

# Classically Computing Performance Bounds on Depolarized Quantum Circuits

Sattwik Deb Mishra<sup>1,†</sup>, Miguel Frías-Pérez<sup>2,†</sup> and Rahul Trivedi<sup>2,3,\*</sup>

<sup>1</sup>*Ginzton Laboratory, Stanford University, 348 Via Pueblo Mall, Stanford, California 94305, USA*

<sup>2</sup>*Max-Planck-Institute of Quantum Optics, Hans-Kopfermann-Str. 1, Garching 85748, Germany*

<sup>3</sup>*Electrical and Computer Engineering, University of Washington, Seattle, Washington 98195, USA*



(Received 20 July 2023; revised 14 February 2024; accepted 29 March 2024; published 19 April 2024)

Quantum computers and simulators can potentially outperform classical computers in finding ground states of classical and quantum Hamiltonians. However, if this advantage can persist in the presence of noise without error correction remains unclear. In this paper, by exploiting the principle of Lagrangian duality, we develop a numerical method to classically compute a certifiable lower bound on the minimum energy attainable by the output state of a quantum circuit in the presence of depolarizing noise. We provide theoretical and numerical evidence that this approach can provide circuit-architecture-dependent bounds on the performance of noisy quantum circuits.

DOI: [10.1103/PRXQuantum.5.020317](https://doi.org/10.1103/PRXQuantum.5.020317)

## I. INTRODUCTION

Fault-tolerant quantum computers hold promise for outperforming classical computers at several computational tasks. One of the most explored computational tasks is the problem of finding the ground state of a given many-body Hamiltonian—a problem that naturally arises in studying equilibrium properties of condensed matter systems [1]. Moreover, classical optimization problems can also be framed as finding ground states of commuting Hamiltonians [2]. Unsurprisingly, quantum algorithms for finding Hamiltonian ground states have been extensively studied [3–5] in search of a possible quantum advantage [6–8]—algorithms based on phase estimation and adiabatic evolution have been proposed for solving this problem, and have even been shown to be efficient for specific classes of Hamiltonians [9,10]. Furthermore, due to the constraints on available quantum hardware, there has been intense activity in exploring hardware-efficient heuristics for solving this problem, such as quantum adiabatic algorithms or variational quantum algorithms [11–14].

Current noisy intermediate-scale quantum devices, however, do not perform quantum error correction and consequently noise places a severe constraint on the performance of these quantum algorithms. From a theoretical

standpoint, it has thus become of interest to develop no-go results by providing theoretical bounds on the minimum energy that a noisy quantum circuit can achieve for a given Hamiltonian—if a classical algorithm [15–17] could obtain an energy better than this lower bound then we can conclude that a reduction in noise rate is necessarily needed for a possible quantum advantage. An approach to assessing the impact of noise on quantum algorithms is to directly simulate the circuit more so since the presence of noise in quantum circuits is expected to make them easier to classically simulate [18,19]. In fact, there have been several recent demonstrations of noisy quantum circuit simulations using tensor network methods [20–22]. However, most of the tensor network methods lack rigorous accuracy guarantees and cannot certify an accurate simulation of the quantum circuit. In particular, they are expected to deviate significantly from the circuit output as the noise rate continues to decrease and thus fall short of rigorously providing a no-go result for quantum advantage.

Alternatively, this problem can be approached analytically using tools from quantum information theory. For instance, Refs. [23–25] analyzed the increase in entropy of the quantum state due to noise, and showed that it can allow for an analytical lower bound on the attainable minimum energy. However, while providing rigorous no-go results, these analyses were circuit-architecture independent and were thus expected to underestimate the impact of noise. Certain circuit architectures are expected to significantly worsen the impact of noise, and this phenomenon has been theoretically demonstrated in random quantum circuits models [26,27]. However, it remains unclear if it is possible to provide an architecture-dependent lower bound for a specific engineered quantum circuit.

\*Corresponding author. [rtriv@uw.edu](mailto:rtriv@uw.edu)

†These authors contributed equally to this work.

*Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

In this article, we propose a method for efficiently computing rigorous bounds on the performance of any *specified* quantum circuit in the presence of a constant rate of depolarizing noise. The key insight behind our proposed method is the formulation of a Lagrangian dual corresponding to the circuit dynamics, which allows us to account for the circuit architecture in addition to the increase in the entropy, or, equivalently, the decrease in the purity of the quantum state. We show that the Lagrangian dual yields a hierarchy of classically computable lower bounds on energy, with respect to a specified Hamiltonian, obtained at the output of a noisy quantum circuit. We provide numerical and analytical evidence that this formulation can capture the circuit-architecture-dependent propagation of errors through the noisy quantum circuit and thus provide more stringent lower bounds than currently available. Our work is, in part, motivated by the application of Lagrangian duality to provide performance bounds on classical physical systems [28–33] and quantum optical devices [34,35].

## II. NOTATION

Given a finite-dimensional Hilbert space  $\mathcal{H}$ , we use  $\mathcal{D}_1(\mathcal{H})$  to denote the set of all density matrices on  $\mathcal{H}$ , and  $\mathcal{M}(\mathcal{H})$  to denote the set of Hermitian linear operators on  $\mathcal{H}$ . Unless otherwise mentioned, for any linear operator  $A$  on  $\mathcal{H}$ ,  $\|A\|$  will denote its operator norm, i.e., the maximum singular value of  $A$ , and  $\|A\|_F = \sqrt{\text{Tr}(A^\dagger A)}$  will denote its Frobenius norm.

We use the computer-science big-O notation for function asymptotics. In particular, given two functions  $f, g : [0, \infty) \rightarrow [0, \infty)$ ,  $f = O(g)$  if, for some  $c > 0$ ,  $f(x) \leq cg(x)$  as  $x \rightarrow \infty$  and  $f = \Omega(g)$  if, for some  $c > 0$ ,  $f(x) \geq cg(x)$  as  $x \rightarrow \infty$ .

## III. DUALITY-BASED BOUNDS

### A. Single-qubit example

As a simple illustrative example of the Lagrangian dual formulation, we first consider a single-qubit circuit [Fig. 1(a)]—consider a qubit initially in  $|0\rangle$ , with a gate  $U = e^{-i\theta Y}$  being applied on it followed by depolarizing noise with probability  $p$ . We would like to find the parameter  $\theta$  to minimize the energy corresponding to the Hamiltonian  $H = \Delta Z$ —in the absence of noise ( $p = 0$ ), it is straightforward to verify that this would be accomplished by setting  $\theta = \pi/2$  to obtain an energy  $-\Delta$ .

However, in the presence of depolarizing noise, the qubit will necessarily be in a mixed state. The extent to which the state is mixed can be quantified with a purity measure, for instance the von Neumann entropy of the qubit state, its trace purity, or even higher-order Rényi entropies [36]. For concreteness, we use the trace purity measure of a state  $\rho$ :  $P(\rho) = \text{Tr}(\rho^2)$ — $P(\rho) = 1$  if

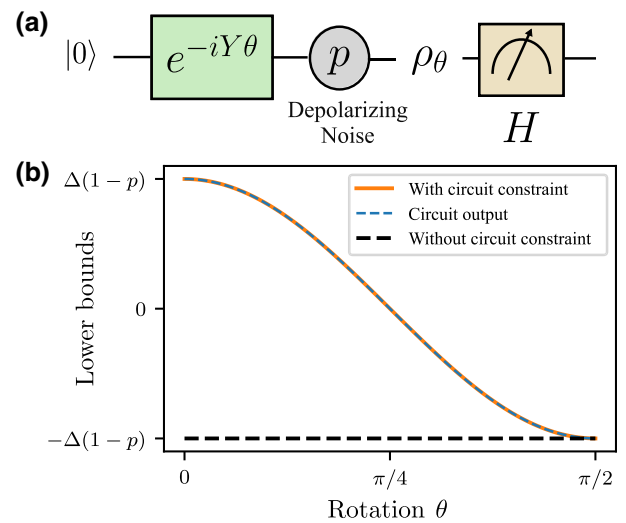


FIG. 1. (a) Schematic of the circuit considered in the single-qubit circuit example. It consists of a  $Y$ -axis rotation followed by depolarizing noise acting with probability  $p$ . (b) Comparison of bounds, with and without accounting for circuit constraints, on the minimum energy corresponding to the Hamiltonian  $H = \Delta Z$  attainable by the output of the single-qubit circuit shown in the schematic.

and only if  $\rho$  is a pure state; otherwise,  $P(\rho) < 1$ . Now, since the state  $\rho_\theta$  at the output of the single-qubit circuit in Fig. 1(a) is obtained by applying the depolarizing noise channel to a single-qubit pure state,  $P(\rho_\theta) = P_0 := p^2/4 + (1 - p/2)^2 < 1$ . Since  $\rho_\theta$  is necessarily mixed, it cannot produce the pure ground state of Hamiltonian  $H$  perfectly, irrespective of the choice of  $\theta$ —in fact, this simple observation can be used to lower bound the energy that can possibly be obtained at the output of the circuit by minimizing it with respect to states with purity at most  $P_0$ , i.e., solving the optimization problem

$$\begin{aligned} & \underset{\rho \in \mathcal{D}_1(\mathbb{C}^2)}{\text{minimize}} && \text{Tr}(H\rho) \\ & \text{subject to} && P(\rho) \leq P_0, \end{aligned} \quad (1)$$

where  $\mathcal{D}_1(\mathbb{C}^d)$  is the space of density matrices on the Hilbert space  $\mathbb{C}^d$ . The optimization problem in Eq. (1) is solved by  $\rho = (1 - p/2)|1\rangle\langle 1| + (p/2)|0\rangle\langle 0|$  with energy  $-\Delta(1 - p)$ . This bound clearly exhibits the intuitively expected dependence on the noise rate  $p$ —if  $p = 0$  then the energy attained coincides with the ground-state energy of  $-\Delta$ , and if  $p = 1$ , it is simply the energy obtained by the maximally mixed state.

However, this bound does not account for the unitary being applied on the qubit, and a better bound can be obtained by explicitly accounting for the circuit. To do so, we use the method of Lagrange duality [28,37]. For this, we extend the problem in Eq. (1) by adding an additional

constraint due to the circuit:

$$\underset{\rho \in \mathcal{D}_1(\mathbb{C}^2)}{\text{minimize}} \quad \text{Tr}(H\rho) \quad (2a)$$

$$\text{subject to} \quad \rho = \mathcal{E}_\theta(\rho_0), \quad (2b)$$

$$P(\rho) \leq P_0. \quad (2c)$$

Here  $\rho_0 = |0\rangle\langle 0|$  and  $\mathcal{E}_\theta$  is the channel corresponding to the unitary  $e^{-iY\theta}$  followed by the single-qubit depolarizing noise. Note that the purity constraint  $P(\rho) \leq P_0$  is redundant and is already implied by the circuit constraint  $\rho = \mathcal{E}_\theta(\rho_0)$ . However, as we will see in the following discussion, while redundant constraints do not impact the solution of an optimization problem, depending on the specific technique used to obtain a lower bound on the problem, they can have a considerable impact.

