q_{y,v}^{(t)} = \big(\prod_{s=1}^t Q(s)\big)_{y,v}$ for $y, v \in \mathcal{G}_n$. Note that the Markov chain modeled here is inhomogeneous, which is different from Sect. 3.1. In particular, the time-dependent transition matrix $Q(t)$ here cannot be expressed in terms of a power P^t of a transition matrix P on the same state space as in Sect. 3.1.

Results. Now define the sets of (δ, t)-bad gate pairs by $\mathcal{B}_{|\psi_0\rangle,\delta}^{|\Psi_0\rangle,t} \triangleq \big\{x \in \mathcal{U}_n \big| \|\rho_t - x\sigma_0 x^\dagger\|_d > \delta\big\}$ for $|\psi_0\rangle, |\Psi_0\rangle \in \mathcal{S}^n$, $t \in \{0, 1, \ldots, \tau\}$, $\delta \geq 0$. Also define the matrices $B_{|\psi_0\rangle,\delta}^{|\Psi_0\rangle,t} \in [0,1]^{|\mathcal{G}_n|\times|\mathcal{G}_n|}$ element-wise by

$$\big(B_{|\psi_0\rangle,\delta}^{|\Psi_0\rangle,t}\big)_{y,v} \triangleq \begin{cases} Q_{y,v}(t) & \text{if} \quad v \notin \mathcal{B}_{|\psi_0\rangle,\delta}^{|\Psi_0\rangle,t}, \\ 0 & \text{otherwise,} \end{cases}$$

for $t = 0, 1, \ldots, \tau$. Recall the notation introduced above Proposition 1. Similarly enumerate in the vector d_{ρ_t} the Schatten d-norms between any of the possibles states of σ_t and the faultless state ρ_t. We state our second result; see [32] for a proof:

Proposition 2 (Error accumulation in nonrandom circuits). *For any $y_0 \in \mathcal{G}_n$, $\delta \geq 0$, $t = 0, 1, \ldots, \tau < \infty$: the expected error is given by $\mathbb{E}_{y_0}[D_t] = e_{y_0}^{\mathrm{T}} \big(\prod_{k=1}^t Q(k)\big) d_{\rho_t}$. Similarly, the probability of error is given by*

$$\mathbb{P}_{y_0}[D_t > \delta] = e_{y_0}^{\mathrm{T}} \Big(\prod_{k=1}^{t} Q(k)\Big) 1_{\mathcal{B}_{|\psi_0\rangle,\delta}^{|\Psi_0\rangle,t}}. \tag{8}$$

Furthermore; if $y_0 \notin \mathcal{B}^{|\Psi_0\rangle,0}_{|\psi_0\rangle,\delta}$*, the probability of maximum error is given by*

$$\mathbb{P}_{y_0}[\max_{0\leq s\leq t} D_s > \delta] = \sum_{s=0}^{t-1}\Big(e_{y_0}^{\mathrm{T}}\big(\prod_{r=0}^{s} B^{|\Psi_0\rangle,r}_{|\psi_0\rangle,\delta}\big)\times\big(Q(s{+}1){-}B^{|\Psi_0\rangle,s+1}_{|\psi_0\rangle,\delta}\big)1_{\mathcal{B}^{|\Psi_0\rangle,s+1}_{|\psi_0\rangle,\delta}}\Big), \quad (9)$$

and otherwise it equals one.

For illustration, we have written a script that will generate a valid P and Q matrices after a user inputs a vector describing (gate-dependent) error probabilities. The code is available on TU/e's GitLab server at https://gitlab.tue.nl/20061069/markov-chains-for-error-accumulation-in-quantum-circuits.

3.2 Continuous, Random Error Accumulation (One-Qubit Case)

In this section, we analyze the case where a single qubit:

1. receives a random perturbation on the Bloch sphere after each s-th unitary gate according to a continuous distribution $p_s(\alpha)$, and
2. depolarizes to the completely depolarized state $I/2$ with probability $q \in [0, 1]$ after each unitary gate,

by considering it an absorbing random walk on the Bloch sphere. The key point leveraged here is that the trace distance is invariant under rotations. Hence a rotationally symmetric perturbation distribution will still allow us to calculate the error probabilities.

Model. Let R_0 be an initial point on the Bloch sphere. Every time a unitary quantum gate is applied, the qubit is rotated and receives a small perturbation. This results in a random walk $\{R_t\}_{t\geq 0}$ on the Bloch sphere for as long as the qubit has not depolarized. Because the trace distance is invariant under rotations and since the rotations are applied both to ρ_t and σ_t, we can ignore the rotations. We let ν denote the random time at which the qubit depolarizes. With the usual independence assumptions, $\nu \sim \text{Geometric}(q)$.

