Optimal Trotterization in universal quantum simulators under faulty control

George C. Knee[*] and William J. Munro

NTT Basic Research Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan
(Received 10 March 2015; published 26 May 2015)

Universal quantum simulation may provide insights into those many-body systems that cannot be described classically and that cannot be efficiently simulated with current technology. The Trotter formula, which decomposes a desired unitary time evolution of the simulator into a stroboscopic sequence of repeated elementary evolutions, is a key algorithmic component which makes quantum simulation of dynamics tractable. The Trotter number n sets the time scale on which a computer running this algorithm is switched from one elementary evolution to another. In the ideal case, the precision of the simulation can be arbitrarily controlled by increasing n. We study a more realistic scenario where each gate is applied imperfectly. The resultant trade-off in errors leads to an ultimate limit on the precision of the simulation. We calculate the optimum Trotter number n^* that achieves this limit, which is the minimum statistical distance from the actual simulation to the ideal one.

I. INTRODUCTION

Simulation involves the control and study of one physical system (the simulator) in order to provide insight into a second (the simuland). It is an indispensable tool in modern science. Under a classical computing paradigm, a huge variety of simulands are investigated using silicon-based digital machines, which represent and process classical information stored as binary data in memory. For a broad class of problems, the simulation is *efficient*: the simulator requires spatial and temporal resources scaling polynomially rather than exponentially in the size of the simuland. When the simuland is such that a fully quantum mechanical model is necessary to describe it, however, the memory requirement for a direct and exact simulation scales exponentially in the number of particles. Examples include quantum chemistry simulations [1] and simulations of condensed-matter systems [2–4]. Often a mean-field approximation or cutoff in long-range entanglement can reduce the complexity and make the problem tractable once more.

The idea of *quantum* simulation, on the other hand, where the computer can store and process *quantum* information, offers the potential to explore quantum simulands without making such assumptions. Quantum simulators are typically thought of as being analog or digital [5]. Analog quantum simulators are purpose built and have limited reprogrammability. They often process information stored in continuous variables and therefore do not benefit from error correction. The concept is similar to the idea of an orrery or a scale model.

Digital or universal quantum simulators, however, are anticipated to operate under the quantum computation paradigm [6–9]. Although they may, in fact, have more modest requirements, for our purposes they will be thought of as fully fledged quantum computers. The advantage of this type of simulation is that a single computer can be reprogrammed to perform any simulation, as long as it has enough memory. Perhaps more importantly, the error in the outcome of the calculation can be controlled by error-correction methods.

Quantum computers are typically constructed with two-level systems known as "quantum bits." Such elements may be realized with trapped ions [3], superconducting circuits [2,10], or other candidate systems [11]. The formal requirements for quantum computation [12] dictate that a fiducial initial state can be prepared, that a universal set [13–15] of operations can be applied, and that projective measurement is possible.

Universal quantum simulation, being an instance of quantum computation, is therefore composed of preparation, evolution, and readout stages. Analyses concerning state preparation and data extraction can be found elsewhere [16,17]. The aim of this paper is to quantify the accuracy of the evolution stage when the control operations are faulty. To that end, in Sec. II we discuss the algorithms used for quantum simulation and recap arguments concerning accuracy with ideal operations; in Sec. III we discuss a general framework for modeling faulty operations. In Sec. IV we discuss various ways to quantify the accuracy of a quantum computation. Our main results are found in Sec. V, where we argue that a trade-off between Trotter error and gate errors is generic to universal quantum simulation and exhibit this with several examples. Our results show how to operate a faulty simulation in the optimal way and the level of precision expected at this optimum.

II. SIMULATION ALGORITHM

Even when a quantum computer with a universal gate set is available, one does not generally know how to combine the gates *efficiently* to achieve a particular desired evolution.

A. Lloyd's algorithm

Lloyd's algorithm [6] is a general but approximate solution to this problem when the desired evolution is known to be generated by a local Hamiltonian:

$$H = \sum_{j=1}^{k} H_j, \qquad (1)$$

where each of the k component Hamiltonians H_j has dimension less than some maximum (call this g). The true evolution is then approximated through a truncation of the Trotter [18]

_____
[*]gk@physics.org

formula:

$$U = \exp[iHt] = \left(\prod_{j=1}^{k} \exp[iH_j t/n] \right)^n + \cdots, \quad (2)$$

where we set $\hbar = 1$ and take H to be time independent (although this can be straightforwardly relaxed). Further, we rescale our units such that time t is dimensionless.

The advantage of Lloyd's approach is that the number of operations required is bounded by a number proportional to $t^2 k g^2/\epsilon$, where t is the simulation time, g is the maximum dimension of the local Hamiltonians, and ϵ is the desired error [6]. The total number k of component Hamiltonians has a better-than-polynomial (rather than exponential) dependence on the particle number, making Lloyd's method "efficient."

It is clear that for any finite value of n (the "Trotter number"), the higher-order terms in the above equation will be nonzero. Their neglect then leads to an error in the simulation: the simulator is not driven to the desired final state but to one that is nearby. As we shall see, "nearby" can be given a concrete mathematical and operational definition.

Higher Trotter numbers result in trajectories of the simulator that result in a final state that is closer to the ideal. Increasing n will generally increase the computational complexity, that is, more control operations. More time may not be required, however, because the gates are correspondingly shorter, although this appears not to be the case in error corrected implementations [19]. Others have considered the dependence of the number of gates and simulation time on the desired error [19], but here our only concern is the overall accuracy of the simulation. Higher-order approximants are available [20,21], along with other techniques that exploit sparsity [22]. The common attributes of these approaches are that they reduce the number of operations to polynomial in the particle number and that their accuracy is controlled (improved) by increasing the number of applied operations. The number of Trotter steps necessary for quantum chemistry simulation was considered in Refs. [23,24].

Throughout this paper we take $V_j = \exp[iH_j t/n]$ as primitive operations applied to the simulator, although, ultimately, these primitives should be understood as being composed from gates drawn from the particular universal set that is available.

B. Time-energy freedom

Because of the $H \to aH, t \to t/a$ symmetry of the Schrödinger equation $i \frac{\partial}{\partial t}|\psi\rangle = -H|\psi\rangle$, when simulating closed-system dynamics, one benefits from the freedom to define

$$H_{\text{simulator}} = a H_{\text{simuland}},$$
$$t_{\text{simulator}} = \frac{1}{a} t_{\text{simuland}}, \quad (3)$$

which will preserve the correspondence between the time evolutions of the simulator and simuland. The simulation time, therefore, may be chosen to be any duration as long as the appropriate global scaling of the energy of the control fields is also performed. Even classical simulators are rarely operated at a speed commensurate with their respective simulands: weather patterns of several weeks are simulated in a matter of hours, and supercomputers spend months calculating chemical

reaction dynamics over time scales many orders of magnitude smaller. In fact, because each unitary $\exp[iH_j t/n]$ is decomposed into the natural elementary gate set of the simulator, one generally expects $a \neq 1$. Clearly, this freedom enables the computation to be sped up (or slowed down) by a constant factor [25], but more importantly, we imagine that such freedom may prove very useful for reducing noise in implementations of quantum simulation, depending on the particular noise which dominates (see below). It is worth noting that assuming this freedom is asking more than is necessary for universal quantum computation, nevertheless, in many quantum computers we expect there to be at least a limited ability to perform the primitive operations at different physical speeds.

III. NOISE TYPES

The error in approximating the true evolution with the first term on the right-hand side of (2) is known to decay at worst as $\propto n^{-1}$ [26], and as Lloyd [6] put it, "n can always be picked sufficiently large to ensure the simulator tracks the correct time evolution to within any [desired nonzero accuracy]." The perfect control of any system (quantum or classical) is only ever an idealization, however. Little is known about the real-world situation, although recently the Trotter decomposition has been shown to be stable in the sense that the overall error can be reduced to a fixed desired level if the precision of the individual steps is good enough [27]. Reference [28] is a study of the influence of noise on certain quantum simulations, calculating the average fidelity of the final state of the computer with the ideal. Here we derive analytical results for arbitrary (generally nonunitary and noncommuting) component Hamiltonians. Further, as we shall show, our results allow us to predict the ultimate performance of a faulty quantum simulator and allow us to prescribe the optimum Trotter number to employ. Our results hold for a whole class of statistical distance metrics, including the most interesting *worst-case* metrics.