To provide a lower bound on this optimization, we construct its Lagrangian  $\mathcal{L}(\sigma, \lambda)$  by introducing Lagrange multipliers  $\sigma \in \mathcal{M}(\mathbb{C}^2)$ , and  $\lambda \geq 0$ :

$$\begin{aligned} \mathcal{L}(\rho, \sigma, \lambda) &= \text{Tr}[H\rho] + \text{Tr}[\sigma(\rho - \mathcal{E}_\theta(|0\rangle\langle 0|))] \\ &\quad + \lambda[P(\rho) - P_0]. \end{aligned} \quad (3)$$

Lagrangian  $\mathcal{L}(\rho, \sigma, \lambda)$  can be considered to be a modified energy function that, in addition to the energy  $\text{Tr}[H\rho]$ , also penalizes violation of the two constraints:  $\rho = \mathcal{E}_\theta(\rho_0)$  imposed by the circuit and  $P(\rho) \leq P_0$  on the purity of state  $\rho$ . Minimizing the Lagrangian with respect to  $\rho$ , we obtain the *dual function*

$$g(\sigma, \lambda) = \min_{\rho} \mathcal{L}(\rho, \sigma, \lambda), \quad (4)$$

which is a function of  $\sigma, \lambda$ , the *dual variables*. It follows from the principle of Lagrange duality that, for any  $\sigma$  and  $\lambda \geq 0$ ,  $g(\sigma, \lambda)$  is a lower bound on the energy attained by the circuit. This can easily be seen from Eq. (3) by noting that, when  $\mathcal{L}$  is evaluated at the circuit output  $\rho_\theta = \mathcal{E}_\theta(|0\rangle\langle 0|)$ , we obtain

$$\mathcal{L}(\rho_\theta, \sigma, \lambda) = \text{Tr}[H\rho_\theta] + \lambda[P(\rho_\theta) - P_0] \leq \text{Tr}[H\rho_\theta],$$

since  $P(\rho_\theta) \leq P_0$  and  $\lambda \geq 0$ . Since from Eq. (4)  $g(\sigma, \lambda)$  is the smallest attainable value of  $\mathcal{L}(\rho, \sigma, \lambda)$  on varying  $\rho$ , we obtain  $g(\sigma, \lambda) \leq \text{Tr}[H\rho_\theta]$ . We emphasize that the dual function  $g(\sigma, \lambda)$ , evaluated at any  $\sigma, \lambda \geq 0$ , is a lower bound on the energy  $\text{Tr}[H\rho_\theta]$  attained by the circuit, and the best lower bound can be obtained by maximizing  $g(\sigma, \lambda)$  with respect to  $\sigma, \lambda$ . Furthermore, since the construction of the dual function explicitly accounts for the circuit constraint, it gives a better bound than obtained from the problem in Eq. (1), i.e., by just accounting for the final purity of the state. This is exhibited in Fig. 1(b), where  $\max_{\sigma, \lambda \geq 0} g(\sigma, \lambda)$  compared with  $-\Delta(1-p)$  and it can be seen that the dual function provides a better lower

bound for most values of  $\theta$ . For the simple example of a single qubit, the duality-based bound that we can compute coincides exactly with the circuit output and thus models it exactly. As we will see in the next sections, this will not be the case for circuits over a large number of qubits.

### B. General formulation

We can now extend the duality lower bound to more general quantum circuits [Fig. 2]—consider a quantum circuit of depth  $d$ , consisting of unitaries  $U_1, U_2, \dots, U_d$  that have been designed to approximate the ground state of a target Hamiltonian  $H$  of  $N$  qubits. In the presence of noise, the state becomes increasingly mixed as the unitaries are applied on it—while it is typically hard to compute exactly how mixed the state is, an analytical upper bound on several purity measures can be obtained. In particular, explicit upper bounds for two purity measures after  $t$  time steps were established in Refs. [23,25]—the information content of the quantum state, as well as its trace purity.

*Lemma 1 (Refs. [23,25]).* Suppose that  $\rho_t$  is the quantum state on  $N$  qubits obtained from an initial pure state after applying  $t$  unitaries and single-qubit depolarizing channels; then

$$I(\rho_t) := N + \text{Tr}[\rho_t \log_2(\rho_t)] \leq N(1-p)^t,$$

$$P_{\text{tr}}(\rho_t) := \text{Tr}(\rho_t^2) \leq 2^{-N[1-(1-p)^t]},$$

where  $p$  is the probability of applying the depolarizing noise at each time step independently on each qubit.

Both the information content and trace purity can be viewed as measures of how mixed the given state is. Both are largest for a pure state [ $I(|\psi\rangle\langle\psi|) = N$  and  $P_{\text{tr}}(|\psi\rangle\langle\psi|) = 1$ ], and are lowest for the maximally mixed state [ $I(I/2^N) = 0$  and  $P_{\text{tr}}(I/2^N) = 2^{-N}$ ].

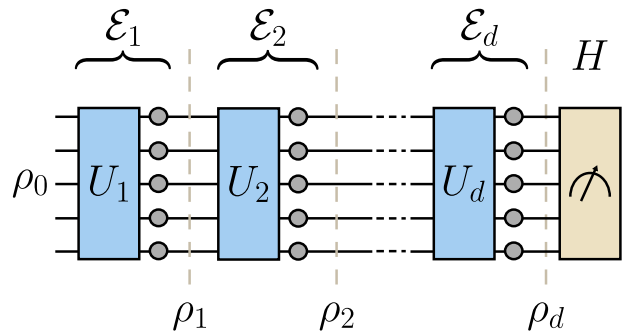


FIG. 2. Schematic depiction of the problem setting considered in this paper. The unitaries (colored boxes) implement a quantum algorithm to prepare an approximation of the ground state of a target Hamiltonian  $H$  in the absence of noise. Each layer of the unitary is followed by single-qubit depolarizing noise (gray circles) on the qubits applied with a probability  $p$ .

In the remainder of this subsection, we denote by  $P_t$  an upper bound on the purity of the state at the  $t$ th time step—we formulate the results of this subsection for general convex purity measures, and specialize them to concrete purity measures (such as the information content or trace purity) in the following sections. Now, as with the single-qubit case, the energy attained at the output of the circuit can now be written as

$$\underset{\rho_1, \rho_2, \dots, \rho_d \in \mathcal{S}}{\text{minimize}} \quad \text{Tr}(H \rho_d) \quad (5a)$$

$$\text{subject to} \quad \rho_t = \mathcal{E}_t(\rho_{t-1}), \quad t \in \{1, \dots, d\}, \quad (5b)$$

$$P(\rho_t) \leq P_t, \quad t \in \{1, \dots, d\}, \quad (5c)$$

where  $\mathcal{E}_t(\cdot)$  is the quantum channel that applies the unitary  $U_t$  for the  $t$ th layer of the circuit followed by depolarizing noise acting individually on the qubits, and  $\rho_0$  is a fixed and known initial state. Importantly, in Eq. (5),  $\mathcal{S}$  is the set of  $N$ -qubit operators over which we allow the states  $\rho_1, \rho_2, \dots, \rho_d$  to vary—this set can be chosen to be *any* set containing density matrices over  $N$  qubits  $\mathcal{D}_1[(\mathbb{C}^2)^{\otimes N}]$  since the circuit constraints  $[\rho_t = \mathcal{E}_t(\rho_{t-1})]$  enforce  $\rho_1, \rho_2, \dots, \rho_d$  to be valid density matrices. For instance,  $\mathcal{S}$  can be chosen to be just the set of  $N$ -qubit Hermitian operators, or the set of  $N$ -qubit Hermitian operators with unity trace. As we will see below, the choice of this set together with the purity measure determines the form of the dual function.

To construct the dual function corresponding to Eq. (5), we introduce the dual variables  $\vec{\sigma} = \{\sigma_1, \sigma_2, \dots, \sigma_d\}$ , which are  $N$ -qubit Hermitian operators, corresponding to the circuit constraints, and  $\vec{\lambda} = \{\lambda_1, \lambda_2, \dots, \lambda_d \geq 0\}$ , corresponding to the purity constraints. The Lagrangian is now constructed by adding penalties corresponding to the circuit constraints and purity constraints at each time step to the output energy:

$$\begin{aligned} \mathcal{L}(\vec{\rho}, \vec{\sigma}, \vec{\lambda}) &= \text{Tr}[H \rho_d] + \sum_{t=1}^d \text{Tr}\{\sigma_t[\rho_t - \mathcal{E}_t(\rho_{t-1})]\} \\ &\quad + \sum_{t=1}^d \lambda_t [P(\rho_t) - P_t], \\ &= \sum_{t=1}^d \text{Tr}(\rho_t H_t) + \lambda_t [P(\rho_t) - P_t] \end{aligned} \quad (6)$$

with  $H_d = H + \sigma_d$ ,  $H_t = \sigma_t - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1})$  for  $t \in \{1, 2, \dots, d-1\}$ . The dual function is obtained by minimizing the

Lagrangian with respect to  $\rho_1, \rho_2, \dots, \rho_d \in \mathcal{S}$ , i.e.,

$$\begin{aligned} g(\vec{\sigma}, \vec{\lambda}) &= \min_{\rho_1, \rho_2, \dots, \rho_d \in \mathcal{S}} \mathcal{L}(\vec{\rho}, \vec{\sigma}, \vec{\lambda}), \\ &= \sum_{t=1}^d \min_{\rho_t \in \mathcal{S}} \{[\text{Tr}(\rho_t H_t) + \lambda_t P(\rho_t)] - \lambda_t P_t\}, \\ &= \sum_{t=1}^d [\mathcal{F}_{\mathcal{S}, P}(H_t, \lambda_t) - \lambda_t P_t], \end{aligned} \quad (7a)$$

where

$$\mathcal{F}_{\mathcal{S}, P}(H, \lambda) = \min_{\rho \in \mathcal{S}} \{\text{Tr}[H \rho] + \lambda P(\rho)\}. \quad (7b)$$

As with the single-qubit example, the dual function is a lower bound on the energy produced at the circuit output for any  $\vec{\sigma}, \vec{\lambda} \geq 0$ , i.e.,

$$g(\vec{\sigma}, \vec{\lambda}) \leq \text{Tr}[\rho_d H] \quad \text{for all } \sigma_t \in \mathcal{M}[(\mathbb{C}^2)^{\otimes N}], \lambda_t \geq 0.$$

The function  $\mathcal{F}_{\mathcal{S}, P}(H, \lambda)$  can be interpreted as a generalized free energy corresponding to Hamiltonian  $H$  at temperature  $\lambda$  that depends on both the domain  $\mathcal{S}$  and the purity measure  $P$ . For instance, if the purity measure is taken to be the information content  $I(\rho) = N - \text{Tr}[\rho \log_2(\rho)]$  then it reduces to the Gibbs free energy with an offset of  $N\lambda$ . However, by choosing different purity measures  $P$  as well as different domains  $\mathcal{S}$ , the dual function allows us to obtain a family of bounds on the noisy quantum circuit. As we will see in the next section, certain choices of  $P$  and  $\mathcal{S}$  provide lower bounds that can be classically computed.

Consider first the best lower bound that can be obtained from the dual function. In the following proposition, we show that the best lower bound attained by the dual function is exactly equal to the energy attained by the quantum circuit, and the choice of dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d$  that yields the largest value dual function corresponds to the Heisenberg picture evolution of Hamiltonian  $H$ .

*Proposition 1.* For the dual function defined in Eq. (7), it follows that its maximum over the dual variables is equal to the output energy of the noisy circuit, i.e.,

$$\underset{\substack{\sigma_1, \sigma_2, \dots, \sigma_d \in \mathcal{M}[(\mathbb{C}^2)^{\otimes N}] \\ \lambda_1, \lambda_2, \dots, \lambda_d \geq 0}}{\text{maximum}} \quad g(\vec{\sigma}, \vec{\lambda}) = \text{Tr}[H \mathcal{E}_d \mathcal{E}_{d-1} \cdots \mathcal{E}_1(\rho_0)],$$

and the maximum is attained at

$$\sigma_d = -H, \quad \sigma_t = -\mathcal{E}_{t+1}^\dagger \mathcal{E}_{t+2}^\dagger \cdots \mathcal{E}_d^\dagger(H),$$

and  $\lambda_1 = \lambda_2 = \cdots = \lambda_d = 0$ .