Define $\mu_t(r)$ for $t < \nu$ as the probability that the random walk is in a solid angle Ω about r (in spherical coordinates) conditional on the qubit not having depolarized yet. That is,

$$\mathbb{P}[R_t \in \mathcal{S} | \nu > t] \triangleq \int_{\mathcal{S}} \mu_t(r)\mathrm{d}\Omega(r).$$

We assume without loss of generality that $R_0 = \hat{z}$. From [40], the initial distribution is then given by

$$\mu_0 = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\cos\theta).$$

Here, the $P_n(\cdot)$ denote the Legendre polynomials. Also introduce the shorthand notation

$$\Lambda_{n,t} \triangleq \prod_{s=1}^{t} \int_0^{\pi} P_n(\cos\alpha) \mathrm{d}p_s(\alpha)$$

for convenience. As we will see in Proposition 3 in a moment, these constants will turn out to be the coefficients of an expansion for the expected trace distance (see (10)). Recall that here, $p_s(\alpha)$ denotes the probability measure of the angular distance for the random walk on the Bloch sphere at time t (see (i) above). In particular: if $p_t(\alpha) = \delta(\alpha)$ for all $t \geq 0$ meaning that each step is taken into a random direction but exactly of angular length α, then $\Lambda_{n,t} = (P_n(\cos\alpha))^t$. From [40], it follows that after t unitary quantum gates have been applied without depolarization having occurred,

$$\mu_t = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \Lambda_{n,t} P_n(\cos\theta).$$

Results. In this section we specify D_t as the trace distance. We are now in position to state our findings; proofs can be found in [32]:

Proposition 3 (Single qubit). *For $0 \leq \delta \leq 1$, $t \in \mathbb{N}_+$: the expected trace distance satisfies*

$$\mathbb{E}[D_t] = \tfrac{1}{2} - (1-q)^t \Big(\tfrac{1}{2} + 2 \sum_{n=0}^{\infty} \frac{\Lambda_{n,t}}{(2n-1)(2n+3)} \Big). \tag{10}$$

The probability of the trace distance deviating is given by

$$\mathbb{P}[D_t \leq \delta] = \mathbb{1}[\tfrac{1}{2} \in [0,\delta]] \big(1 - (1-q)^t\big) + (1-q)^t \sum_{n=0}^{\infty} (2n+1) \Lambda_{n,t} \sum_{r=1}^{n+1} (-1)^{r+1} \delta^{2r} C_{r-1} \binom{n+r-1}{2(r-1)}.$$

Here, the C_r denote the Catalan numbers. Finally; the probability of the maximum trace distance deviating is lower bounded by

$$\mathbb{P}[\max_{0 \leq s \leq t} D_s \leq \delta | \nu > t] \geq 0 \vee \Big(1 - t + \delta^2 \sum_{s=1}^{t} \sum_{n=0}^{\infty} (2n+1) \Lambda_{n,s} \frac{n!}{(2)_n} P_n^{(1,-1)}(1-2\delta^2) \Big).$$

4 Simulations

We will now briefly illustrate and validate our results numerically. For a more indepth numerical investigation, see [32].

Consider two nonrandom circuits: the first is a periodical single-qubit circuit that repeats a Hadamard, Pauli-X, Pauli-Y and Pauli-Z gate $k = 25$ times, and

the second a two-qubit circuit that is repeated $k = 5$ times; see also Fig. 3. Here the controlled-NOT gate $\mathrm{CNOT} = \big((1,0,0,0);(0,1,0,0);(0,0,0,1);(0,0,1,0)\big)$. Consider also the following two error models in which the errors depend on the gates:
(i) For the single-qubit circuit, presume $\mathbb{P}[\Lambda = I] = 0.990, \mathbb{P}[\Lambda = Z] = 0.010$.
(ii) For the two-qubit circuit, when labeling the qubits by A and B, suppose

$$\mathbb{P}[\Lambda_A = I] = 0.990,\ \mathbb{P}[\Lambda_A = X] = 0.006, \mathbb{P}[\Lambda_A = Y] = 0.003,\ \mathbb{P}[\Lambda_A = Z] = 0.001;$$
$$\mathbb{P}[\Lambda_B = I] = 0.980,\ \mathbb{P}[\Lambda_B = X] = 0.002, \mathbb{P}[\Lambda_B = Y] = 0.014,\ \mathbb{P}[\Lambda_B = Z] = 0.004.$$

In order to evaluate Proposition 2, we set the error threshold $\delta = 1/10$.