We construct "faulty Trotterized quantum channels" and consider the following generalized noise map, with $\mathcal{V}_j(\rho) = e^{iH_j t/n} \rho e^{-iH_j t/n}$ representing the component unitary processes of Trotterization (here describing the transformation of a $d \times d$ density matrix ρ describing the quantum state of the simulator):

$$\mathcal{E}^{\text{faultyTrotter}}(\rho) = \bigcirc_{i=1}^{n} \bigcirc_{j=1}^{k} \mathcal{E}_{ij} \circ \mathcal{V}_j(\rho)$$
$$= \bigcirc_{i=1}^{n} (\cdots \circ \mathcal{E}_{i2} \circ \mathcal{V}_2 \circ \mathcal{E}_{i1} \circ \mathcal{V}_1)(\rho)$$
$$= \bigcirc_{i=1}^{n} \mathcal{E}_i^{\text{fTss}}(\rho). \quad (4)$$

The $\circ$ symbol is used here to denote the concatenation of quantum channels: here we have a triple concatenation (first due to faulty evolutions $\mathcal{E}_{ij}$ following the clean ones, second due to applying each of the k local Hamiltonians in turn, and third by repeating this process n times). The interleaved operations $\mathcal{E}_{ij}$ are unwanted evolutions; they carry an index i to emphasize that they may vary over the course of the experiment, and there may also be a dependence on $\{H_j\}$, t, and n. In the course of our derivations it is useful to consider the map over a single Trotter iteration (faulty Trotter single step) $\mathcal{E}_i^{\text{fTss}}$.

It will be convenient to consider the supermatrix representation of these maps: this is defined by $\mathbf{T}_\mathcal{E} \vec{\rho} = \overrightarrow{\mathcal{E}(\rho)}$,

where $\vec{\rho}$ is the vectorized (column-stacked) density matrix and $\mathbf{T}_{\mathcal{E}}$ is a supermatrix of dimension d^2. This representation makes the semigroup structure of quantum maps apparent. The concatenation of channels is merely matrix multiplication:

$$
\begin{aligned}
\mathbf{T}^{\text{faultyTrotter}} &= \prod_{i=1}^{n} \prod_{j=1}^{k} \mathbf{T}_{\mathcal{E}_{ij}} \mathbf{T}_{\mathcal{V}_j} \\
&= \prod_{i=1}^{n} \left(\cdots \mathbf{T}_{\mathcal{E}_{i2}} \mathbf{T}_{\mathcal{V}_2} \mathbf{T}_{\mathcal{E}_{i1}} \mathbf{T}_{\mathcal{V}_1} \right) \\
&= \prod_{i=1}^{n} \mathbf{T}_i^{\text{fTss}}.
\end{aligned} \tag{5}
$$

We write the perfect implementation of the Trotter technique as $\mathcal{V}$ with unitary matrix $V = (\sqrt[n]{V})^n$ and supermatrix $\mathbf{T}_{\mathcal{V}} = (\sqrt[n]{\mathbf{T}_{\mathcal{V}}})^n$, essentially by removing all the faulty maps from the faulty Trotter channel.

We write the supermatrix representation of the ideal map $\mathcal{U}$ as $\mathbf{U} = U \otimes U^*$ and break up this unitary evolution as

$$
\mathcal{U}(\rho) = \bigcirc^n \sqrt[n]{\mathcal{U}}(\rho), \tag{6}
$$

with the notation justified by the supermatrix describing each ideal Trotter step $\mathbf{U} = (\sqrt[n]{\mathbf{U}})^n$.

IV. STATISTICAL DISTANCE MEASURES FOR QUANTUM CHANNELS

Hauke *et al.* [29] pose the question, Can we trust quantum simulators? Rather than give a binary answer, one can instead ask the question, To what extent can we trust quantum simulators? We aim to provide an answer to the above question in the form of a discrimination probability. To measure the faithfulness of quantum operations, one can appeal to generalizations of either classical fidelity or classical statistical distance between probability distributions. We prefer the latter here: the *trace distance* between quantum *states* (density matrices) has an operational meaning because it determines the success probability for the state discrimination problem [30]. One is provided with either ρ_A or ρ_B with equal probability and is asked to guess which after a single measurement. The probability of success is

$$
p_{\text{distinguish}} = \tfrac{1}{2} + \tfrac{1}{4} ||\rho_A - \rho_B||. \tag{7}
$$

Here

$$
||\rho_A - \rho_B|| = \text{Tr}|\rho_A - \rho_B| = \text{Tr}\sqrt{(\rho_A - \rho_B)(\rho_A - \rho_B)^\dagger}. \tag{8}
$$

The norm of the difference between the operators is a metric, giving the quantum statistical distance between the operators. This definition implies that the discrimination is informed by measurement results arising from the optimum choice of positive operator-valued measure (POVM), or generalized measurement procedure [31]. In fact,

$$
||\rho_A - \rho_B|| = 2 \max_{E} \text{Tr}[E(\rho_A - \rho_B)], \tag{9}
$$

where E is a positive operator (or POVM element). A similar maximization over states in turn induces a distance on quantum *channels*, defined as

$$
||\mathcal{E}_A - \mathcal{E}_B||_{\diamond,1} = \max_{\rho \in S, S_{\text{sep}}} ||(\mathcal{E}_A \otimes \mathbb{I})(\rho) - (\mathcal{E}_B \otimes \mathbb{I})(\rho)||. \tag{10}
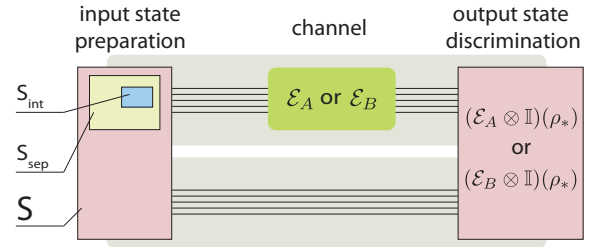$$

FIG. 1. (Color online) The distinguishability of two quantum channels $\mathcal{E}_A$ and $\mathcal{E}_B$ can be defined as the distinguishability of the outputs of each channel when the input is an optimally chosen quantum state ρ_*. The optimum is generally drawn from a set S describing joint states of an enlarged system. The subset S_{sep}, contains states with no system-ancilla entanglement and a further subset S_{int}, which contains only physical states of interest.

The associated task is to submit an optimal input state ρ_* to undergo evolution under $\mathcal{E}_A$ or $\mathcal{E}_B$ (chosen at random) and to perform the state discrimination task (above) on the output state. For full generality the enlarged search space S (having dimensions d^2) is needed to allow for the possibility of entanglement with an ancilla (of dimension no larger than that defined by the channels themselves) assisting in the channel discrimination task [32]. When the full search space S is available, $|| \cdot ||_{\diamond}$ is the diamond norm [32]; otherwise, when the maximization is over states which factorize into system, ancilla states, the norm is the unstabilized induced trace norm $|| \cdot ||_1$ (and the ancilla plays no role). A further restriction considered by Lloyd [6] would be to define the set of states of "interest" S_{int}, but this has the drawback of requiring detailed knowledge of the particular simulation at hand and makes additional assumptions on the set of initial states of interest (see Fig. 1).