*Proof.* The proof of this proposition follows simply by noting that, from the definition,

$$\mathcal{F}_{S,P}(0, 0) = 0.$$

Now, if  $\sigma_d = -H$ , and  $\sigma_i = -\mathcal{E}_{i+1}^\dagger \mathcal{E}_{i+2}^\dagger \cdots \mathcal{E}_d^\dagger(H)$ , then  $H_i = 0$ . Hence, at this value of  $\vec{\sigma}$  and at  $\lambda = 0$ , we obtain  $g(\vec{\sigma}, \vec{\lambda}) = \text{Tr}[\rho_0 \mathcal{E}_1^\dagger \mathcal{E}_2^\dagger \cdots \mathcal{E}_d^\dagger(H)] = \text{Tr}[H \mathcal{E}_d \mathcal{E}_{d-1} \cdots \mathcal{E}_1(\rho_0)]$ . Since  $\text{Tr}[H \mathcal{E}_d \mathcal{E}_{d-1} \cdots \mathcal{E}_1(\rho_0)]$  is also an upper bound of  $g(\vec{\sigma}, \vec{\lambda})$ , the proposition follows. ■

This proposition establishes that finding the best dual bound is equivalent to exactly simulating the circuit, which we expect to be hard to do on classical computers. This hardness fundamentally stems from the fact that the dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d$  are operators in an exponentially large space. However, since the dual function  $g(\vec{\sigma}, \vec{\lambda})$  is a lower bound on the output energy for any  $\vec{\sigma}, \vec{\lambda}$ , a natural approach to evaluate a lower bound would be restrict  $\sigma_i$  to subsets of  $\mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$  where the dual function could be evaluated efficiently—the specific subset would depend on the choice of the purity measure. In Sec. IV below, we will see that the dual function obtained on choosing the purity measure to be the trace purity and the domain  $\mathcal{S} = \mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$  of  $N$ -qubit Hermitian operators can be evaluated efficiently if  $\sigma_1, \sigma_2, \dots, \sigma_d$  are chosen to be matrix product operators (MPOs) of bond dimension  $\text{poly}(N)$ . In Sec. V below, we consider the dual function obtained on choosing the purity measure to be the information content of the state, in which case  $\sigma_1, \sigma_2, \dots, \sigma_d$  can be restricted to the space of geometrically local Hamiltonians, allowing for an exact evaluation of the dual function.

Restricting the dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d$  to a subset of  $\mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$  raises the question of whether the maximum value that the dual function can attain within this restricted set of dual variables gives a better lower bound on the energy compared to neglecting the circuit constraints and just accounting for the purity of the final state, i.e., does the duality-based bound still account for the circuit architecture. Our next proposition answers this question affirmatively, and shows that a better lower bound can be obtained as long as the restricted set of dual variables contains 0.

*Proposition 2.* Suppose that  $\mathcal{S}_\sigma \subset \mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$ , such that  $0 \in \mathcal{S}_\sigma$ ; then

$$\underset{\substack{\sigma_1, \sigma_2, \dots, \sigma_d \in \mathcal{S}_\sigma \\ \lambda_1, \lambda_2, \dots, \lambda_d \geq 0}}{\text{maximum}} g(\vec{\sigma}, \vec{\lambda}) \geq \underset{\rho \in \mathcal{S}, P(\rho) \leq P_d}{\text{minimum}} \text{Tr}(H\rho).$$

*Proof.* Since  $0 \in \mathcal{S}_\sigma$ ,

$$\underset{\lambda \geq 0}{\text{maximum}} g(\{0 \cdots 0\}, \{0 \cdots \lambda\}) \leq \underset{\substack{\sigma_1, \sigma_2, \dots, \sigma_d \in \mathcal{S}_\sigma \\ \lambda_1, \lambda_2, \dots, \lambda_d \geq 0}}{\text{maximum}} g(\vec{\sigma}, \vec{\lambda}).$$

Now, we can note that

$$g(\{0 \cdots 0\}, \{0 \cdots \lambda\}) = \mathcal{F}_{S,P}(H, \lambda) - \lambda P_d.$$

It can be noted that  $g(\{0 \cdots 0\}, \{0 \cdots \lambda\})$  is simply the dual function of the convex problem

$$\begin{aligned} & \underset{\rho \in \mathcal{S}}{\text{minimize}} \quad \text{Tr}(H\rho) \\ & \text{subject to} \quad P(\rho) \leq P_d. \end{aligned}$$

Furthermore, this convex problem trivially satisfies Slater's conditions [37,38]. This can be checked by noting that Slater's conditions are satisfied if there is a  $\rho \in \mathcal{S}$  such that  $P(\rho) < P_d$ —this follows by noting that  $P(I/2^N) < P_d$  and  $I/2^N \in \mathcal{D}_1[(\mathbb{C}^2)^{\otimes N}] \subseteq \mathcal{S}$ . Since Slater's conditions are satisfied, this problem is strongly dual and, consequently, the optimal duality bound is equal to the solution of the optimization problem, i.e.,

$$\underset{\lambda \geq 0}{\text{maximum}} g(\{0 \cdots 0\}, \{0 \cdots \lambda\}) = \underset{\rho \in \mathcal{S}, P(\rho) \leq P_d}{\text{minimum}} \text{Tr}(H\rho),$$

which proves the proposition. ■

While this proposition indicates that accounting for the circuit constraints while constructing the lower bound results in an improvement over only accounting for the final purity even with a restricted space of dual variables, it says nothing about the extent to which the bound improves. We expect the improvement to be strongly dependent on the purity function  $P$ , the domain set  $\mathcal{S}$ , and the dual set  $\mathcal{S}_\sigma$  used in formulating and evaluating the bound. In the next section, we consider a specific formulation of the dual function that uses the trace purity measure, and show that the lower bound obtained on accounting for the circuit constraints can be exponentially better than if the circuit constraints were not accounted for.

## IV. TRACE PURITY-BASED LOWER BOUND

### A. Formulation

In this section, we now consider a specific choice of the purity function and the domain  $\mathcal{S}$  that results in a dual function that can be computed exactly when the dual variables are parametrized as matrix product operators with  $\text{poly}(N)$  bond dimension. We choose the purity measure to be the trace purity  $P(\rho) = P_{\text{tr}}(\rho) = \text{Tr}(\rho^2)$ , and the domain  $\mathcal{S}$  in Eq. (5) to be the space of Hermitian  $N$ -qubit operators  $\mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$ . It then follows that  $\mathcal{F}_{S,P}(H, \lambda)$

defined in Eq. (7b) evaluates to

$$\mathcal{F}_{S,P}(H, \lambda) = -\frac{\text{Tr}(H^2)}{4\lambda},$$

and, therefore, we obtain

$$g(\vec{\sigma}, \vec{\lambda}) = -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] - \sum_{t=1}^d \left( \frac{\text{Tr}(H_t^2)}{4\lambda_t} + \lambda_t P_t \right),$$

where  $H_d = H + \sigma_d$  and  $H_t = \sigma_t - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1})$ . Furthermore, for this dual function, it is possible to perform the maximization over  $\vec{\lambda}$  analytically to obtain

$$\begin{aligned} h(\vec{\sigma}) &= \max_{\vec{\lambda} \geq 0} g(\vec{\sigma}, \vec{\lambda}) \\ &= -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] - \sum_{i=1}^d \sqrt{P_i \text{Tr}(H_i^2)}. \end{aligned} \quad (8)$$

From the expression for  $h(\vec{\sigma})$ , we immediately note that if  $\sigma_1, \sigma_2, \dots, \sigma_d$  are restricted to be matrix product operators with bond dimension  $D$  then  $h(\vec{\sigma})$  can be evaluated classically in time  $NdD^4$ . However, as we established in Proposition 1, the best lower bound is  $h(\vec{\sigma})$  when evaluating it at  $\vec{\sigma}$  corresponding to a Heisenberg picture evolution of Hamiltonian  $H$ . While, for most problems of interest (e.g., where  $H$  is a local or spatially local Hamiltonian),  $H$  can be represented as a matrix product operator of a modest bond dimension, the unitaries involved in the circuit can, in general, grow its bond dimension exponentially. A natural choice of  $\sigma_1, \sigma_2, \dots, \sigma_d$  would then be to perform time-evolving block decimation (TEBD) [39–43] on the Heisenberg evolution and compress the operators in each step into bond dimension  $D$ , i.e., at  $\vec{\sigma}^{h,D} = \{\sigma_1^{h,D}, \sigma_2^{h,D}, \dots, \sigma_d^{h,D}\}$ ,

$$\sigma_d^{h,D} = -H$$

$$\text{and } \sigma_t^{h,D} = \Pi_D \mathcal{E}_{t+1}^\dagger(\sigma_{t+1}^{h,D}) \quad \text{for } t \in \{1, 2, \dots, d-1\}, \quad (9)$$

where  $\Pi_D$  compresses an  $N$ -qubit operator to an operator with bond dimension  $D$  [44].

### 1. Duality bound and TEBD truncation errors

If a Heisenberg picture TEBD simulation, for some bond dimension  $D$ , of the noisy quantum circuit is exact, then, by Proposition 1, the duality-based bound  $h(\vec{\sigma}^{h,D})$  is exactly equal to the expected energy at the output of the circuit. In practice, for small bond dimensions  $D$ , the TEBD algorithm is not exactly correct, but incurs an error. However, as shown below, an upper bound on this error can also be efficiently computed for the TEBD algorithm. Consequently, tracking the error incurred in the TEBD algorithm

allows us to calculate another lower bound on the output of the quantum circuit, i.e., if the TEBD algorithm produces an estimate  $E_{\text{TEBD}}$ , the output energy  $E$  of the circuit within an additive error  $\delta$ , then  $E_{\text{TEBD}} - \delta$  also lower bounds the energy  $E$ . A natural question to ask is if the duality-based bounds are more informative than the bound obtained from just a TEBD simulation.

Consider now the problem of estimating the TEBD error following Ref. [40]. The TEBD estimate of the energy at the circuit output,  $E_{\text{TEBD}}$ , can be expressed as

$$E_{\text{TEBD}} = \text{Tr} \left[ \mathcal{E}_1(\rho_0) \left( \prod_{t=2}^d \Pi_D \mathcal{E}_t^\dagger \right) (H) \right],$$

while the true energy at the circuit output can be expressed as

$$E = \text{Tr} \left[ \mathcal{E}_1(\rho_0) \left( \prod_{t=2}^d \mathcal{E}_t^\dagger \right) (H) \right].$$

Denoting by  $\rho_t$  the state of the qubits in the quantum circuit at time step  $t$ ,  $\rho_t = \mathcal{E}_t \mathcal{E}_{t-1} \dots \mathcal{E}_1(\rho_0)$ , we note that

$$\begin{aligned} E - E_{\text{TEBD}} &= \text{Tr} \left[ \rho_1 \left( \prod_{t=2}^d \mathcal{E}_t^\dagger - \prod_{t=2}^d \Pi_D \mathcal{E}_t^\dagger \right) (H) \right], \\ &= \sum_{t=2}^d \text{Tr} \left[ \rho_1 \left( \prod_{s=2}^{t-1} \mathcal{E}_s^\dagger \right) (\mathcal{E}_t^\dagger - \Pi_D \mathcal{E}_t^\dagger) \right. \\ &\quad \left. \times \left( \prod_{s=t+1}^d \Pi_D \mathcal{E}_s^\dagger \right) (H) \right], \\ &= \sum_{t=2}^d \text{Tr} [\rho_{t-1} (\sigma_{t-1}^{h,D} - \mathcal{E}_t^\dagger(\sigma_t^{h,D}))] \\ &= \sum_{t=1}^{d-1} \text{Tr}(\rho_t H_t), \end{aligned}$$

where, in the last step, we have used the fact that, by definition,  $H_t = \sigma_{t-1}^{h,D} - \mathcal{E}_t^\dagger(\sigma_t^{h,D})$ . Now, an upper bound on the error  $|E - E_{\text{TEBD}}|$  can be obtained via

$$|E - E_{\text{TEBD}}| \leq \sum_{t=1}^{d-1} |\text{Tr}(\rho_t H_t)| \leq \sum_{t=1}^{d-1} \|\rho_t\|_F \|H_t\|_F, \quad (10)$$

where  $\|A\|_F = \sqrt{\text{Tr}(A^\dagger A)}$  and we have used the fact that, by Holder's inequality,  $|\text{Tr}(\rho_t H_t)| \leq \|\rho_t\|_1 \|H_t\|_F \leq \|H_t\|_F$  since  $\|\rho_t\|_1 = 1$  and  $\|H_t\|_F \leq \|H_t\|_F$ . We point out that an important reason why we express the error bound in terms of the Frobenius norm of  $H_t$ , instead of its operator norm, is because the Frobenius norm can be efficiently computed if  $H_t$  is a matrix product operator of a small bond dimension

(which is the case while performing the TEBD simulation). The deviation bound in Eq. (10) implies a lower bound

$$E \geq E_{\text{TEBD}} - \delta = -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] - \sum_{i=1}^d \sqrt{\text{Tr}(H_i^2)}. \quad (11)$$

This bound is significantly worse than the duality-based bound in Eq. (8) as  $P_t \ll 1$  for all time steps  $t$ . The key reason why just accounting for a worst-case accumulation of TEBD errors yields a loose lower bound is that the upper bound in  $|E - E_{\text{TEBD}}|$  does not account for the decrease in the trace purity of the quantum state in the presence of noise, which is explicitly factored into the formulation of the dual.