The theoretical and simulation results on the two circuits are shown in Fig. 3. Note that the simulation curves almost coincide with the theoretical curves; the deviation is only due to numerical limits. Furthermore, because different gates influence error accumulation to different degrees, the periodical ladder shape

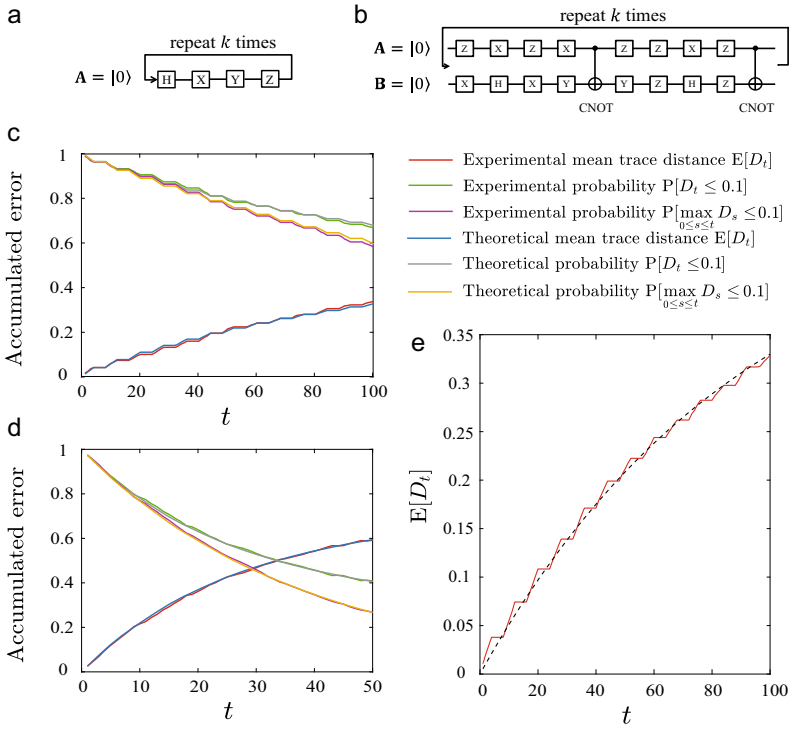

Fig. 3. Theoretical and simulation results for error accumulation on a single-qubit circuit (figures a, c, and e) and a two-qubit circuit (figures b and d). The numerical results are calculated from 2000 independent runs, and almost indistinguishable from the formulae. The dashed, black curve in figure e is a fit of $D_t \approx \frac{1}{2}(1-(1-\mu)^t)$ to the data—this formula describes the so-called *depolarization channel*, a basic model of depolarization of one qubit. The fit parameter is $\mu^{\mathrm{fit}} \approx 0.011$.

occurs in Fig. 3. Observe furthermore that this periodical ladder shape is not captured by a fit method that only takes into account the decay of t applications of a single depolarizing channel.

5 Minimizing Errors in Quantum Circuit Through Optimization

The rate at which errors accumulate may be different for different quantum circuits that can implement the same algorithm. Using techniques from optimization and (9), we can therefore search for the quantum circuit that has the lowest error rate accumulation while maintaining the same final state. To see this, suppose we are given a circuit $\mathcal{U}_\tau = \{U_1, U_2, \ldots, U_\tau\}$. For given ρ_0 this brings the quantum state to some quantum state ρ_τ. Other circuits may go to the same final state and have a lower probability of error at time τ. We will therefore aim to

$$\begin{aligned} &\underset{G_1,\ldots,G_\tau \in \mathcal{G}_n}{\text{minimize}} && u(\{G_1, \ldots, G_\tau\}) \\ &\text{subject to} && G_\tau \cdots G_1 = U_\tau \cdots U_1. \end{aligned} \tag{11}$$

Here, one can for example choose for the objective function $u(\cdot)$ the probability of error (8), or probability of maximum error (9). To solve (11), we design a simulated annealing algorithm in Sect. 5.1 to improve the quantum circuit.