These metrics then give the bias away from a half in the probability of discriminating between the real and ideal channels, given a single optimal initial state preparation (for the diamond norm over a larger space), a single channel use, and a single sample from an optimally chosen measurement basis (again for the diamond norm on a larger space):

$$
p_{\text{distinguish}} \leqslant \tfrac{1}{2} + \tfrac{1}{4} ||\mathcal{E}_A - \mathcal{E}_B||_{\diamond,1}. \tag{11}
$$

They are therefore worst-case metrics. Yet another norm to consider is the J norm $|| \cdot ||_J$: this provides a bound on the average trace distance over a uniformly measured state space [33]. The J distance between two quantum maps is merely the trace distance of the associated states in the Jamiolkowski isomorphism [34]:

$$
||\mathcal{E}_A - \mathcal{E}_B||_J = ||J(\mathcal{E}_A) - J(\mathcal{E}_B)||, \tag{12}
$$

where $J(\mathcal{E}) \propto \sum_{ij} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j|$ with the constant of proportionality such that $J(\mathcal{E})$ has unit trace. The J distance is related to the average probability of discriminating the real and ideal simulations, given an optimal measurement:

$$
\bar{p}_{\text{distinguish}} \leqslant \tfrac{1}{2} + \tfrac{1}{4} ||\mathcal{E}_A - \mathcal{E}_B||_J. \tag{13}
$$

For a comprehensive survey of available metrics, see Refs. [33,35].

There are a number of senses in which these metrics are pessimistic measures. First, for the worst-case metrics, an optimal input state (possibly entangled with an ancilla) may not arise in a quantum simulator: readout methods such as the phase-estimation algorithm [36], for example, do not employ such states. Currently, the full set of preparations and measurements at the disposal of a quantum computer is not used in quantum simulation algorithms: in the future, more involved procedures (including readout using an ancilla) may turn out to be more powerful at extracting information. Second, for the maximization over POVMs, in reality, the measurements are likely to be restricted to those observables necessary for calculating macroscopic quantities, e.g., the ground-state energy of the system, the Shannon entropy, effective temperature, net magnetization, and so on. Consider that these measurement choices might be completely impervious to certain types of error, but the errors are ruthlessly sought out and exposed under the trace-norm or diamond-norm distances. In other words, there may be an equivalence class of (possibly mutually orthogonal) microstates with the same macrostate (or property of interest). An "incorrect" microstate could well be good enough for the purposes of the simulation (even yielding precisely the correct "answer") while being highly (or even perfectly) distinguishable from the correct state. Consider, for example, a very large simulator suffering only a bit flip on a qubit corresponding to the least significant digit of the readout register (minimal effect on readout result) or even a simulator which suffers a severe scrambling of phase relations prior to measurement in the computational basis (potentially zero effect on readout result). The game of simulation is not necessarily an adversarial one; nevertheless, due to the often random and potentially significant effect of errors, worst-case metrics are usually regarded as the most relevant measures to employ.

All metrics satisfy some important properties which we will make use of: the triangle inequality $||\mathcal{E}_A - \mathcal{E}_C|| \leqslant ||\mathcal{E}_A - \mathcal{E}_B|| + ||\mathcal{E}_B - \mathcal{E}_C||$ and convexity $||\sum_i p_i \mathcal{E}_i|| \leqslant \sum_i p_i ||\mathcal{E}_i||$. Another important property satisfied by all metrics considered in this paper is chaining: $||\mathcal{E}_{A_1} \circ \mathcal{E}_{A_2} - \mathcal{E}_{B_1} \circ \mathcal{E}_{B_2}|| \leqslant ||\mathcal{E}_{A_1} - \mathcal{E}_{B_1}|| + ||\mathcal{E}_{A_2} - \mathcal{E}_{B_2}||$. The diamond norm and J distance are special in that they are *stable*: they satisfy $||\mathbb{I} \otimes \mathcal{E}||_{\diamond, J} = ||\mathcal{E}||_{\diamond, J}$. See Gilchrist *et al.* [33] for further details.

The convexity property is an important one and is enough to ensure that the maximum statistical distance is achieved on a pure state [33]. Further, although often understood as the inability to increase distinguishability through averaging (i.e., as a handicap in information processing tasks such as parameter estimation [31]), here it implies that the performance of a faulty quantum computation may be *improved* by repeating the computation and averaging the results. As we shall see below, this can lead to quite significant improvements because the fluctuations may be suppressed.

A related fact concerns appropriate worst-case benchmarks. While the ideal evolution $\mathcal{U}$ is unitary and will take a pure state to a pure state, a completely noisy channel $\mathcal{E}^{\rho \to \mathbb{I}/d}(\rho) = \mathbb{I}/d$ destroys the purity of any state (having dimension d). One can prove [37] that under the unstabilized trace norm one has

$$||\mathcal{E}^{\rho \to \mathbb{I}/d} - \mathcal{U}||_1 = 2 - \frac{2}{d}, \quad (14)$$

whereas with the diamond norm [38] and J distance (which we concentrate on for the remainder of this article) one has

$$||\mathcal{E}^{\rho \to \mathbb{I}/d} - \mathcal{U}||_{\diamond, J} = 2 - \frac{2}{d^2}. \quad (15)$$

These values represent benchmarks: the distinguishability between the output of the ideal simulation and complete noise. Any simulation giving a higher distance than this will be worse than a completely random output! Note that both norms converge to the algebraic maximum of 2 as $d \to \infty$. The algebraic maximum is achieved, for example, by an unwanted bit flip.

V. RESULTS

We are now in a position to calculate the accuracy of a realistic quantum simulator. To that end, we apply the norms introduced above to calculate the statistical distance from the ideal map to the faulty Trotter map. To see why we expect an optimum Trotter number to emerge in realistic universal quantum simulators, consider first a simple application of the triangle inequality to the distance between ideal and realistic channels over a single Trotter step:

$$||\sqrt[n]{\mathcal{U}} - \mathcal{E}^{\text{fTss}}|| \leqslant ||\sqrt[n]{\mathcal{U}} - \sqrt[n]{\mathcal{V}}|| + ||\sqrt[n]{\mathcal{V}} - \mathcal{E}^{\text{fTss}}||. \quad (16)$$

Next, assume the first term is bounded by a quantity $\propto n^{-s}$ for some $s \geqslant 2$: this captures the first-order Trotter formula that we study here ($s = 2$), as well as higher-order expansions [20]. Now assume the second term is a constant with respect to n. This immediately leads, via the chaining property, to

$$D := ||\mathcal{U} - \mathcal{E}^{\text{faultyTrotter}}|| \leqslant \frac{\mathscr{C}}{n^{s-1}} + \mathscr{D}n \quad (17)$$

for some constants $\mathscr{C}$ and $\mathscr{D}$. It is simple to show that the optimum Trotter number is

$$n^* = \sqrt[s]{\frac{\mathscr{C}(s-1)}{\mathscr{D}}} \quad (18)$$

and the statistical distance at this optimum is

$$D(n^*) = \frac{\mathscr{D}s}{s-1} \sqrt[s]{\frac{\mathscr{C}(s-1)}{\mathscr{D}}}. \quad (19)$$

Depending on the values of $\mathscr{C}$ and $\mathscr{D}$ (which also depend on s in the general case), one is able to also optimize over s to find the best-order Trotter formula to use. For the remainder of this paper, however, we set $s = 2$, and these quantities reduce to

$$n^* = \sqrt{\frac{\mathscr{C}}{\mathscr{D}}} \quad (20)$$

and

$$D(n^*) = 2\sqrt{\mathscr{C}\mathscr{D}}. \quad (21)$$

As we shall show, $\mathscr{C}$ is fixed by the pair $\{H, t\}$, while $\mathscr{D}$ is determined by gate errors. Below we investigate some example noise models that influence $\mathscr{D}$.

A. Mistimed control

Trotterization fundamentally requires "switching" between unitary gates: in our simplification this is thought of as switching each component Hamiltonian H_j on for a specific duration. Consider that the duration is increased (decreased)

by a random number Δ_{ij}, normally distributed around zero with variance σ^2. This makes

$$\mathcal{E}_{ij}^{\mathrm{MTC}}(\rho) = e^{iH_j\Delta_{ij}}\rho e^{-iH_j\Delta_{ij}}. \tag{22}$$

Such imperfection is ubiquitous in the control of quantum systems, and an equivalent imperfection (laser intensity fluctuations) was cited as the dominant error source by Lanyon *et al.* in a recent implementation of universal quantum simulation [3]. Under this noise model, the total map $\mathcal{E}^{\mathrm{faultyTrotter}}$, being a concatenation of random unitaries, is itself a random unitary map. Note that the magnitude of the errant time shift Δ_{ij} is independent of t or n, but the error map is more severe when $||H_i||$ is large, i.e., the simulator is driven "hard."