In Fig. 3, we numerically exhibit the difference between the bound in Eq. (11) and  $h(\vec{\sigma}^{h,D})$  for a one-dimensional (1D) circuit on  $N = 40$  qubits [Fig. 3(a)], which is chosen to find the ground state of a commuting 1D Hamiltonian (see the figure caption for the exact circuit and Hamiltonian). As can be seen from Fig. 3(b), the lower bound computed from the trace purity-based dual is significantly larger, and thus more representative of the impact of noise on the output energy, than the lower bound provided by Eq. (11). We point out that the dual variables  $\vec{\sigma}^{h,D}$  obtained by TEBD in the Heisenberg picture are not necessarily the globally optimal choice in the space of all MPOs with bond dimension  $D$  to evaluate the dual function  $h(\vec{\sigma})$ . The function  $h(\vec{\sigma})$  can potentially be optimized beyond the TEBD-based value to obtain better lower bounds. In practice, we observe that local optimization of  $h(\vec{\sigma})$  with a gradient-based method starting from the initial point of  $\vec{\sigma} = \vec{\sigma}^{h,D}$  yields only a modest improvement over  $h(\vec{\sigma}^{h,D})$ .

Next, we study the improvement that the duality-based bounds that account for the circuit constraint provide over bounds in existing literature that just account for the information content at the circuit output. In particular, we numerically compare the best lower bound  $\ell_D^{\text{dual}}$  that we can obtain by evaluating  $h(\vec{\sigma})$  at  $\sigma_1, \sigma_2, \dots, \sigma_d \in \text{MPO}_D$  (the space of all  $N$ -qubit MPOs of bond dimension  $D$ ),

$$\ell_D^{\text{dual}} = \underset{\sigma_1, \sigma_2, \dots, \sigma_d \in \text{MPO}_D}{\text{maximize}} \quad h(\vec{\sigma}),$$

to the lower bound  $\ell^I$  analyzed in Ref. [24],

$$\ell^I = \underset{\rho: I(\rho) \leq N(1-p)^d}{\text{minimize}} \quad \text{Tr}(H\rho),$$

i.e., where they accounted only for the decreased information content  $I(\rho) = N - \text{Tr}[\rho \log_2 \rho]$  of the final state as per Lemma 1. First, we show that there exists a Hamiltonian and a 1D circuit where  $\ell_D^{\text{dual}}$ , with  $D = O(N)$  [45], scales superexponentially with the depth of the circuit and thus captures the propagation of errors through the circuit, while  $\ell^I$  scales at most exponentially with the circuit depth.

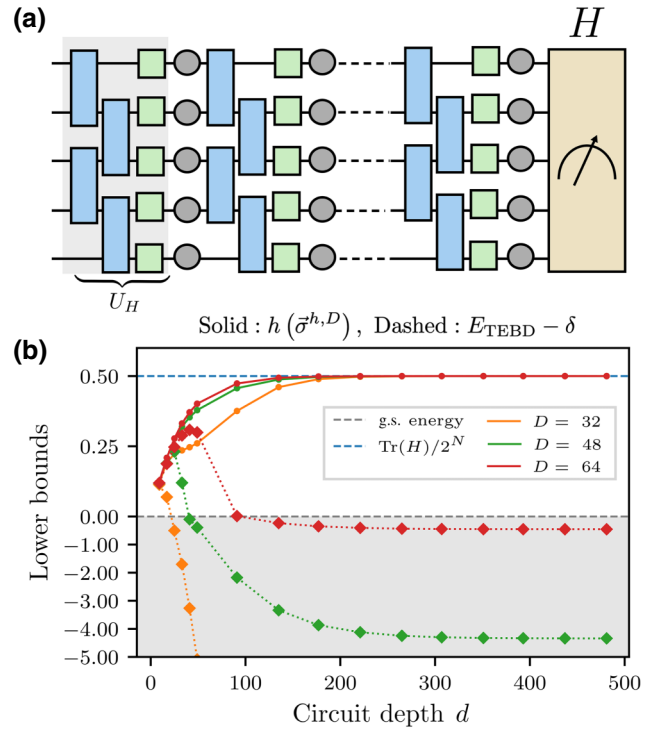


FIG. 3. (a) Schematic of benchmark circuits considered for one-dimensional spin systems: colored boxes indicate unitaries and gray circles depolarizing noise. Two-qubit unitaries are chosen to be  $\exp(-i\theta X \otimes X)$  and single-qubit unitaries are independently Haar random. The Hamiltonian is chosen to be  $H = -U_H(\sum_i Z_i)U_H^\dagger$ , where  $U_H$  is the first layer of unitaries, making  $H$  a 4-local commuting Hamiltonian. The first layer of unitaries  $U_H$  thus transforms the initial state  $|0\rangle^{\otimes N}$  into the ground state of  $H$ . The last  $(d-1)/2$  layers are chosen to be the inverse of the previous  $(d-1)/2$  layers—in the absence of noise, the output of the circuit is the ground state of  $H$ . (b) Plot shows the trace purity-based dual bound  $[h(\vec{\sigma}^{h,D})$  in Eq. (8); solid lines, circular markers] and the bound obtained by only considering the TEBD errors  $[E_{\text{TEBD}} - \delta$  in Eq. (11); dotted lines, diamond markers] for the ground-state (g.s.) energy of the target Hamiltonian, as a function of the circuit depth  $d$  for a system of  $N = 40$  spins, with two-qubit gate parameter  $\theta = 0.05$ , a depolarizing noise rate of  $p = 3\%$ , and varying MPO ansatz bond dimensions  $D$ . The gray dashed line indicates the g.s. energy, the gray shaded area indicates the region of trivial bounds (less than the g.s. energy), and the blue dashed line indicates the energy of the completely mixed state  $\mathbb{1}/2^N$ . The y axis is scaled by a constant multiplicative factor in the trivial region for visual clarity. The Hamiltonian is shifted and scaled such that its spectrum is in  $[0, 1]$ .

**Proposition 3.** There exists a 1D circuit and an  $N$ -qubit Hamiltonian  $H$  with  $\text{Tr}(H) = 0$  and  $\|H\| = N$ , such that  $\ell^I = -N(1-p)^{O(d)}$ , while  $\ell_D^{\text{dual}} = -N(1-p)^{\Omega(d^2)} + O(\sqrt{N})$  for a choice of  $D \leq O(N)$ .

*Sketch of proof (see Appendix A for details).* Consider a Clifford circuit chosen at random from the ensemble of entangle-unentangle circuits analyzed in Ref. [26]. It was

shown for this ensemble that, for a 1D circuit, on average, the energy of the output state with respect to a 2-local Hamiltonian converges to the energy of the maximally mixed state as  $\sim \text{poly}(N) \times (1-p)^{\Omega(d^2)}$ . Consider now the Hamiltonian  $H = -\sum_{i=1}^N Z_i$  and initial state  $|0\rangle^{\otimes N}$ . In the Heisenberg picture, each  $Z_i$  will be mapped to exactly one Pauli string under the action of Clifford gates [46,47]. Since a Pauli string is representable as an MPO of bond dimension 1,  $\sigma_i$  obtained from Heisenberg picture evolution will be a sum of  $N$  MPOs of bond dimension 1 and will thus be an MPO of bond dimension at most  $N$ . Thus, from Proposition 1, the purity-based dual exactly matches the energy of the output of the quantum circuit, which will scale as  $-N(1-p)^{\Omega(d^2)}$  [26,48]. As the bound without circuit constraints is agnostic to the unitaries in a circuit, it is also a lower bound on the circuit where all the unitaries are just the identity operation. For this trivial circuit, the energy of the state after  $d$  layers of just depolarizing noise scales as  $-N(1-p)^d$ . Hence, the lower bound without circuit constraints  $\ell^I = -N(1-p)^{O(d)}$ . ■

## B. Numerical studies

For non-Clifford circuits, MPO parametrization with bond dimension  $\text{poly}(N)$  is no longer expected to fully capture Heisenberg picture evolution. Even so, we numerically demonstrate for a 1D spin system that the circuit dual function given in Eq. (8) can be used to compute circuit-specific bounds that perform better than bounds that only take into account the information content of the output state. We compute the bounds by evaluating the dual function  $h(\vec{\sigma})$  at the dual variables obtained from TEBD on the Heisenberg picture as in Eq. (9). Figure 4 shows numerical studies of the bounds  $h(\vec{\sigma}^{h,D})$  computed in this manner—we consider a 1D spin system of size  $N = 40$  and circuits designed to prepare the ground state of a commuting local Hamiltonian (see the figure caption for details). The plots in Fig. 4 show the bounds for MPO ansatzes with different bond dimensions  $D$  plotted against the circuit depth  $d$  for circuits with noise rates  $p = 3\%$  [Fig. 4(a)],  $p = 5\%$  [Fig. 4(b)],  $p = 10\%$  [Fig. 4(c)], and  $p = 20\%$  [Fig. 4(d)]. However, for the lowest noise rate  $p = 3\%$ , the circuit dual bounds at intermediate depths are trivial, i.e., lower than the ground-state energy of  $H$ —this can be attributed to the fact that the intermediate depth regime is the regime where the MPO ansatz is least representative. For shorter depths, the bond dimension of the Heisenberg picture operator would have not grown very much, while, for very long depths, the action of the depolarizing noise reduces the bond dimension of the Heisenberg picture operator.

Figure 4 also compares the trace purity-based dual bound to the information content-based bound. However, since the duality-based bound is exactly computable on a classical computer, to make a fair comparison, we need

to use a certifiable method for computing the information content-based bound. In particular, using Lagrangian duality, the information content-based bound can be reframed in terms of the Gibbs free energy of the problem Hamiltonian, i.e.,

$$\begin{aligned} \ell^I &= \underset{\rho: I(\rho) \leq N(1-p)^d}{\text{minimize}} \text{Tr}(H\rho) \\ &= \underset{\lambda \geq 0}{\text{maximize}} \lambda S_d + G(H, \lambda), \end{aligned} \quad (12)$$

where  $S_d = N - N(1-p)^d$  and  $G(H, \lambda) = -\lambda \log \text{Tr} \exp(-H/\lambda)$  is the Gibbs free energy of  $H$  at temperature  $\lambda$ . However, since  $H$  is generally a many-body Hamiltonian, an accurate evaluation of  $G(H, \lambda)$  can only be guaranteed at sufficiently high temperatures [49]. Thus, instead of evaluating the bound  $\ell^I$  in Eq. (12), we introduce a lower bound  $\lambda_c$  on the temperature  $\lambda$  and evaluate

$$\ell_{\lambda_c}^I = \underset{\lambda \geq \lambda_c}{\text{maximize}} \lambda S_d + G(H, \lambda), \quad (13)$$

For spatially local Hamiltonians,  $\lambda_c$  can be chosen depending on the norms of the local terms in the Hamiltonian, the dimensionality of the lattice, and the interaction range. In our calculations, we make the choice of  $\lambda_c = 8e^3$ —this is based on Ref. [49] that, to the best of our knowledge, provides the only rigorous algorithm that works for evaluating  $G(H, \lambda)$  at temperatures above  $\lambda_c$ .