The minimization problem in (11) is well-defined and has a few attractive features. For starters, the minimization problem automatically detects shorter circuits if the probability of error when applying the identity operator $I^{\otimes n}$ is relatively small. The optimum may then for example occur at a circuit of the form

$$G_\tau G_{\tau-1} G_{\tau-2} \cdots G_2 G_1 = I^{\otimes n} G_{\tau-1} I^{\otimes n} \cdots I^{\otimes n} G_1,$$

which effectively means that only the two gates $G_{\tau-1}G_1$ are applied consecutively. The identity operators in this solution essentially describe the passing of time. Now, critically, note that while the minimization problem does consider all shorter circuits of depth at most τ, this does not necessarily mean that the physical application of one specific group element $G \in \mathcal{G}_n$ is always the best. Concretely, in spite of the fact that any quantum circuit of the form $G_\tau \cdots G_1 = G \in \mathcal{G}_n$ performs the single group element $G \in \mathcal{G}_n$, it is not necessarily true that

$$u(\{G, I^{\otimes n}, \ldots, I^{\otimes n}\}) < u(\{G_1, \ldots, G_\tau\}).$$

The reason for this is that the error distribution on the direct group element G may be worse than using a circuit utilizing multiple other group elements. In other words, the optimal circuit need not always be the 'direct' circuit, but of course it can be. (In Sect. 5.2 we also consider the situation in which an experimentalist can only apply a subset $\mathcal{A} \subseteq \mathcal{G}_n$ that need not necessarily be a group, and in such a case the direct group element G may not even be a viable solution to the experimentalist if $G \notin \mathcal{A}$.) Typically, the minimization

problem will prefer shorter circuits if the probability of error when applying the identity operator $I^{\otimes n}$ is relatively small and the error distributions of all gate distributions are relatively homogeneous.

5.1 Simulated Annealing

We will generate candidate circuits as follows. Let $\{G_1^{[\eta]}, \dots, G_\tau^{[\eta]}\}$ denote the circuit at iteration η. Choose an index $I \in [\tau - 1]$ uniformly at random, choose $G \in \mathcal{G}$ uniformly at random. Then set

$$G_i^{[\eta+1]} = \begin{cases} G & \text{if } i = I, \\ G_{I+1}^{[\eta]} G_I^{[\eta]} G^{\leftarrow} & \text{if } i = I+1, \\ G_i^{[\eta]} & \text{otherwise.} \end{cases}$$

Here, $G^{\leftarrow}$ denotes the (left) inverse group element, i.e., $G^{\leftarrow} G = I^{\otimes n}$. The construction thus ensures that

$$G_{I+1}^{[\eta+1]} G_I^{[\eta+1]} = \left(G_{I+1}^{[\eta]} G_I^{[\eta]} G^{\leftarrow}\right) G = G_{I+1}^{[\eta]} G_I^{[\eta]}$$

so that the circuit's intent does not change: $G_\tau^{[\eta+1]} \cdots G_1^{[\eta+1]} = G_\tau^{[\eta]} \cdots G^{[\eta]}$.

We will use the Metropolis algorithm. Let

$$E = \left\{ \{G_1, \dots, G_\tau\} | G_\tau \cdots G_1 = U_\tau \cdots U_1 \right\}$$

denote the set of all viable circuits. For two arbitrary circuits $i, j \in E$, let

$$\Delta(i,j) \triangleq \sum_{s=1}^{\tau-1} \mathbb{1}[i_s \neq j_s, i_{s+1} \neq j_{s+1}]$$

denote the number of consecutive gates that differ between both circuits. Under this construction, the *candidate-generator matrix* of the Metropolis algorithm is given by

$$q_{ij} = \begin{cases} \frac{1}{(\tau-1)|\mathcal{G}|} & \text{if } \Delta(i,j) \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Since the candidate-generator matrix is symmetric, this algorithm means that we set $\alpha_{i,j}(T) = \exp\left(-\frac{1}{T} \max\{0, u(j) - u(i)\}\right)$ as the *acceptance probability* of circuit j over i. Here $T \in (0, \infty)$ is a positive constant. Finally, we need a cooling schedule. Let $M \triangleq \sup_{\{i,j \in E | \Delta(i,j) \leq 1\}} \{u(j) - u(i)\}$. Based on [8], if we choose a cooling schedule $\{T_\eta\}_{\eta \geq 0}$ that satisfies $T_\eta \geq \frac{\tau M}{\ln \eta}$, then the Metropolis algorithm will converge to the set of global minima of the minimization problem in (11).

Lemma 2. *Algorithm 1 converges to the global minimizer of* (11) *whenever* $T_\eta \geq \tau M / \ln \eta$ *for* $\eta = 1, 2, \cdots$.