As shown in Appendix A, by assuming $||H_i[t/n + \Delta_{ij}]|| \ll 1$ one can find the difference between real and ideal supermatrices for this error model, finding

$$\sqrt[n]{\mathbf{U}} - \mathbf{T}_i^{\mathrm{fTss}}$$

$$= \sum_{j<l}(\mathbb{I}\otimes[H_j,H_l]^* + [H_j,H_l]\otimes\mathbb{I})\frac{t^2}{2n^2}$$

$$+ i\sum_j(\mathbb{I}\otimes H_j^* - H_j\otimes\mathbb{I})\Delta_{ij}$$

$$+ \sum_{j<l}(\mathbb{I}\otimes H_j^* H_l^* + H_j H_l\otimes\mathbb{I})$$

$$\times\left[\frac{t}{n}\Delta_{ij} + \frac{t}{n}\Delta_{il} + \Delta_{ij}\Delta_{il}\right]$$

$$+ \frac{1}{2}\sum_j(\mathbb{I}\otimes H_j^{2*} + H_j^2\otimes\mathbb{I})\left[2\frac{t}{n}\Delta_{ij} + \Delta_{ij}^2\right]$$

$$- \sum_{jl}H_j\otimes H_l^*\left[\frac{t}{n}\Delta_{ij} + \frac{t}{n}\Delta_{il} + \Delta_{ij}\Delta_{il}\right] + \cdots \tag{23}$$

Note that the terms linear in Δ work to introduce "coherent" or reversible errors that may be averaged away because of our assumptions $\overline{\Delta_{ij}} = 0$ and $\overline{\Delta_{ik}\Delta_{im}} = \delta_{km}\sigma^2$; hence

$$\overline{\sqrt[n]{\mathbf{U}} - \mathbf{T}_i^{\mathrm{fTss}}} = \sum_{j<l}(\mathbb{I}\otimes[H_j,H_l]^* + [H_j,H_l]\otimes\mathbb{I})\frac{t^2}{2n^2}$$

$$+ \frac{1}{2}\sum_j(\mathbb{I}\otimes H_j^{2*} + H_j^2\otimes\mathbb{I})\sigma^2$$

$$- \sum_{jl}H_j\otimes H_l^*\sigma^2 + \cdots \tag{24}$$

leaving the "incoherent" or irreversible errors (quadratic in Δ) remaining. Clearly, averaging can help reduce fluctuations introduced by the noise. This is expected, because of the convexity property of the norm. The resultant (less severe) map is denoted $\overline{\mathrm{MTC}}$, 'for averaged mistimed control'. Figure 2 shows (through numerical simulations) how averaging can reduce the simulation error for this noise model, as measured by the diamond-norm distance. Taking the norm of (24) gives the statistical distance, and employing the triangle inequality gives

$$D_{\mathrm{MTC}}^{ss} \lessapprox \mathscr{A}\frac{t^2}{2n^2} + \mathscr{B}\sigma^2. \tag{25}$$
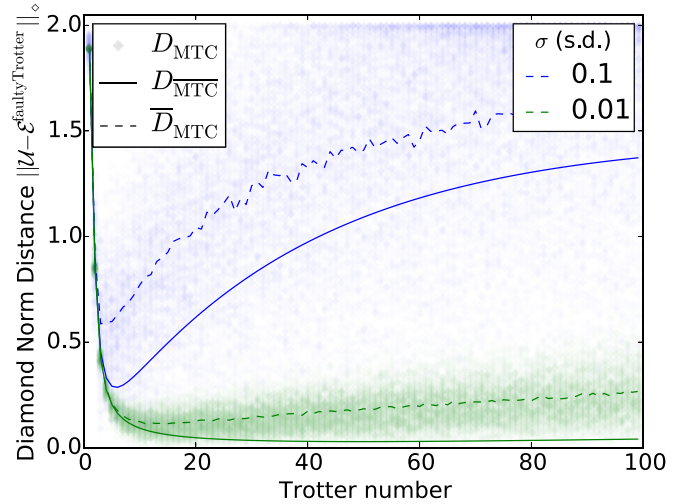

FIG. 2. (Color online) Diamond norm distance vs Trotter number for a quantum simulation of the Hamiltonian $H = \sigma_x + \sigma_y$ over a duration $t = 2$. Each transparent marker represents a single Monte Carlo simulated random unitary evolution of the simulator. The solid lines represent the performance of the averaged map, and the dashed lines are the average of the individual Monte Carlo runs: lower is better. Note how the position of the averaged map is below the average position of the markers. The performance of the average map is better than the average performance of the maps of which it is composed. The two colors correspond to different noise characteristics of the mistimed control noise model.

By the chaining property we reach

$$D_{\overline{\mathrm{MTC}}} \lessapprox \mathscr{A}\frac{t^2}{2n} + \mathscr{B}n\sigma^2, \tag{26}$$

with

$$\mathscr{A} = \left\|\sum_{j<l}\left(\mathbb{I}\otimes[H_j,H_l]^* + [H_j,H_l]\otimes\mathbb{I}\right)\right\|_?$$

$$\mathscr{B} = \left\|\frac{1}{2}\sum_j\left(\mathbb{I}\otimes H_j^{*2} + H_j^2\otimes\mathbb{I} - 2H_j\otimes H_j^*\right)\right\|_?. \tag{27}$$

Comparing (26) with (17), we see that $\mathscr{A}t^2/2 = \mathscr{C}$ and $\mathscr{B}\sigma^2 = \mathscr{D}$. Note that we leave the choice of norm free here: $? \in \{\diamond, 1, J, \ldots\}$. We made use of only the triangle inequality and the chaining property. As we will show, if $\mathscr{A} : \mathscr{B}$ happens to be invariant under the choice of norm, the optimum Trotter number is also an invariant since it only depends on this ratio. Note also that these are norms of supermatrices: the supermatrices may need to be converted to another form so that the norms are evaluated.

Applying the triangle inequality to $\mathscr{A}$ and $\mathscr{B}$, we note that the first quantity is upper bounded by the sum of, at worst, $\frac{1}{2}(k^2 - k)$ norms and the second by the sum of only k norms. Since k is polynomial in the particle number, the Trotter error (sometimes called the digital error) therefore grows only polynomially in the size of the simuland.

Taking $n \in \mathbb{R}$ for the moment, simple analysis yields (i) the optimum Trotter number

$$n_{\mathrm{MTC}}^* = \sqrt{\mathscr{A}(2\mathscr{B})^{-1}}t\sigma^{-1}, \tag{28}$$

(ii) the statistical distance at this optimum

$$D_{\overline{\text{MTC}}}(n^*) = \sqrt{2\mathscr{A}\mathscr{B}}t\sigma, \qquad (29)$$

and (iii) the maximum simulation time after which the statistical distance is above the accepted level $D_{\max}$, given by

$$t_{\max} = D_{\max}/(\sigma\sqrt{2\mathscr{A}\mathscr{B}}). \qquad (30)$$

The meaning of equations (28), (29), and (30) is inherited from the choice of norm in the definition of $\mathscr{A}$ and $\mathscr{B}$.

Of course, in reality $n^* \in \mathbb{Z}$: the true optimum must be a whole number of Trotter steps, so n^* should be rounded in the direction of the sign of $\lfloor n^*\rfloor\lceil n^*\rceil - \mathscr{A}t^2/(2\mathscr{B}\sigma^2)$. See Appendix C for more details.

Although our formulas apply generally, as a case study we study a special case of the Ising Hamiltonian,

$$H_1 = \sum_r^N \sigma_z^r, \quad H_2 = \sum_{\langle r,s\rangle} \sigma_x^r \sigma_x^s \qquad (31)$$

with σ_x^r and σ_z^r being the Pauli matrices acting on qubit r of N, with identity matrices implied on other qubits. The notation $\sum_{\langle r,s\rangle}$ denotes a sum over nearest neighbors. Note that $[H_1,H_2] \neq 0$. As shown in Fig. 3(a), our approximate analytical upper bound (26) is a good fit for exact numerics in the correct parameter regime (for example, we chose $t = 0.1$ and $N = 2$).