We see from Fig. 4 that the dual provides a tighter lower bound on the output than the bound based on just the information content of the output state. The information content-based bounds shown in Fig. 4 are also trivial (i.e., lower than the ground-state energy) for intermediate *and* short depths—this is due to the temperature lower bound that needs to be introduced to ensure computability of the Gibbs free energy. We also observe that the separation between the information content-based and circuit dual bounds increases with the bond dimension  $D$  as the MPO ansatz becomes more expressive with increasing bond dimension. In the limit of large circuit depth at nonzero depolarizing noise rates, the state of the circuit approaches the completely mixed state, and we observe that both bounds also approach the energy corresponding to the completely mixed state.

Finally, we demonstrate that the dual bounds are able to capture the extent of entanglement being generated in a circuit. In the brick-wall quantum circuits we consider, the two-qubit gates in the circuit  $\exp(-i\theta X \otimes X)$  are parametrized by an angle  $\theta$  that controls the entanglement being produced—for example, at  $\theta = 0, \pi/2$ , there is no entanglement at all. Figure 5 shows the bounds as a function of the angular parameter  $\theta$  and the noise rate  $p$  for constant bond dimensions  $D$  and circuit depths  $d$ , for a 1D system of  $N = 32$  spins and Figs. 6(d) and 6(e)



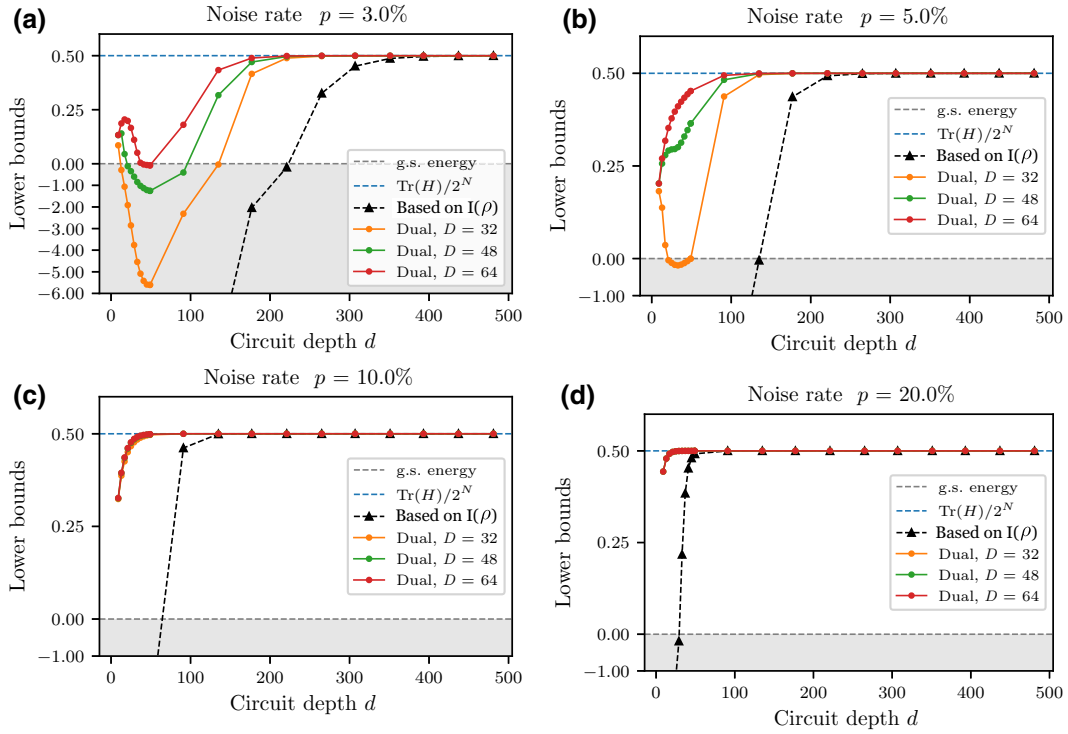


FIG. 4. Comparison of the trace purity–based dual bound  $[h(\bar{\sigma}^{h,D})]$  in Eq. (8) and the bound based on just the information content of the output state  $[\ell_{\lambda_c}^I]$  in Eq. (13) for 1D many-body spin systems. Both bounds are lower bounds on the ground-state (g.s.) energy of the same target Hamiltonian as considered for the results in Fig. 3 (see the description of the Hamiltonian in the caption of Fig. 3). Plots show the bounds as a function of brick-wall quantum circuit depth  $d$  [Fig. 3(a)] for a system of  $N = 40$  spins, varying MPO ansatz bond dimensions  $D$ , with depolarizing noise rates of (a)  $p = 3\%$ , (b)  $p = 5\%$ , (c)  $p = 10\%$ , and (d)  $p = 20\%$ . Note that at  $p = 10\%$  and  $20\%$ , approximately the same duality-based lower bound is obtained for different bond dimensions. Two-qubit unitaries in the brick-wall circuit are chosen to be  $\exp(-i\theta X \otimes X)$  with  $\theta = 0.1$  and single-qubit unitaries are independently Haar random. The gray dashed line indicates the g.s. energy, the gray shaded area indicates the region of trivial bounds (less than the g.s. energy), and the blue dashed line indicates the energy of the completely mixed state  $\mathbb{1}/2^N$ . The y axis is scaled by a constant multiplicative factor in the trivial region for visual clarity. The Hamiltonian is shifted and scaled such that its spectrum is in  $[0, 1]$ .

show the same for a 2D system of  $N = 36$  spins in a  $6 \times 6$  lattice. For the 2D system, we consider the MPO ansatz to have a “snakelike” bond structure on the 2D lattice [Fig. 6(c)]—such a snakelike structure is a numerically convenient approach for performing TEBD for 2D systems. This ansatz is useful for moderate system sizes, but due to gates along the vertical edges of the lattice, the bond dimension required grows rapidly. For larger system sizes, we expect that a tensor network ansatz that matches the architecture of the circuit [50–52] would give better bounds. For both the 1D and 2D systems, the target Hamiltonians are shifted and scaled such that the ground-state energies are zero and any bounds lower than zero are considered trivial and represented as zero in the plots—the black regions in the plots thus correspond to trivial bounds. We observe that, near  $\theta = 0, \pi/2$ , where the entanglement is small, the MPO ansatz of constant bond dimension used for the bounds is able to capture it and we obtain nontrivial bounds for small noise rates  $p \approx 6\%$ . For values of  $\theta$  away from these limits, the region of triviality is larger, but

nontrivial dual bounds can still be obtained for higher noise rates.

While we have restricted ourselves to an MPO ansatz even for the 2D problem, we point out that we can likely get better results by using a projected entangled-pair operator (PEPO) ansatz for the dual variables. However, contracting PEPOs is a computation that is, in the worst case, hard on classical computers [53] and we might lose the rigorous guarantee associated with the lower bound. However, several numerical studies do point out that, outside of the worst case, projected entangled-pair states and operators could still be contracted accurately with classically efficient heuristics [54–56]. Furthermore, there are certain physically meaningful subclasses of tensor networks on 2D lattices where classical algorithms for contraction, often with rigorous guarantees, are available and thus could be used as ansatzes for the dual variables. An example of such a subclass are sequentially generated projected entangled-pair tensors [57]. Other subclasses, such as string bond tensors [58,59] and entangled plaquette tensors [60,61],

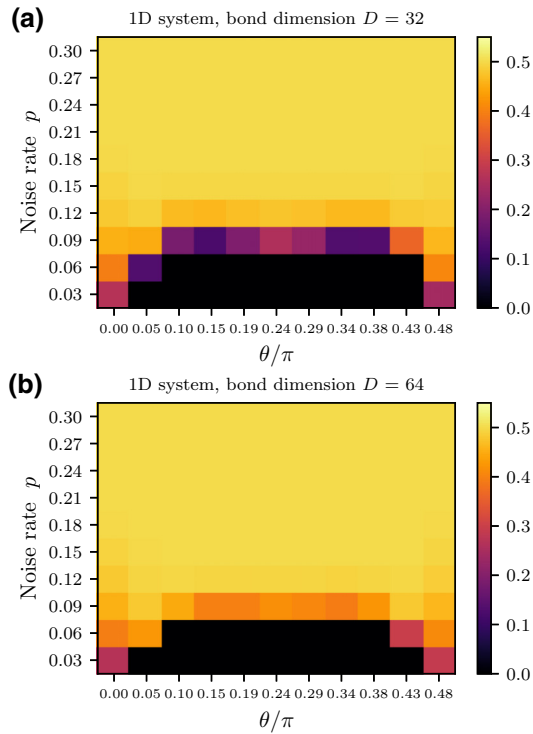


FIG. 5. Trace purity-based dual bounds  $[h(\vec{\sigma}^{h,D})]$  in Eq. (8) for a 1D system of  $N = 32$  spins, the same target Hamiltonian as considered for the results in Fig. 3 (see the description of the Hamiltonian in the caption of Fig. 3), and brick-wall quantum circuits [Fig. 3(a)] where two-qubit unitaries are chosen to be  $\exp(-i\theta X \otimes X)$  and single-qubit unitaries are independently Haar random. Plots show dual bounds as a function of the noise rate  $p$  and circuit parameter  $\theta$  for circuit depth  $d = 25$  and bond dimensions (a)  $D = 32$  and (b)  $D = 64$ . The Hamiltonian is shifted and scaled such that its spectrum is in  $[0, 1]$ .

can be classically contracted using Monte Carlo methods. We leave a detailed theoretical and numerical investigation of these ansatzes for future work.

### C. Nondepolarizing noise models

Up until now, we have modeled the noise present in the circuit as depolarizing. However, noise in several experimental systems might have a more complex structure. In this subsection, we consider these other noise models and show that the duality-based bounding procedure can be reformulated slightly to provide informative bounds even without the assumption of the depolarizing noise. The only assumption that we make that the noise channel under consideration has a full Kraus rank (i.e., the Kraus operators describing the noise channel span the entire space of single-qubit operators). This assumption could be seen as a reasonable physical assumption for sufficiently generic noise models—if the Kraus operators are interpreted effectively as operators that randomly act on the qubit when it is experiencing noise, the Kraus operators not spanning the

full space of linear operators would mean that the noise is special and does *not* apply an entire subspace of operators on the qubit. Nevertheless, for channels that do not have full Kraus rank, the methods presented in this section do not apply and we leave it as an open problem for future work.

First, consider noise channels that are unital and primitive (i.e., have identity as a fixed point). In this case, the noise channel  $\mathcal{N}$  with noise rate  $p$  can always be expressed as

$$\mathcal{N}(\rho) = (1 - p)\rho + V\left(\sum_{P \in \{X, Y, Z\}} p_P P U \rho U^\dagger P\right)V^\dagger \quad (14)$$

for single-qubit unitaries  $U, V$  and  $p_X, p_Y, p_Z \in (0, 1)$  with  $p = p_X + p_Y + p_Z < 1$ . In this case, as detailed in Appendix B, it follows from an application of a corollary of Ref. [62] that Lemma 1 can be extended to this class of channels with the noise rate being chosen as  $\min(p_X, p_Y, p_Z)$ .