Input: A group $\mathcal{G}$, a circuit $\{U_1, \ldots, U_\tau\}$, and number of iterations w
Output: A revised circuit $\{G_1^{[w]}, \ldots, G_\tau^{[w]}\}$
begin
  Initialize $\{G_1^{[0]}, \ldots, G_\tau^{[0]}\} = \{U_1, \ldots, U_\tau\}$;
  for $\eta \leftarrow 1$ **to** w **do**
    Choose $I \in [\tau - 1]$ uniformly at random;
    Choose $G \in \mathcal{G}$ uniformly at random;
    Set $J_I = G, J_{I+1} = G_{I+1}^{[\eta]} G_I^{[\eta]} G^{\leftarrow}, J_i = G_i^{[\eta]} \, \forall_{i \neq I, I+1}$;
    Choose $X \in [0, 1]$ uniformly at random;
    if $X \leq \alpha_{G^{[\eta]}, J}(T_\eta)$ **then**
      Set $G^{[\eta+1]} = J$;
    else
      Set $G^{[\eta+1]} = G^{[\eta]}$;
    end
  end
end

Algorithm 1: Pseudo-code for the simulated annealing algorithm described in Section 5.1.

5.2 Examples

Gate-Dependent Error Model. We are going to improve the one-qubit circuit in Fig. 3 using Algorithm 1. The gates are limited to the Clifford group $\mathcal{C}_1$ and the errors will be limited to the Pauli channel. The error probabilities considered here are gate-dependent and written out explicitly in [32, Appendix H]. The cooling schedule used here will be set as $T_\eta = C/\ln(\eta + 1)$, and the algorithm's result when using $C = 0.004$ is shown in Fig. 4a. Figure 4a illustrates that the improved circuit can indeed lower the error accumulation rate. The circuit with the lowest error accumulation rate that was found is shown in [32, Appendix H].

Gates in a Subset of One Group. The gates that are available in practice may be restricted to some subset $\mathcal{A} \subseteq \mathcal{G}$ not necessarily a group. Under such constraint, we could generate candidate circuits as follows: Let $\{G_1^{[\eta]}, \ldots, G_\tau^{[\eta]}\}$ denote the circuit at iteration η. In each iteration, two neighboring gates will be considered to be replaced by two other neighboring gates. There are $m \leq (\tau - 1)$ neighboring gate pairs $(G_1^{[\eta]}, G_2^{[\eta]}), \ldots, (G_{m-1}^{[\eta]}, G_m^{[\eta]})$ that can be replaced by two different neighboring gates. Choose an index $I \in [m - 1]$ uniformly at random, and replace $(G_I^{[\eta]}, G_{I+1}^{[\eta]})$ by any gate pair from $\{(\tilde{G}_1, \tilde{G}_2) \in \mathcal{A}^2 \mid G_I^{[\eta]} G_{I+1}^{[\eta]} = \tilde{G}_1 \tilde{G}_2\}$ uniformly at random. Pseudo-code for this modified algorithm can be found in [32, Algorithm 2]. It must be noted that this algorithm is not guaranteed to converge to the global minimizer of (11) (due to limiting the gates available); however, it may still find use in practical scenarios where one only has access to a restricted set of gates.

We now aim to decrease the probability of maximum error (9) by changing the two-qubit circuit shown in Fig. 3. The error model is the same as that in Sect. 4–B. The set of gates available for improving the circuit is here limited to $\{I, X, Y, Z, H, CNOT\}$. The result here for the two-qubit circuit is obtained by again using the cooling schedule $T_\eta = C/\ln(\eta + 1)$ but now letting the parameter $C = 0.002$. Figure 4b shows that a more error-tolerant circuit can indeed be found using this simulated annealing algorithm. The improved circuit is shown in [32, Appendix H].