We note three regimes where the fit is worse: (i) when $||H_i t/n|| \approx 1$ (the perturbative expansion breaks down), (ii) in the trade-off region near the optimum Trotter number n^* (here the chaining inequality is loose), and (iii) when the distance measure approaches its maximum (the errors are saturating, and the chaining inequality between consecutive Trotter steps is loose). Despite the looseness of the fit in these areas, the location of n^* and $D(n^*)$ are well captured.

Aside from combatting the gate noise via averaging, one can apply the time-energy freedom (3), noting that

$$a \neq 1 \Rightarrow D_{\overline{\text{MTC}}} \lessapprox \mathscr{A}\frac{t^2}{2n} + \mathscr{B}n(a\sigma)^2. \qquad (32)$$

Clearly, one can now take $a < 1$, i.e., retard the simulator to decrease the absolute error at each Trotter step. This has the same effect as reducing σ: the first term is unaffected.

B. Trotter-step-induced depolarization

Next we study a different noise model, which we call the Trotter-step-induced depolarization, or DEPOL model. Depolarizing noise is commonly used to phenomenologically model noise because it makes analytic results tractable. To gain an idea of how each operation might introduce generic noise into the simulator, we assume that after a single Trotter step the combined effect of imperfections in the applied gates is to depolarize the entire simulator uniformly. The faulty map is then

$$\mathcal{E}^{\text{faultyTrotter}}(\rho) = \bigcirc_i \mathcal{E}_i^{\text{DEPOL}} \circ (\cdots \mathcal{V}_2 \circ \mathcal{V}_1)(\rho) \qquad (33)$$

$$= \bigcirc_i \mathcal{E}_i^{\text{fTss}}, \qquad (34)$$

with

$$\mathcal{E}_i^{\text{DEPOL}}(\rho) = (1-p)\rho + \frac{p}{d}\mathbb{I}. \qquad (35)$$

Here $\mathbb{I}$ is the $d \times d$ identity matrix. Note there is no dependence on any variable other than p, the probability of a depolarizing error during one Trotter step. This quantity therefore captures the severity of the faulty control. Now

$$\mathcal{E}^{\text{fTss}} - \sqrt[n]{\mathcal{U}} = (1-p)\sqrt[n]{\mathcal{V}} + p\mathcal{E}^{\rho\rightarrow\mathbb{I}/d} - \sqrt[n]{\mathcal{U}}$$
$$= (1-p)[\sqrt[n]{\mathcal{V}} - \sqrt[n]{\mathcal{U}}] + p[\mathcal{E}^{\rho\rightarrow\mathbb{I}/d} - \sqrt[n]{\mathcal{U}}], \quad (36)$$

and we can use convexity to reach

$$||\mathcal{E}^{\text{fTss}} - \sqrt[n]{\mathcal{U}}|| \leqslant (1-p)||\sqrt[n]{\mathcal{V}} - \sqrt[n]{\mathcal{U}}|| + p||\mathcal{E}^{\rho\rightarrow\mathbb{I}/d} - \sqrt[n]{\mathcal{U}}||. \qquad (37)$$

Now, the first term follows trivially from the above calculation and equals $(1-p)D_{\overline{\text{MTC}}}^{\text{ss}}|_{\sigma=0}$. The second term is also known (see above). This means that (again employing the chaining inequality over the n iterations of the single step channel and assuming $||H_i||\,t/n \ll 1$)

$$D_{\text{DEPOL}} \lessapprox (1-p)\mathscr{A}\frac{t^2}{2n} + np\left(2 - \frac{2}{d^2}\right), \qquad (38)$$

which we can simplify, in the case of $d \gg 1$, to

$$D_{\text{DEPOL}} \lessapprox (1-p)\mathscr{A}\frac{t^2}{2n} + 2pn. \qquad (39)$$

Qualitatively, this is very similar to the mistimed control (MTC) model. One cannot, however, leverage the freedom in the time-energy correspondence to effectively reduce p, the characteristic of the noise, because (39) is invariant under that transformation. One has

$$n_{\text{DEPOL}}^* = t\sqrt{\frac{(1-p)\mathscr{A}}{2p\left(2 - \frac{2}{d^2}\right)}}, \qquad (40)$$

or, in the case of $d \gg 1$,

$$n_{\text{DEPOL}}^* = t\sqrt{\frac{(1-p)\mathscr{A}}{4p}}. \qquad (41)$$

Note that, as above, one should round n^* to the nearest integer in the direction of the sign of $\lfloor n^*\rfloor\lceil n^*\rceil - (1-p)\mathscr{A}t^2/4p$. One has

$$D_{\text{DEPOL}}(n^*) = 2t\sqrt{\mathscr{A}p(1-p)}. \qquad (42)$$

See Fig. 3(b) for comparison with numerical results, which show that the DEPOL noise model has behavior qualitatively similar to that of the $\overline{\text{MTC}}$ noise model and also that our approximate analytical approach continues to provide a good fit for the true performance.

C. Decoherence

To model the overall noise of a simulator under perfect control, we depolarize the simulator in a manner that is independent of the number of operations. This means the number of Trotter steps does not affect the total noise characteristics. An argument very similar to the one above gives

$$\mathcal{E}^{\text{faultyTrotter}} - \mathcal{U} = [1-p(t)][\mathcal{V} - \mathcal{U}] + p(t)[\mathcal{E}^{\rho\rightarrow\mathbb{I}/d} - \mathcal{U}], \qquad (43)$$

and now $p(t)$ increases over time. Typically, $1 - p(t)$ represents exponential decay of population and coherence
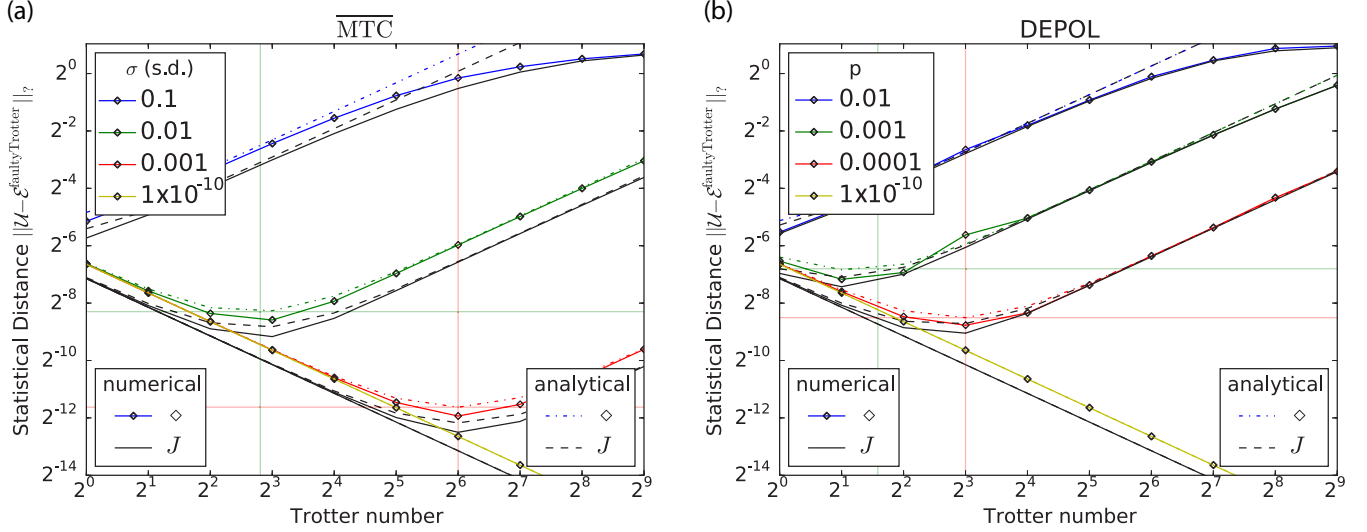
(a)

(b)