*Lemma 2 (Follows from Corollary 5.6 of Ref. [62]).* Suppose that  $\rho_t$  is the quantum state on  $N$  qubits obtained from an initial pure state after applying  $t$  unitaries followed by single-qubit noise channels of the form of Eq. (14); then

$$P_{\text{tr}}(\rho_t) := \text{Tr}(\rho_t^2) \leq 2^{-N\{1 - [1 - \min(p_X, p_Y, p_Z)]^t\}}. \quad (15)$$

The case of nonunital noise channels is more complex, and it is not possible to get architecture-independent bounds on the entropy or the trace purity of the time-dependent state of the quantum circuit. We restrict ourselves to the case where the nonunital noise channel under consideration has a unique fixed point  $\tau$ —this noise channel, then, tends to drive the output of a quantum circuit on  $N$  qubits to the state  $\tau^{\otimes N}$ . Therefore, instead of using a trace purity constraint while formulating the dual, as we have for unital noise channels, we instead use a constraint on  $\|\rho_t - \tau^{\otimes N}\|_F^2$ , where  $\|X\|_F^2 = \text{Tr}(X^\dagger X)$  is the Frobenius norm of  $X$ . To obtain a simple analytical upper bound on  $\|\rho_t - \tau^{\otimes N}\|_F^2$ , we assume that the two-qubit gates in the quantum circuit are diagonal in the basis of eigenvectors of  $\tau$ . This is the case, e.g., if the nonunital noise channel is amplitude damping and the two-qubit gates used in the quantum circuit are all control phase gates. We point out that if prior information about the fixed point of the noise channel is known (e.g., from a previous noise tomography) then a universal gate set can always be chosen such that the two-qubit gates satisfy this requirement.

More concretely, suppose that the unitary in time step  $t$  has single-qubit gates  $V_{t,\alpha}^{(1)}$ . We also assume that the noise channel,  $\mathcal{N}$ , is nonunital and has a full Kraus rank, which we expect to be true for noisy systems if the noise is sufficiently generic. Denoting the fixed point of  $\mathcal{N}$  by  $\tau$ , we

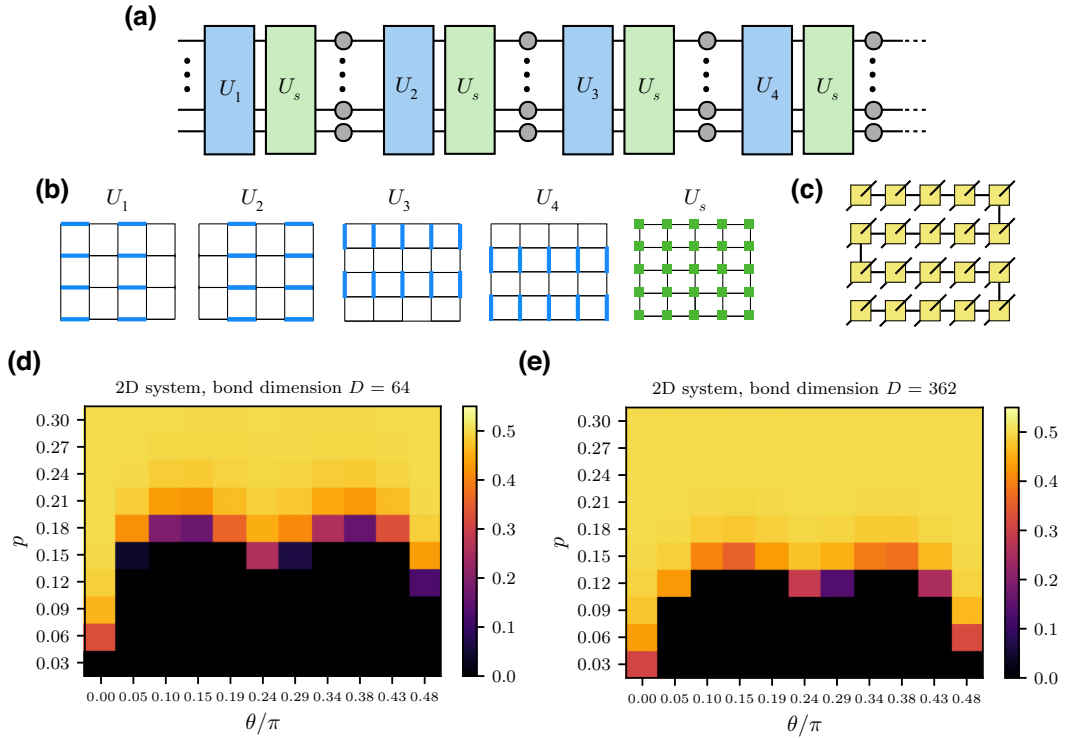


FIG. 6. (a) Schematic of quantum circuits considered for 2D spin systems on a square lattice: colored boxes indicate unitaries and gray circles depolarizing noise. Each unitary layer consists of two-qubit unitaries  $\exp(-i\theta X \otimes X)$  (blue boxes) followed by independently Haar random single-qubit unitaries (green boxes). The first  $d/2$  layers serve to increase the entanglement in the state. The remaining layers invert the action of the previous  $d/2$  such that, in the absence of noise, the output of the circuit is the ground state of  $H = -\sum_{\langle i,j \rangle} Z_i Z_j$ , where  $Z_i$  is the Pauli-Z operator for the  $i$ th spin and  $\langle i,j \rangle$  indicates nearest neighbors. (b) Circuits considered have a brick-wall structure: two-qubit unitary layers cycle between gates on odd horizontal edges ( $U_1$ ), even horizontal edges ( $U_2$ ), odd vertical edges ( $U_3$ ), and even vertical edges ( $U_4$ ). Single-qubit gates ( $U_s$ ) are applied on every qubit after every two-qubit gate layer. (c) Structure of the MPO considered for 2D spin systems: yellow squares indicate tensors at each site in the 2D lattice and lines emerging from them indicate tensor indices. Horizontal lines indicate bond indices with dimension  $D$  and diagonal lines indicate physical indices. (d),(e) Trace purity-based dual bounds  $[h(\bar{\sigma}^{h,D})$  in Eq. (8)] for the ground-state energy of the target Hamiltonian  $H$  as a function of the noise rate  $p$  and circuit parameter  $\theta$  for a system of  $N = 36$  spins arranged in a  $6 \times 6$  lattice, circuit depth  $d = 32$ , and MPO bond dimensions (d)  $D = 64$  and (e)  $D = 362$ . The Hamiltonian is shifted and scaled such that its spectrum is in  $[0, 1]$ .

show in Appendix B that if  $\mathcal{N}$  has a full Kraus rank then

$$\mathcal{N} = \Lambda_{\tau,q} \mathcal{N}' \quad (16)$$

for some channel  $\mathcal{N}'$  that also has  $\tau$  as a fixed point and  $\Lambda_{\tau,q}(X) = (1-q)X + q\text{Tr}(X)\tau$  for  $0 < q < 1$ . The parameter  $q$  can be interpreted as the probability with which the noise channel traces out the qubit and replaces it with  $\tau$ , and it can be computed by solving the semidefinite program

$$\max_q q \quad (17a)$$

$$\text{such that } \Phi_{\mathcal{N}} - qI \otimes \tau \geq 0, \quad (17b)$$

$$q > 0, q < 1, \quad (17c)$$

where  $\Phi_{\mathcal{N}}$  is the Choi state corresponding to  $\mathcal{N}$ . For such noise channels and unitary circuits, we have the following result.

*Lemma 3 (Follows from Lemma 1 of Ref. [24]).* Suppose that  $\rho_t$  is the quantum state on  $N$  qubits obtained from an initial state  $\rho_0$  after applying  $t$  unitaries followed by a single-qubit noise channel  $\mathcal{N}$  of the form of Eq. (16); then it follows that

$$D(\rho_t || \tau^{\otimes N}) \leq D(\rho_0 || \tau^{\otimes N})(1-q)^t + 2 \sum_{s=0}^{t-1} \sum_{\alpha} (1-q)^{t-s} \times \log_2(\|\tau^{-1/2} V_{\alpha,s} \tau V_{\alpha,s}^{\dagger} \tau^{-1/2}\|), \quad (18)$$

where  $D(\rho_1 || \rho_2) = \text{Tr}[\rho_1 \log_2 \rho_1 - \rho_1 \log_2 \rho_2]$  is the quantum relative entropy between  $\rho_1$  and  $\rho_2$ .

We outline a full proof of this lemma in Appendix B. To translate the upper bound on the quantum relative entropy to an upper bound on  $\|\rho_t - \tau^{\otimes N}\|_F$ , we note that

$$\|\rho_t - \tau^{\otimes N}\|_F^2 \leq 2\|\rho_t - \tau^{\otimes N}\|_1 \leq [2D(\rho_t || \tau^{\otimes N})]^{1/2},$$

where we have used the fact that  $\|O\|_2^2 \leq \|O\|_1 \|O\|$  and Pinsker's inequality  $\|\rho_1 - \rho_2\|_1 \leq \sqrt{\frac{1}{2}D(\rho_1||\rho_2)}$ . We remark that our bounds become trivial (i.e.,  $\rightarrow \infty$ ) when  $\tau$  is not full rank, which would also imply by the quantum Perron-Frobenius theorem that the noise channel does not have a full Kraus rank.

Following the same procedure as in Sec. III, we can now formulate the following optimization problem for the energy at the output quantum circuit:

$$\underset{\rho_1, \rho_2, \dots, \rho_d \in \mathcal{S}}{\text{minimize}} \quad \text{Tr}(H\rho_d) \quad (19a)$$

$$\text{subject to} \quad \rho_t = \mathcal{E}_t(\rho_{t-1}), \quad t \in \{1, \dots, d\}, \quad (19b)$$

$$\|\rho_t - \tau^{\otimes N}\|_F^2 \leq d_t^2, \quad t \in \{1, \dots, d\}. \quad (19c)$$

Here, instead of a trace purity bound as in Eq. (5), we use an upper bound on the Frobenius norm distance between  $\rho_t$  and  $\tau^{\otimes N}$  with  $d_t$  given by Lemma 3. Again, introducing the dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d \in \mathcal{M}[(\mathbb{C}^2)^{\otimes N}]$  and  $\lambda_1, \lambda_2, \dots, \lambda_t \geq 0$ , we can construct the Lagrangian

$$\begin{aligned} \mathcal{L}(\vec{\rho}, \vec{\sigma}, \vec{\lambda}) = & \text{Tr}[H\rho_d] + \sum_{t=1}^d \text{Tr}\{\sigma_t[\rho_t - \mathcal{E}_t(\rho_{t-1})]\} \\ & + \sum_{t=1}^d \lambda_t[\|\rho_t - \tau^{\otimes N}\|_F^2 - P_t]. \end{aligned} \quad (20)$$

Minimizing the Lagrangian over  $\vec{\rho}$  and then maximizing it over  $\vec{\lambda}$  yields the dual function

$$h(\vec{\sigma}) = -\text{Tr}(\sigma_1 \mathcal{E}_1(\rho_0)) + \sum_{t=1}^d \left[ \text{Tr}(H_t \tau^{\otimes N}) - d_t \sqrt{\text{Tr}(H_t^2)} \right],$$

where, again,  $H_d = H + \sigma_d$  and  $H_t = \sigma_t - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1})$  for  $t \in \{1, 2, \dots, d-1\}$ . Similar to the case in the previous subsections, due to Lagrangian duality,  $g(\vec{\sigma})$  evaluated at any  $\vec{\sigma}$  provides a lower bound on the energy at the output of the circuit. Following the strategy in the previous sections, we again evaluate  $g(\vec{\sigma})$  at  $\vec{\sigma}^h$  given by a TEBD algorithm in the Heisenberg picture [Eq. (9)]. As an example shown in Fig. 7, the bounds obtained in a 1D circuit on  $N = 30$  qubits, at a noise rate of 3%, with noise modeled by the nonunitary replacement channel  $\mathcal{N}(X) = (1-q)\rho + q\text{Tr}(X)\tau$ , where we assume that  $\tau = I/2 + \epsilon Z$ . Here  $\epsilon$  controls how “nonunitary” the noise channel is— $\epsilon = 0$  corresponds to the previously studied case of a depolarizing noise channel, and in the limit of  $\epsilon \rightarrow \pm 1/2$  we obtain an amplitude damping channel (which does not have a full Kraus rank). We also compute the bounds obtained by considering only the quantum relative entropy of the circuit output ( $\rho_d$ ) with respect to the noise channel fixed point  $\tau^{\otimes N}$  as given by Eq. (18)—to obtain this,