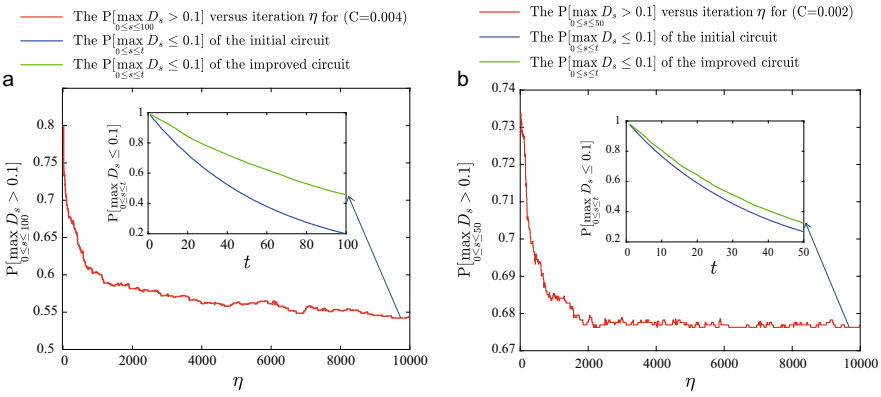

Fig. 4. a) Circuit optimization when using Algorithm 1. The error probabilities are gate-dependent. Note that the probability of maximum error (9) decreases as the number of iterations η increases when using Algorithm 1 ($C = 0.004$). b) Circuit optimization when the available gates are limited. The set of gates available is chosen limited to $\{I, X, Y, Z, H, CNOT\}$. Here we started from the two-qubit circuit shown in Fig. 3.

Deutsch–Jozsa Algorithm. Let us give further proof of concept through the Deutsch–Jozsa Algorithm for one classical bit [11,13]. This quantum algorithm determines if a function $f : \{0,1\} \to \{0,1\}$ is constant or balanced, i.e., if $f(0) = f(1)$ or $f(0) \neq f(1)$. It is typically implemented using the quantum circuit in Fig. 5. If no errors occur in this quantum circuit, then the first qubit would measure $|0\rangle$ or $|1\rangle$ w.p. one if f constant or balanced, respectively. If errors occur in this quantum circuit, then there is a strictly positive probability that the first qubit measures $|1\rangle$ or $|0\rangle$ in spite of f being constant or balanced, respectively, and thus for the algorithm to incorrectly output that f is constant or balanced. This *misclassification probability* ν of the algorithm depends on the underlying error distributions, and can be calculated by adapting (8)'s derivation.

We suppose now that errors occur according to a distribution in which two-qubit Clifford gates are more error prone than single-qubit gates. We can then revise the quantum circuit in Fig. 5 using a simulated annealing algorithm that aims at minimizing (11) by randomly swapping out poor gate pairs for better gate pairs. This simulated annealing algorithm, like any other, is sensitive to the choice of *cooling schedule* [8], here set as $T_\eta = C\big(\gamma/\eta + (1-\gamma)/\ln(\eta+1)\big)$

with $C > 0$, $\gamma \in [0, 1]$; the integer η indexes the iterations. Figure 5 shows the ratio $\Theta \triangleq \nu_{\text{original circuit}}/\nu_{\text{revised circuit}}$ as a function of C, γ for $f_a(x) = x$, $f_b(x) = 1 - x$, $f_c(x) = 0$, $f_d(x) = 1$ where $x \in \{0, 1\}$. Note that $\Theta \geq 1$ always, ≥ 1.60 commonly, and sometimes even ≥ 2.20.

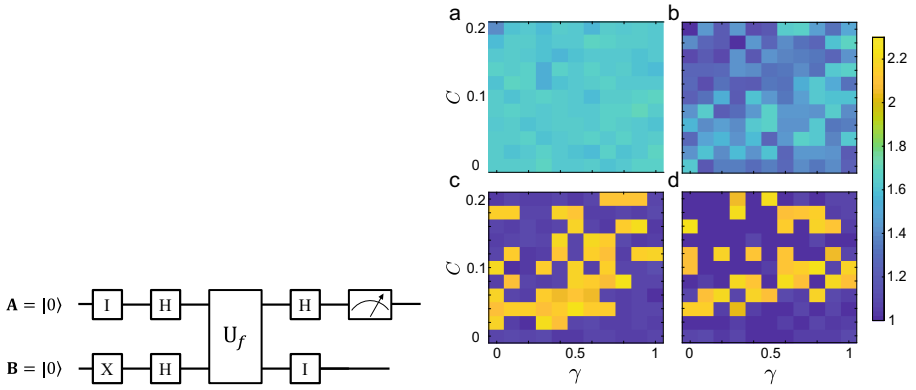

Fig. 5. (left) The Deutsch–Jozsa Algorithm for one classical bit in quantum circuit form. (right) Relative improvement when using Algorithm 1. For every pair (C, γ) here, Θ was calculated using a Monte Carlo simulation with 10^5 independent repetitions for the best circuit found throughout $w = 10^3$ iterations of the annealing algorithm. $u(\cdot)$ was set to the misclassification probability for a, c; and to (9) for b, d.