FIG. 3. (Color online) Log-Log plot of the statistical distance vs the Trotter number for a faulty simulation of the Ising Hamiltonian [Eq. (31)], with $N = 2$ and $t = 0.1$. The noise model is either (a) timing errors (the $\overline{\text{MTC}}$ noise model) or (b) depolarizing noise (the DEPOL model). The horizontal axis represents the frequency of applied operations, and the vertical axis is related to the probability that the actual simulation could be distinguished from the ideal simulation (lower is better). Solid colored lines (numerically calculated) correspond to the diamond norm distance, and the dotted lines are approximate analytical formulas. The different colors correspond to either (a) the standard deviation σ of the noise or (b) the probability p of a depolarizing error at each Trotter step, each shown in the legend. Black lines are the J distance (solid lines are exact numerics; dotted lines are approximate analytics). The crosshairs show the analytically predicted optimum Trotter number and performance level at that optimum. Note the $n \to \infty$ behavior shows the norms saturating to the completely noisy benchmark (see main text).

with time [31]. The first term follows trivially and is equal to $(1 - p)D_{\overline{\text{MTC}}}|_{\sigma=0}$. We have

$$D_{\text{DECOH}} \leqslant [1 - p(t)]\frac{\mathscr{A}t^2}{2n} + p(t)\left(2 - \frac{2}{d^2}\right). \quad (44)$$

Once more employing the time-energy freedom (3), we find

$$D_{\text{DECOH}} \leqslant \left[1 - p\left(\frac{t}{a}\right)\right]\frac{\mathscr{A}t^2}{2n} + p\left(\frac{t}{a}\right)\left(2 - \frac{2}{d^2}\right); \quad (45)$$

note that accelerating the simulator $a > 1$ will improve performance if $p(t)$ is monotonically increasing. Note here there is no optimum Trotter number because the noise is not worsened by increasing the number of operations.

VI. DISCUSSION

By modeling some simple imperfections in the control of a Trotterized universal quantum simulator, we have shown how the accuracy of the quantum channel that is applied depends on various parameters. These parameters include those outside of direct experimental control, such as the severity of environmental decoherence or the spread in control timing, and also controllable quantities such as the Trotter number and the simulator-simuland time ratio. We suggested ways in which these quantities can be optimized over in order to improve the accuracy of the simulation, which we quantified by calculating the statistical distance to the ideal map. The optimization implies only that the uncontrollable quantities be estimated from control experiments and that $\mathscr{C}$ (simply related to $\mathscr{A}$, which measures the total noncommutativity of the local pieces of the Hamiltonian) and $\mathscr{D}$ (which measures the control error per Trotter step) be calculated only once (possibly numerically).

We discussed the choice of a number of norms for this purpose, including the worst-case induced-trace norm and stabilized (diamond) norm, as well as the average-case J distance.

In particular we found there to be a finite optimum Trotter number to employ, setting a maximum performance and a maximum simulation time. By way of a general argument, we have shown these features to be generic to faulty simulators operated with a Trotterized (or similar) algorithm. In order to predict these quantities, one can appeal to a microscopic model of the simulator; having a sufficiently good model for a physical system is arguably a prerequisite for using it as a computational device [39].

The problem of noise in quantum computers has, in one sense, been solved by the error-correction threshold theorems [40–43]. An error-correction threshold is a critical value of a measure of the accuracy of quantum control. Once it has been surpassed, error-correction techniques work to decrease the overall net error. These theorems show that, when experimental operations become clean enough, encoding logical qubits in a larger number of redundant physical qubits allows the error in the overall computation to be suppressed at will [44]. But as interest in simulators grows [45] and small-scale (say, 64-qubit) devices begin to appear, the noise problem will remain until many-thousand-qubit devices with subthreshold physical error rates can be engineered. Even after this is achieved, imperfections will not have completely disappeared: the question of how the accuracy of overall computation depends on the necessarily finite *residual* (i.e., error-corrected) error rate [46] remains of high importance, and our analyses here will still apply.

We will briefly comment on two recent experimental demonstrations in superconducting systems. The Martinis

group performed an investigation of fermionic models with four transmon qubits [2], but they failed to control the error in the simulation at all, operating in the gate-error-dominated regime $n > n^*$. The Wallraff group, investigating interacting spin systems using only two transmon qubits, successfully swept the Trotter number through the optimum point [47] for various times t. By increasing the total simulation time they effectively increased the Trotter error until it competed with the other errors. Our results predict that, if both the Hamiltonian and error per Trotter step (i.e., $\mathscr{C}$ and $\mathscr{D}$) are kept constant, then n^*/t should be a constant: the experimental results are in agreement with this prediction. The worst trace distance in these experiments was of the order of 10^{-1}. These experiments show the difficulty in achieving good confidence in quantum simulators, especially as the number of qubits is increased. Using our results (20) and (21), we estimate that for the Wallraff experiment (fixing the Hamiltonian and simulation time) reaching the sorts of precision we are accustomed to with modern classical computers (the current standard "single precision" is $2^{-24} \approx 10^{-7}$ [48]) will require a factor of at least 10^6 more Trotter steps and therefore a 10^{12} improvement in the error per Trotter step. This is due to the fact that the overall error is proportional to the square root of the error per Trotter step.

Furthermore, as pointed out in Ref. [49], a full simulation using readout with the phase-estimation algorithm calls for simulation not only of U but of U^{2m}, further increasing the required number of Trotter steps and therefore precision per step. As Wecker *et al.* show [23], the number of logical gates per Trotter step (for example, in a quantum chemistry simulation) is roughly $O(N^5)$. This fact means that the error per gate may need to be several orders of magnitude smaller still.

That simulations are only likely to provide novel insight when featuring many dozens of qubits implies that yet another improvement in gate quality is necessary. Consider an $N = 64$ qubit simulation with nearest-neighbor interactions. Compared to the recent $N = 2$ experiment [47], there will be a thousandfold increase in the number of commutators between H_i; many of these may be expected to vanish, but nevertheless, we might expect $\mathscr{A}$ (and therefore $\mathscr{C}$) to increase by two or even three orders of magnitude, requiring an improvement in $\mathscr{D}$ of comparable magnitude to compensate.

Future work will investigate the multiparameter optimization when all noise types are present, the use of alternative Trotter-type approximants, and the performance of simulators when the set of interesting states and applied measurements is chosen from a restricted set.

APPENDIX A: ANALYTICS

We are interested in the difference between the faulty Trotter map described by the $\overline{\mathrm{MTC}}$ noise model and the ideal map. It is convenient here to choose the supermatrix representation. For a single shot map over a single Trotter step

$$\sqrt[n]{\mathbf{U}} - \mathbf{T}_i^{\mathrm{fTss}} = e^{i \sum_j H_j t/n} \otimes e^{-i \sum_j H_j^* t/n} - \prod_j e^{i H_j (t + n\Delta_{ij})/n} \otimes \prod_j e^{-i H_j^* (t + n\Delta_{ij})/n}$$

$$= \left(\mathbb{I} + \frac{it}{n} \sum_j H_j - \frac{t^2}{2n^2} \left\{ \sum_{j>l} + \sum_{j<l} + \sum_{j=l} \right\} H_j H_l + \cdots \right)$$

$$\otimes \left(\mathbb{I} - \frac{it}{n} \sum_j H_j^* - \frac{t^2}{2n^2} \left\{ \sum_{j>l} + \sum_{j<l} + \sum_{j=l} \right\} H_j^* H_l^* + \cdots \right)$$

$$- \left(\mathbb{I} + i \sum_j H_j \left[\frac{t}{n} + \Delta_{ij} \right] - \frac{1}{2} \left\{ \sum_{j=l} + 2 \sum_{j<l} \right\} H_j H_l \left[\frac{t}{n} + \Delta_{ij} \right] \left[\frac{t}{n} + \Delta_{il} \right] + \cdots \right)$$

$$\otimes \left(\mathbb{I} - i \sum_j H_j^* \left[\frac{t}{n} + \Delta_{ij} \right] - \frac{1}{2} \left\{ \sum_{j=l} + 2 \sum_{j<l} \right\} H_j^* H_l^* \left[\frac{t}{n} + \Delta_{ij} \right] \left[\frac{t}{n} + \Delta_{il} \right] + \cdots \right) \quad (A1)$$