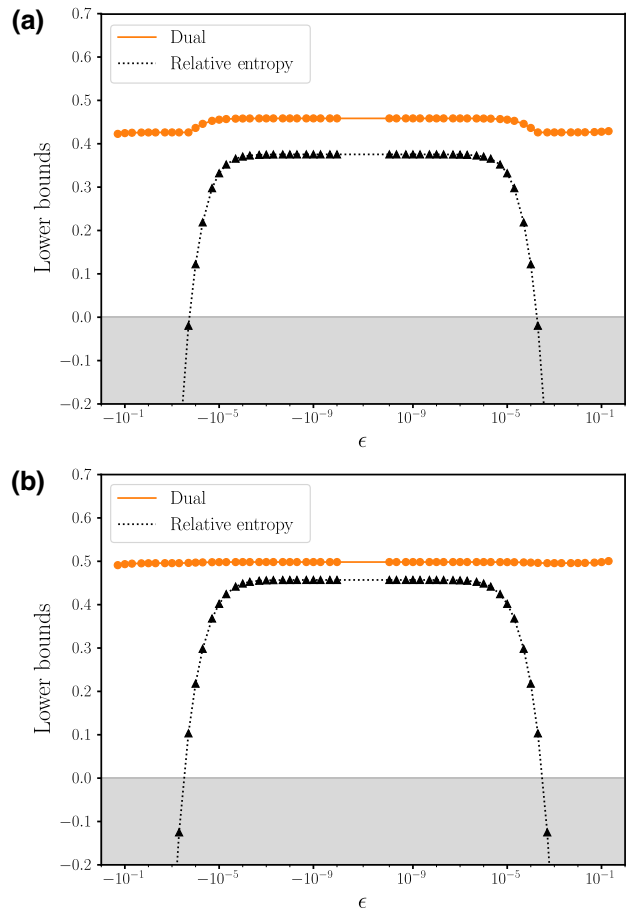


FIG. 7. Bounds on the output energy for the case of a nonunitary noise channel. The structure of the circuits are identical to those considered in Fig. 3(a) with parameters  $N = 32$ ,  $d = 102$ , and  $\theta = 0.1$ . The local noise after each layer of unitaries is given by  $\mathcal{N}(X) = (1-q)X + q\text{Tr}(X)\tau_\epsilon$ , where  $\tau_\epsilon = (1/2 + \epsilon)|0\rangle\langle 0| + (1/2 - \epsilon)|1\rangle\langle 1|$ , with  $q = 0.03$  in (a) and  $q = 0.05$  in (b). In orange, we show the bounds derived on the output energy from our dual problem with an MPO ansatz with  $D = 128$  for the dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d$ . The black lines show the lower bound  $\text{Tr}(H\tau^{\otimes N}) - \|H\| \|\rho_d - \tau^{\otimes N}\|_1$  that is obtained by disregarding the circuit architecture.

we note that, given an upper bound on  $D(\rho_d||\tau^{\otimes N})$ , we can translate it to an upper bound on  $\|\rho_d - \tau^{\otimes N}\|_1$  via Pinsker's inequality, i.e.,

$$\|\rho_d - \tau^{\otimes N}\|_1 \leq \sqrt{\frac{1}{2}D(\rho_d||\tau^{\otimes N})},$$

and therefore the expected energy  $\text{Tr}(H\rho_d)$  at the circuit observable can deviate from  $\text{Tr}(H\tau^{\otimes N})$  by at most  $\|H\| \|\rho_d - \tau^{\otimes N}\|_1$ . Thus we obtain the lower bound

$$\begin{aligned} \text{Tr}(H\rho_d) & \geq \text{Tr}(H\tau^{\otimes N}) - \|H\| \|\rho_d - \tau^{\otimes N}\|_1, \\ & \geq \text{Tr}(H\tau^{\otimes N}) - \|H\| \sqrt{\frac{1}{2}D(\rho_d||\tau^{\otimes N})}. \end{aligned} \quad (21)$$

We find that using the duality-based bound continues to give informative bounds that are significantly better compared to the bounds attained by just accounting for the distance between the noisy output state and the noise channel fixed point  $\tau^{\otimes N}$ . Specifically, as the noise channel becomes increasingly nonunitary then the dual formulation continues to provide nontrivial bounds since it accounts for the circuit architecture, and the bounds attained without accounting for the circuit architecture become trivial for even slightly nonunitary channels.

## V. USING DUALITY WITH INFORMATION CONTENT

In the previous sections, we investigated the impact of noise using the trace purity as a measure of the mixedness in the noisy circuit. The trace purity-based dual function given in Eq. (8) contains terms with Frobenius norms  $\sqrt{\text{Tr}(H_t^2)}$  that, in the worst case, could grow exponentially with the system size  $N$ . Hence, the trace purity-based dual tends to become trivial in the limit of large system size and intermediate circuit depths. An alternative better conditioned purity measure is the information content based on the von Neumann entropy,  $I(\rho) = N - S(\rho) = N + \text{Tr}[\rho \log_2(\rho)]$ . In this section, we formulate a duality-based bound using the information content as a purity measure. However, as we illustrate below, the duality bound here is harder to compute than that based on the trace purity for a general spin model—to still gauge the efficacy of this bound, we numerically study in the simpler but physically relevant case of Gaussian fermions. While our results are suggestive that using the information content-based bounds could be useful for spin models, the associated dual function is harder to compute classically—we leave it as an open problem to develop classical algorithms to compute informative bounds using this strategy for spin models.

Considering the information content, the free energy defined in Eq. (7b) becomes the Gibbs free energy with an offset,

$$\begin{aligned} \mathcal{F}_{S,P}(H, \lambda) &= \inf_{\rho \geq 0, \text{Tr}(\rho)=1} [\text{Tr}[H\rho] + \lambda I(\rho)], \\ &= N\lambda + \inf_{\rho \geq 0, \text{Tr}(\rho)=1} [\text{Tr}(H\rho) - \lambda S(\rho)], \\ &= N\lambda - \lambda \log \text{Tr} \exp(-H/\lambda), \end{aligned} \quad (22)$$

which together with Eq. (7) yields

$$\begin{aligned} \tilde{g}(\vec{\sigma}, \vec{\lambda}) &= -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] + \sum_{t=1}^d [-\lambda_t \log \text{Tr} \exp(-H_t/\lambda_t) \\ &\quad + \lambda_t(N - I_t)], \end{aligned} \quad (23)$$

where  $I_t = N(1 - p)^t$  is the analytical bound on the information content under depolarizing noise defined in Lemma 1.

To benchmark the performance of the information content-based dual, we consider Gaussian fermionic systems where the dual function given in Eq. (23) can be computed exactly. We study  $N$  fermions arranged on a lattice and choose  $H$  to be a quadratic Hamiltonian,

$$H = i \sum_{\alpha, \alpha', x, x'} h_{x, x'}^{\alpha, \alpha'} c_x^\alpha c_{x'}^{\alpha'},$$

where  $c_x^\alpha, c_x^{\alpha'}$  are the Majorana operators for the fermion at point  $x$  on the lattice, and the  $h_{x, x'}^{\alpha, \alpha'}$  are real numbers specifying  $H$ . We additionally assume the unitaries in the circuit that prepares the ground state of  $H$  from an initial vacuum state to be Gaussian unitaries.

Since both Gaussian unitaries and the depolarizing channel map a quadratic Hermitian operator to another quadratic Hermitian operator, Proposition 1 indicates that the dual function is maximized for  $\sigma_t$ , which themselves are quadratic Hermitian operators. This motivates the following ansatz for  $\sigma_t$ :

$$\sigma_t = i \sum_{\substack{\alpha, \alpha', x, x' \\ d(x, x') \leq r}} s_{x, x', t}^{\alpha, \alpha'} c_x^\alpha c_{x'}^{\alpha'}$$

for real  $s_{x, x', t}^{\alpha, \alpha'}$ . In our study, we restrict  $\sigma_t$  to be local operators with interaction range  $r$  while maximizing  $g(\vec{\sigma}, \vec{\lambda})$  to obtain the lower bound—when  $r \sim$  lattice size, we expect to obtain the best possible lower bound, but, since the ansatz always includes the point  $\vec{\sigma} = 0$ , we expect from Proposition 2 to obtain a bound better than that predicted by only considering the information content of the output state, even for small  $r$ . Choosing  $\sigma_t$  to be quadratic Hermitian operators allows for exact, classically efficient computation of the Gibbs free-energy terms in the dual function and, furthermore, even the circuit output can be computed exactly by considering the covariance matrix describing the state; see Appendix B for details. We obtain bounds by maximizing  $g(\vec{\sigma}, \vec{\lambda})$  through a gradient-based local optimization algorithm, starting from the initial point where the  $\sigma_t$  are chosen to be the Heisenberg picture evolution of  $-H$ , but projected onto the space of quadratic fermionic Hamiltonians with interaction range  $r$  after each time step—much like the compression into MPOs of bond dimension  $D$  in Eq. (9).

Figure 8 shows a numerical study of the bounds that we obtain—we consider systems with  $\sim 50$  fermions arranged both on 1D [Fig. 8(b)] and 2D lattices [Fig. 8(c)] and experiencing depolarizing noise at a rate of 5%. We choose  $H$  to be an Su–Schrieffer–Heeger (SSH) model, nearest-neighbor Hamiltonian with alternating hopping strengths [Fig. 8(a)]. For the 1D benchmarks, we choose

$$H = \sum_x (v_{a,b} a_x^\dagger b_x + v_{b,a} b_x^\dagger a_{x+1} + \text{H.c.}), \quad (24a)$$

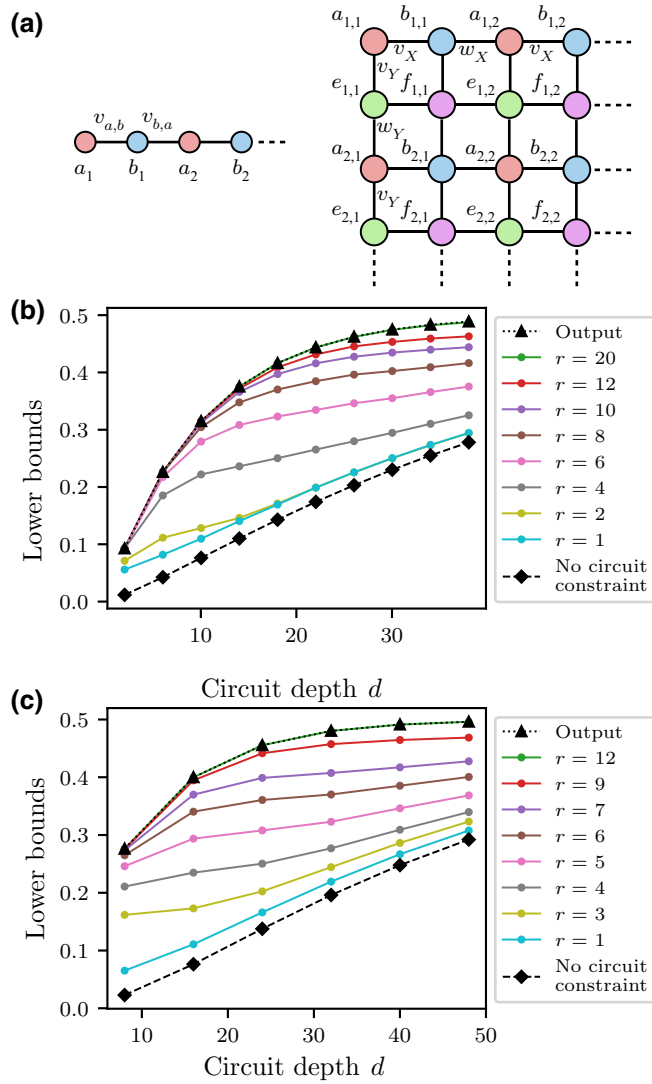


FIG. 8. (a) Schematic of SSH model quadratic fermionic Hamiltonians with alternating hopping strengths in one and two dimensions. (b),(c) Comparison of information content–based dual bounds with and without circuit constraints, and the output energies of noisy Gaussian circuits for systems consisting of (b)  $N = 48$  fermions arranged in a 1D lattice, (c)  $N = 49$  fermions arranged in a  $7 \times 7$  2D lattice. Dual bounds are shown for ansatzes with varying interaction range  $r$ . The horizontal axis represents the depth  $d$  of a Gaussian brick-wall circuit that outputs the ground state of the SSH model. Fermions are independently subject to depolarizing noise with probability  $p = 5\%$  after every unitary layer. The Hamiltonians are shifted and scaled such that their spectrum is in  $[0, 1]$ .