6 Conclusion

In conclusion; we have proposed and studied a model for discrete Markovian error accumulation in a multi-qubit quantum computation, as well as a model describing continuous errors accumulating in a single qubit. By modeling the quantum computation with and without errors as two coupled Markov chains, we were able to capture a weak form of time-dependency, allow for fairly generic error distributions, and describe multi-qubit systems. Furthermore, by using techniques from discrete probability theory, we could calculate the probability that error measures such as the fidelity and trace distance exceed a threshold analytically. To combat the numerical challenge that may occur when evaluating our expressions, we additionally provided an analytical bound on the error probabilities that is of lower numerical complexity. Finally, we showed how our expressions can be used to decide how many gates one can apply before too many errors accumulate with high probability, and how one can lower the rate of error accumulation in existing circuits by using techniques from optimization.

Acknowledgments. We are grateful to Bart van Schooten, who contributed the code on TU/e's GitLab server. Finally, this research received financial support from the Chinese Scholarship Council (CSC) in the form of a CSC Scholarship.

References

1. Aaronson, S., Gottesman, D.: Improved simulation of stabilizer circuits. Phys. Rev. A **70**(5), 052328 (2004)
2. Amy, M.: Formal methods in quantum circuit design (2019)
3. Ball, H., Stace, T.M., Flammia, S.T., Biercuk, M.J.: Effect of noise correlations on randomized benchmarking. Phys. Rev. A **93**(2), 022303 (2016)
4. Bennett, C.H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A., Wootters, W.K.: Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Phys. Rev. Lett. **70**(13), 1895 (1993)
5. Bennett, C.H., Wiesner, S.J.: Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. Phys. Rev. Lett. **69**(20), 2881 (1992)
6. Bhatia, R.: Matrix Analysis, vol. 169. Springer Science & Business Media (2013). https://doi.org/10.1007/978-1-4612-0653-8
7. Bravyi, S., Englbrecht, M., König, R., Peard, N.: Correcting coherent errors with surface codes. npj Quantum Inf. **4**(1), 55 (2018)
8. Brémaud, P.: Discrete Probability Models and Methods, vol. 78. Springer (2017). https://doi.org/10.1007/978-3-319-43476-6
9. Brown, W.G., Eastin, B.: Randomized benchmarking with restricted gate sets. Phys. Rev. A **97**(6), 062323 (2018)
10. Carignan-Dugas, A., Boone, K., Wallman, J.J., Emerson, J.: From randomized benchmarking experiments to gate-set circuit fidelity: how to interpret randomized benchmarking decay parameters. New J. Phys. **20**(9), 092001 (2018)
11. Cleve, R., Ekert, A., Macchiavello, C., Mosca, M.: Quantum algorithms revisited. Proc. Royal Soc. London. Ser. A: Math. Phys. Eng. Sci. **454**(1969), 339–354 (1998)
12. Cramer, J., et al.: Repeated quantum error correction on a continuously encoded qubit by real-time feedback. Nat. Commun. **7**, 11526 (2016)
13. Deutsch, D., Jozsa, R.: Rapid solution of problems by quantum computation. Proc. Royal Soc. London. Ser. A: Math. Phys. Sci. **439**(1907), 553–558 (1992)
14. Epstein, J.M., Cross, A.W., Magesan, E., Gambetta, J.M.: Investigating the limits of randomized benchmarking protocols. Phys. Rev. A **89**(6), 062321 (2014)
15. Fong, B.H., Merkel, S.T.: Randomized benchmarking, correlated noise, and ising models. arXiv preprint arXiv:1703.09747 (2017)
16. Fowler, A.G., Hollenberg, L.C.: Scalability of Shor's algorithm with a limited set of rotation gates. Phys. Rev. A **70**(3), 032329 (2004)
17. França, D.S., Hashagen, A.: Approximate randomized benchmarking for finite groups. J. Phys. A: Math. Theoret. **51**(39), 395302 (2018)
18. Fujii, K.: Stabilizer formalism and its applications. In: Quantum Computation with Topological Codes, pp. 24–55. Springer (2015). https://doi.org/10.1007/978-981-287-996-7
19. Gottesman, D.: The Heisenberg representation of quantum computers. arXiv preprint quant-ph/9807006 (1998)
20. Gottesman, D.: Efficient fault tolerance. Nature **540**, 44 (2016)
21. Greenbaum, D., Dutton, Z.: Modeling coherent errors in quantum error correction. Quantum Sci. Technol. **3**(1), 015007 (2017)
22. Greenberger, D.M., Horne, M.A., Zeilinger, A.: Going beyond Bell's theorem. In: Bell's Theorem, Quantum Theory and Conceptions of the Universe, pp. 69–72. Springer (1989). https://doi.org/10.1007/978-94-017-0849-4_10
23. Gutiérrez, M., Svec, L., Vargo, A., Brown, K.R.: Approximation of realistic errors by Clifford channels and Pauli measurements. Phys. Rev. A **87**(3), 030302 (2013)