$$= -\frac{t^2}{2n^2} \left\{ \sum_{j>l} + \sum_{j<l} + \sum_{j=l} \right\} \left(\mathbb{I} \otimes H_j^* H_l^* + H_j H_l \otimes \mathbb{I} \right) - i \sum_j \Delta_{ij} (H_j \otimes \mathbb{I} - \mathbb{I} \otimes H_j^*)$$

$$+ \frac{1}{2} \left\{ \sum_{j=l} + 2 \sum_{j<l} \right\} \left(\mathbb{I} \otimes H_j^* H_l^* + H_j H_l \otimes \mathbb{I} \right) \left[\frac{t^2}{n^2} + \frac{t}{n} \Delta_{ij} + \frac{t}{n} \Delta_{il} + \Delta_{ij} \Delta_{il} \right]$$

$$- \sum_{jl} H_j \otimes H_l^* \left[\frac{t}{n} \Delta_{ij} + \frac{t}{n} \Delta_{il} + \Delta_{ij} \Delta_{il} \right] + \cdots \quad (A2)$$

which, after employing the identity

$$-\left\{\sum_{j>l}+\sum_{j<l}+\sum_{j=l}\right\}+\left\{\sum_{j=l}+2\sum_{j<l}\right\}H_jH_l \equiv \left\{\sum_{j<l}-\sum_{j>l}\right\}H_jH_l \equiv \sum_{j<l}[H_j,H_l],$$

gives

$$\sqrt[n]{\mathbf{U}}-\mathbf{T}_i^{\text{fTss}}=\sum_{j<l}(\mathbb{I}\otimes[H_j,H_l]^*+[H_j,H_l]\otimes\mathbb{I})\frac{t^2}{2n^2}+i\sum_{j}(\mathbb{I}\otimes H_j^*-H_j\otimes\mathbb{I})\Delta_{ij}$$

$$+\sum_{j<l}(\mathbb{I}\otimes H_j^*H_l^*+H_jH_l\otimes\mathbb{I})\left[\frac{t}{n}\Delta_{ij}+\frac{t}{n}\Delta_{il}+\Delta_{ij}\Delta_{il}\right]$$

$$+\frac{1}{2}\sum_{j}(\mathbb{I}\otimes H_j^{2*}+H_j^2\otimes\mathbb{I})\left[2\frac{t}{n}\Delta_{ij}+\Delta_{ij}^2\right]$$

$$-\sum_{jl}H_j\otimes H_l^*\left[\frac{t}{n}\Delta_{ij}+\frac{t}{n}\Delta_{il}+\Delta_{ij}\Delta_{il}\right]+\cdots \tag{A3}$$

which is the result in the main text.

APPENDIX B: NUMERICS

The diamond-norm distance between two unitary maps described by matrices U and V is simply the diameter of the smallest enclosing circle of the eigenvalues of $UV^\dagger$ [50]. In our numerical work we used PYTHON code from Minase [51].

To calculate the diamond norm for a general map expressed in supermatrix, Choi, or Jamiolkowski form, we used a semidefinite programming [52,53] algorithm due to Watrous [54,55], implemented by Johnston in a MATLAB package called QETLAB [56].

1. Total noise and depolarization

What is the supermatrix form of $\mathcal{E}^{\rho\to\mathbb{I}/d}(\rho)=\mathbb{I}/d$? One can exploit the trace of the density matrix to get the following supermatrix:

$$\mathbf{T}^{\rho\to\mathbb{I}/d}=\begin{pmatrix}\frac{1}{d}&0&\ldots&0&\frac{1}{d}&0&\ldots\\0&0&\ldots&0&0&0&\ldots\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\ddots\\0&0&\ldots&0&0&0&\ldots\\\frac{1}{d}&0&\ldots&0&\frac{1}{d}&0&\ldots\\0&0&\ldots&0&0&0&\ldots\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\ddots\end{pmatrix}; \tag{B1}$$

each ellipsis denotes that there are d zero entries between the nonzero elements. Clearly, each off-diagonal is annihilated, and each diagonal goes to $(\text{Tr}\rho)/d$. From here we can easily construct a supermatrix representation of $\mathcal{E}^{\text{DEPOL}}$, for example, by scaling $\mathbf{T}^{\rho\to\mathbb{I}/d}$ by p and adding $(1-p)\mathbb{I}_{d^2\times d^2}$.

2. Averaged mistimed control

The quantum channel under consideration is what we call the *averaged* mistimed control ($\overline{\text{MTC}}$) model. It is the average

map that is applied when the timing suffers a nonzero spread.

$$\overline{\mathbf{T}}^{\text{faultyTrotter}}=\prod_i\overline{\mathbf{T}}_i^{\text{nTss}}=\prod_i\prod_j\overline{\mathbf{T}}_{\mathcal{E}_{ij}}\mathbf{T}_{\mathcal{U}_{ij}}; \tag{B2}$$

here we have exploited the independence of the random (matrix) variables to distribute the expectation over the concatenation of channels. Now for the $\overline{\text{MTC}}$ model the interleaved maps can be written as

$$\overline{\mathbf{T}}_{\mathcal{E}_{ij}}:\rho\to\int p(\Delta_{ij})e^{i\hat{H}_j\Delta_{ij}}\rho e^{-i\hat{H}_j\Delta_{ij}}d\Delta_{ij}, \tag{B3}$$

with the probability distribution $p(\Delta_{ij})$ taken as a product of independent Gaussians with zero mean and standard deviation of σ. Now, expanding the density matrix ρ in the energy eigenbasis, each operation merely generates a phase, and we can evaluate the integral (here we suppress the Hamiltonian index j for clarity):

$$\rho\to\frac{1}{\sqrt{2\pi}\sigma}\int e^{-\Delta_{ij}^2/\sigma^2}e^{i\hat{H}\Delta_{ij}}$$

$$\times\left(\sum_{mn}\tilde{\rho}_{mn}|E_m\rangle\langle E_n|\right)e^{-i\hat{H}\Delta_{ij}}d\Delta_{ij} \tag{B4}$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int e^{-\Delta_{ij}^2/\sigma^2}e^{iE_m\Delta_{ij}}e^{-iE_n\Delta_{ij}}$$

$$\times\left(\sum_{mn}\tilde{\rho}_{mn}|E_m\rangle\langle E_n|\right)d\Delta_{ij} \tag{B5}$$

$$=\sum_{mn}e^{-\frac{1}{2}(E_m-E_n)^2\sigma^2}\tilde{\rho}_{mn}|E_m\rangle\langle E_n|. \tag{B6}$$

We defined $\tilde{\rho}_{mn}=\langle E_m|\rho|E_n\rangle$ with $H|E_m\rangle=E_m|E_m\rangle$. Observe how the map acts on energy eigenstates: as an elementwise product $\tilde{\rho}\to\Lambda\cdot\tilde{\rho}$ with a matrix having entries $\Lambda_{mn}=e^{-\frac{1}{2}(E_m-E_n)^2\sigma^2}$. If we define W as the unitary matrix changing from the canonical into the energy eigenbasis, then the map acts on any ρ as $\rho\to W^\dagger(\Lambda\cdot W\rho W^\dagger)W$.

The supermatrix $\mathbf{T}_\Lambda$ corresponding to the elementwise product can be found by reshaping Λ into a vector and then

constructing a diagonal matrix from this vector. The advantage of using a supermatrix is the composition of channels is simply matrix multiplication. The map can be defined for any density matrix in the following way:

$$\overline{\mathbf{T}}_{\mathcal{E}_{ij}} \vec{\rho} = \mathbf{T}_{W^\dagger} \mathbf{T}_\Lambda \mathbf{T}_W \vec{\rho} \tag{B7}$$

and $\mathbf{T}_W = W \otimes W^*$ and so on.

By using these techniques, one is able to build up the total (averaged) faulty Trotter map in supermatrix form, which is convenient for numerical calculations.

APPENDIX C: FINDING n^* AS AN INTEGER

Using calculus, i.e., assuming n to be a real number, will generally give a noninteger solution for n^*. We can always round this up (or down) to get an integer, and in the worst case we may round in the wrong direction. Here we briefly show

how to make sure the integer chosen is the correct one. To ensure the lowest value for D, clearly, we should choose the "rounded-down" value $\lfloor n^* \rfloor$ and the "rounded-up" value $\lceil n^* \rceil$ when the quantity

$$D(\lceil n^* \rceil) - D(\lfloor n^* \rfloor) \tag{C1}$$

is positive and negative, respectively. When it is exactly zero, we can choose either. Let $\lfloor n^* \rfloor = k$ and let

$$D(k) = \frac{\mathscr{C}}{k} + \mathscr{D}k, \tag{C2}$$

and our condition simplifies to

$$k(k+1) - \frac{\mathscr{C}}{\mathscr{D}}. \tag{C3}$$

This justifies the claim in the main text that one should round n in the direction of the sign of $\lceil n^* \rceil \lfloor n^* \rfloor - (\mathscr{C}/\mathscr{D})$.

[1] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Simulating chemistry using quantum computers, Annu. Rev. Phys. Chem. **62**, 185 (2011).

[2] R. Barends *et al.*, Digital quantum simulation of fermionic models with a superconducting circuit, arXiv:1501.07703 (2015).

[3] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, Universal digital quantum simulation with trapped ions, Science **334**, 57 (2011).

[4] X. Ma, B. Dakic, W. Naylor, A. Zeilinger, and P. Walther, Quantum simulation of the wavefunction to probe frustrated Heisenberg spin systems, Nat. Phys. **7**, 399 (2011).

[5] I. Buluta and F. Nori, Quantum simulators, Science **326**, 108 (2009).

[6] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[7] R. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[8] K. L. Brown, W. J. Munro, and V. M. Kendon, Using quantum computers for quantum simulation, Entropy **12**, 2268 (2010).

[9] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, Nature (London) **464**, 45 (2010).

[10] U. Las Heras, A. Mezzacapo, L. Lamata, S. Filipp, A. Wallraff, and E. Solano, Digital quantum simulation of spin systems in superconducting circuits, Phys. Rev. Lett. **112**, 200501 (2014).

[11] I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Rev. Mod. Phys. **86**, 153 (2014).

[12] D. P. DiVincenzo, The physical implementation of quantum computation, Fortschr. Phys. **48**, 771 (2000).

[13] A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, Elementary gates for quantum computation, Phys. Rev. A **52**, 3457 (1995).

[14] D. Deutsch, A. Barenco, and A. Ekert, Universality in quantum computation, Proc. R. Soc. London, Ser. A **449**, 669 (1995).

[15] S. Lloyd, Almost any quantum logic gate is universal, Phys. Rev. Lett. **75**, 346 (1995).

[16] N. Zuniga-Hansen, Y.-C. Chi, and M. S. Byrd, Effects of noise, correlations, and errors in the preparation of initial states in quantum simulations, Phys. Rev. A **86**, 042335 (2012).

[17] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[18] H. F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. **10**, 545 (1959).

[19] K. R. Brown, R. J. Clark, and I. L. Chuang, Limitations of quantum simulation examined by simulating a pairing hamiltonian using nuclear magnetic resonance, Phys. Rev. Lett. **97**, 050504 (2006).

[20] W. Janke and T. Sauer, Properties of higher-order Trotter formulas, Phys. Lett. A **165**, 199 (1992).

[21] M. Suzuki, General theory of higher-order decomposition of exponential operators and symplectic integrators, Phys. Lett. A **165**, 387 (1992).

[22] D. Berry, G. Ahokas, R. Cleve, and B. Sanders, Efficient quantum algorithms for simulating sparse Hamiltonians, Commun. Math. Phys. **270**, 359 (2007).

[23] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Gate-count estimates for performing quantum chemistry on small quantum computers, Phys. Rev. A **90**, 022305 (2014).

[24] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doberty, and M. Troyer, The Trotter step size required for accurate quantum simulation of quantum chemistry, Quantum Inf. Comput. **15**, 361 (2015).

[25] A simulator cannot be operated in sublinear time; that is, it cannot take less than a time proportional to t_simuland [22].

[26] M. Reed and B. Simon, *Methods of Modern Mathematical Physics: Functional Analysis*, Methods of Modern Mathematical Physics Vol. 1 (Academic Press, San Diego, California, 1980).

[27] I. Dhand and B. C Sanders, Stability of the Trotter-Suzuki decomposition, J. Phys. A **47**, 265206 (2014).

[28] W. Dür, M. J. Bremner, and H. J. Briegel, Quantum simulation of interacting high-dimensional systems: The influence of noise, Phys. Rev. A **78**, 052325 (2008).

[29] P. Hauke, F. M. Cucchietti, L. Tagliacozzo, I. Deutsch, and M. Lewenstein, Can one trust quantum simulators?, Rep. Prog. Phys. **75**, 082401 (2012).

[30] C. A. Fuchs, Distinguishability and accessible information in quantum theory, Ph.D. thesis, University of New Mexico, 1995, arXiv:quant-ph/9601020.

[31] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, 2004).

[32] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics (American Mathematical Society, Providence, Rhode Island, 2002).

[33] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, Phys. Rev. A **71**, 062310 (2005).

[34] A. Jamiolkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Phys. **3**, 275 (1972).

[35] In [33] they refer to the diamond norm as the *S* distance.

[36] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, Science **309**, 1704 (2005).

[37] Due to the property of unitary invariance we can assume here that the unitary channel is the identity channel.

[38] M. Sacchi, Optimal discrimination of quantum operations, Phys. Rev. A **71**, 062340 (2005).

[39] C. Horsman, S. Stepney, R. C. Wagner, and V. Kendon, When does a physical system compute?, Proc. R. Soc. London, Ser. A **470**, 20140182 (2014).

[40] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, STOC '97* (ACM, New York, 1997), pp. 176–188.

[41] A. Yu. Kitaev, Quantum computations: algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[42] P. Aliferis, D. Gottesman, and J. Preskill, Quantum accuracy threshold for concatenated distance-3 codes, Quantum Inf. Comput. **6**, 97 (2006).

[43] E. Knill, R. Laflamme, and W. H. Zurek, Resilient quantum computation, Science **279**, 342 (1998).

[44] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[45] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Towards quantum chemistry on a quantum computer, Nat. Chem. **2**, 106 (2010).

[46] R. Raussendorf, Key ideas in quantum error correction, Philos. Trans. R. Soc. A **370**, 4541 (2012).

[47] Y. Salathé, M. Mondal, M. Oppliger, J. Heinsoo, P. Kurpiers, A. Potočnik, A. Mezzacapo, U. Las Heras, L. Lamata, E. Solano, S. Filipp, and A. Wallraff, Digital quantum simulation of spin models with circuit quantum electrodynamics, arXiv:1502.06778.

[48] M. L. Overton, *Numerical Computing with IEEE Floating Point Arithmetic* (Society for Industrial and Applied Mathematics, Philadelphia, 2001).

[49] C. R. Clark, T. S. Metodi, S. D. Gasster, and K. R. Brown, Resource requirements for fault-tolerant quantum simulation: The ground state of the transverse Ising model, Phys. Rev. A **79**, 062314 (2009).

[50] N. Johnston, D. W. Kribs, and V. I. Paulsen, Computing stabilized norms for quantum operations via the theory of completely bounded maps, Quantum Inf. Comput. **9**, 16 (2009).

[51] N. Minase, Smallest enclosing circle, http://www.nayuki.io/page/smallest-enclosing-circle.

[52] M. X. Goemans and D. P. Williamson, Approximation algorithms for max-3-cut and other problems via complex semidefinite programming, J. Comput. Syst. Sci. **68**, 442 (2004).

[53] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2009).

[54] J. Watrous, Simpler semidefinite programs for completely bounded norms, arXiv:1207.5726.

[55] J. Watrous, Semidefinite programs for completely bounded norms, Theory Comput. **5**, 217 (2009).

[56] N. Johnston, QETLAB: A MATLAB toolbox for quantum entanglement, http://www.qetlab.com.