24. Gutmann, H.: Description and control of decoherence in quantum bit systems. Ph.D. thesis, lmu (2005)
25. Harper, R., Hincks, I., Ferrie, C., Flammia, S.T., Wallman, J.J.: Statistical analysis of randomized benchmarking. Phys. Rev. A **99**(5), 052350 (2019)
26. Huang, E., Doherty, A.C., Flammia, S.: Performance of quantum error correction with coherent errors. Phys. Rev. A **99**(2), 022313 (2019)
27. Janardan, S., Tomita, Yu., Gutiérrez, M., Brown, K.R.: Analytical error analysis of Clifford gates by the fault-path tracer method. Quantum Inf. Process. **15**(8), 3065–3079 (2016). https://doi.org/10.1007/s11128-016-1330-z
28. Kliuchnikov, V., Maslov, D.: Optimization of Clifford circuits. Phys. Rev. A **88**(5), 052307 (2013)
29. Knill, E.: Quantum computing with realistically noisy devices. Nature **434**(7029), 39 (2005)
30. Koenig, R., Smolin, J.A.: How to efficiently select an arbitrary Clifford group element. J. Math. Phys. **55**(12), 122202 (2014)
31. Linke, N.M., et al.: Fault-tolerant quantum error detection. Sci. Adv. **3**(10), e1701074 (2017)
32. Ma, L., Sanders, J.: Markov chains and hitting times for error accumulation in quantum circuits. arXiv preprint arXiv:1909.04432 (2021)
33. Magesan, E., Gambetta, J.M., Emerson, J.: Scalable and robust randomized benchmarking of quantum processes. Phys. Rev. Lett. **106**(18), 180504 (2011)
34. Magesan, E., Puzzuoli, D., Granade, C.E., Cory, D.G.: Modeling quantum noise for efficient testing of fault-tolerant circuits. Phys. Rev. A **87**(1), 012324 (2013)
35. Maslov, D., Dueck, G.W., Miller, D.M., Negrevergne, C.: Quantum circuit simplification and level compaction. IEEE Trans. Comput.-Aided Des. Integrated Circ. Syst. **27**(3), 436–444 (2008)
36. Moll, N., et al.: Quantum optimization using variational algorithms on near-term quantum devices. Quantum Sci. Technol. **3**(3), 030503 (2018)
37. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information: 10th Anniversary Edition, 10th edn. Cambridge University Press, New York (2011)
38. Preskill, J.: Quantum computing: pro and con. Proc. Royal Soc. London. Ser. A: Math. Phys. Eng. Sci. **454**(1969), 469–486 (1998)
39. Proctor, T., Rudinger, K., Young, K., Sarovar, M., Blume-Kohout, R.: What randomized benchmarking actually measures. Phys. Rev. Lett. **119**(13), 130502 (2017)
40. Roberts, P.H., Ursell, H.D.: Random walk on a sphere and on a Riemannian manifold. Philos. Trans. Royal Soc. London. Ser. A, Math. Phys. Sci. **252**(1012), 317–356 (1960)
41. Ruskai, M.B.: Pauli exchange errors in quantum computation. Phys. Rev. Lett. **85**(1), 194 (2000)
42. Selinger, P.: Generators and relations for n-qubit Clifford operators. Logical Meth. Comput. Sci. **11** (2013)
43. Wallman, J.J.: Randomized benchmarking with gate-dependent noise. Quantum **2**, 47 (2018). https://doi.org/10.22331/q-2018-01-29-47
44. Wallman, J.J., Barnhill, M., Emerson, J.: Robust characterization of loss rates. Phys. Rev. Lett. **115**(6), 060501 (2015)
45. Wood, C.J., Gambetta, J.M.: Quantification and characterization of leakage errors. Phys. Rev. A **97**(3), 032306 (2018)
46. Xia, T., et al.: Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits. Phys. Rev. Lett. **114**(10), 100503 (2015)