and for 2D benchmarks, we choose

$$\begin{aligned}
 H = & \sum_{x,y} \left( \sum_{p,q \in \{(a,b),(e,f)\}} (v_{p,q} p_{x,y}^\dagger q_{x,y} + v_{q,p} q_{x,y}^\dagger p_{x+1,y} + \text{H.c.}) \right. \\
 & \left. + \sum_{p,q \in \{(a,e),(b,f)\}} (v_{p,q} p_{x,y}^\dagger q_{x,y} + v_{q,p} q_{x,y}^\dagger p_{x,y+1} + \text{H.c.}) \right), \quad (24b)
 \end{aligned}$$

where we choose  $v_{a,b} = v_{ef} = v_x, v_{b,a} = v_{f,e} = w_x, v_{a,e} = v_{b,f} = v_y$ , and  $v_{e,a} = v_{f,b} = w_y$ . For the numerical studies shown in Fig. 8,  $H$  is shifted and scaled such that the ground- and highest-excited-state energies are zero and one, respectively. We consider circuits of depth  $d$  consisting of two-mode Gaussian unitaries arranged in a brick-wall layout, where the first  $d/2$  layers are composed of randomly generated two-mode Gaussian unitaries that serve to increase the entanglement in the state. The remaining layers invert the action of the previous  $d/2$  such that, in the absence of noise, the output state is the initial state, which is chosen to be the ground state of  $H$ . In Fig. 8, for comparison, we also include the exact output of the noisy Gaussian circuit, as well as the bound obtained by neglecting the circuit constraints and only considering the information content of the output state. As expected, we find that on accounting for the circuit constraint, we obtain bounds that are more representative of the output. We also observe that the dual bounds get closer to the output as the dual ansatz’s interaction range  $r$  increases, since the ansatz becomes more expressive.

## VI. CONCLUSION AND OUTLOOK

In conclusion, we demonstrate a method to rigorously lower bound the performance of any given quantum circuit subject to a constant rate of depolarizing noise. We achieve this by constructing a Lagrangian dual specific to the circuit, which takes into consideration not only the decreasing purity of the state through the circuit due to noise, but also the details of the gates in the circuit, allowing the study of the effect of entanglement generation in the circuit that can worsen the detrimental effects of noise. We presented numerical studies in spin systems and showed that it is possible to efficiently calculate circuit-specific lower bounds that are tighter than bounds obtained by just considering the information content of the output state. We provided an interpretation of the trace purity–based dual evaluated at dual variables obtained from TEBD in the Heisenberg picture in terms of the compression errors. We also showed that the dual can be formulated in terms of the information content of the state instead of the trace purity—we computed information content–based circuit dual bounds for Gaussian fermionic systems where the Gibbs free energy can be computed exactly.

Our method opens the door to promising avenues of future research. We already demonstrate numerically that this approach can provide informative lower bounds even when the TEBD simulation is inaccurate—e.g., in Figs. 4(a) and 4(b), it can be seen that the obtained lower bound changes with MPO bond dimension, indicating that the bond dimension is not large enough to accurately simulate the circuit dynamics, but still continues to provide nontrivial energy lower bounds. We expect that we are able to handle circuits with moderate entanglement using

our strategy, but, indeed, we expect to obtain trivial lower bounds for highly entangled circuits (e.g., circuits performing fault-tolerant quantum computation). Moving forward, larger-scale tensor network numerics can allow us to also provide energy lower bounds in higher-dimensional circuits, with a number of qubits reaching state-of-the-art experiments even if the tensor network methods are not good enough for accurately simulating circuit dynamics. Specifically, for 2D circuits, using projected entangled-pair operators (or their subclasses) as the ansatz for the dual variables is an important avenue for future exploration. Furthermore, better numerical optimization algorithms can be explored for improving the lower bounds than those from evaluating the dual function at a specific point. Moreover, extensions of the methods of this paper to continuous time would better capture the experimental system, and even allow us to apply this method for understanding quantum adiabatic algorithms [4]. Finally, the Lagrangian dual formulation of lower bounds, apart from being a numerical tool, could also shed rigorous theoretical insights into understanding the resilience of quantum circuit architectures to noise.

### ACKNOWLEDGMENTS

We thank J. Ignacio Cirac, Mari Carmen Bañuls, and Guillermo González-García for helpful discussions. R.T. acknowledges a Max Planck Harvard Research Center for Quantum Optics (MPHQ) Postdoctoral Fellowship.

### APPENDIX A: PROOF OF PROPOSITION 4

We first recall the main result of González-García *et al.* [26] in which error propagation in a family of random quantum circuits was analyzed. Specifically, they considered brick-wall quantum circuits of depth  $d$  (assumed to be even) with unitaries  $U_1, U_2, \dots, U_d$ , where

$$U_{d/2+1} = U_{d/2}^\dagger, \quad U_{d/2+2} = U_{d/2-1}^\dagger, \\ U_{d/2+3} = U_{d/2-2}^\dagger, \quad \dots, \quad U_d = U_1^\dagger.$$

The unitaries  $U_1, U_2, \dots, U_{d/2}$  are chosen randomly depending on the circuit architecture. We specifically consider the 1D case, where  $U_1, U_3, U_5, \dots$  are formed by applying random two-qubit gates between qubits (1, 2), (3, 4), (5, 6),  $\dots$  and  $U_2, U_4, U_6, \dots$  are formed by applying random two-qubit gates between qubits (2, 3), (4, 5), (6, 7),  $\dots$ . All the two qubit gates are chosen independently at random from an ensemble that forms a 2-design. Furthermore, we consider the noisy setting where depolarizing noise with probability  $p$  is applied to each qubit after every unitary layer. González-García *et al.* [26] established the following result characterizing the average energy of the output state for a 2-local Hamiltonian. While the result of Ref. [26] holds for arbitrary

2-local Hamiltonians, we only consider the Hamiltonian  $H = -\sum_{i=1}^N Z_i$ .

*Lemma 4 (Ref. [26]).* The expectation value  $E$  of the Hamiltonian  $H = -\sum_{i=1}^N Z_i$  with respect to the output state of a circuit chosen randomly from the ensemble described above satisfies

$$\text{Prob}[|E + N(1-p)^{\Omega(d^2)}| \leq \alpha_0 \sqrt{N}] \geq 1 - 2e^{-2\alpha_0^2/2d^2}.$$

*Proof of Proposition 4.* Consider the Hamiltonian  $H = -\sum_{i=1}^N Z_i$ , and choose the two-qubit gates to be Haar-random Clifford gates. Since Haar-random Clifford gates form a 2-design [63], we can use Lemma 4. Hamiltonian  $H$  satisfies  $\text{Tr}(H) = 0$  and has operator norm  $\|H\| = N$ . We evaluate the dual function at the dual variables obtained by Heisenberg picture evolution of  $-H$ ,

$$\sigma_d = -H, \quad \sigma_t = -\mathcal{E}_{t+1}^\dagger \mathcal{E}_{t+2}^\dagger \cdots \mathcal{E}_d^\dagger(H).$$

Note that, for Clifford circuits,  $Z_i$  will be mapped to a single Pauli string [47], which is expressible as an MPO with bond dimension 1, and, consequently,  $\sigma_t$  will have a bond dimension of at most  $N$ . Since the Heisenberg picture evolution can be captured exactly with an MPO ansatz of bond dimension  $O(N)$ , the dual bound  $\ell_D^{\text{dual}}$  with  $D \leq O(N)$  is exactly equal to the energy  $E$  of the output of the noisy quantum circuit. Now, from Lemma 4, it follows that there must exist at least one 1D circuit such that

$$\ell_D^{\text{dual}} = -N(1-p)^{\Omega(d^2)} + O(\sqrt{N})$$

with  $D \leq O(N)$ .

We now consider the lower bound  $\ell^I$  obtained by only considering the information content of the state at the output of the circuit and neglecting the circuit constraints:

$$\ell^I = \underset{\rho: I(\rho) \leq N(1-p)^d}{\text{minimize}} \text{Tr}(H\rho).$$

A negative upper bound on  $\ell^I$  that converges to  $\text{Tr}(H/2^N) = 0$  exponentially with  $d$  can be obtained by computing  $\text{Tr}(H\rho)$  at

$$\rho = [p_d|0\rangle\langle 0| + (1-p_d)I/2]^{\otimes N}, \quad \text{where } p_d = (1-p)^d.$$

Note that this  $\rho$  satisfies  $I(\rho) = N(1-p)^d$ , and, consequently,  $\text{Tr}(H\rho)$  is an upper bound on  $\ell^I$ . For  $H = \sum_{i=1}^N Z_i$ , we then obtain  $\text{Tr}(H\rho) = -N(1-p)^d$ , which implies that  $\ell^I \leq -N(1-p)^{O(d)}$ . ■

### APPENDIX B: NONDEPOLARIZING NOISE CHANNELS

In this appendix, we obtain bounds on the trace purity in the presence of nondepolarizing noise channels. The

bounds presented here are already contained in or can be straightforwardly obtained from existing results (e.g., in Refs. [24,25,62]). We include this appendix for a self-contained derivation of the results used in the main text.

### 1. Unital noise channels

We first consider unital noise channels that are also primitive. It is a standard result from the characterization of qubit channels that such a noise channel can be expressed as

$$\mathcal{N}(X) = (1-p)X + \sum_{a \in \{x,y,z\}} p_a \mathcal{U}[A\mathcal{V}(X)A], \quad (\text{B1})$$

where  $p = p_x + p_y + p_z$  can be considered to be the noise rate, and  $\mathcal{V}(\cdot) = V \cdot V^\dagger$ ,  $\mathcal{U}(\cdot) = U \cdot U^\dagger$  are unitary channels with  $V, U$  being unitaries. Furthermore, if  $\mathcal{N}$  is primitive (i.e., has  $I/2$  as the unique fixed point) then  $p_x, p_y, p_z > 0$ . To obtain a bound on the trace purity of the qubit's state in a quantum circuit impacted by such a noise, we need the following lemma from Ref. [62]. Here,  $\|\cdot\|_{2 \rightarrow 2}$  of a superoperator refers to the Schatten-2 norm, i.e.,

$$\|\mathcal{E}\|_{2 \rightarrow 2}^2 = \sup_{X \neq 0} \frac{\text{Tr}[\mathcal{E}(X)^\dagger \mathcal{E}(X)]}{\text{Tr}(X^\dagger X)}.$$

Furthermore, for  $\sigma \succ 0$ ,  $D_2(\rho||\sigma)$  is the sandwiched 2-Rényi divergence that is given by

$$D_2(\rho||\sigma) = \log_2 \text{Tr}(\sigma^{-1/2} \rho \sigma^{-1/2} \rho).$$

In particular, note that if  $\rho$  is an  $N$ -qubit density matrix and  $\sigma = (I/2)^{\otimes N}$ , we obtain

$$D_2(\rho||\sigma) = N + \log \text{Tr}(\rho^2).$$

*Lemma 5 (Corollary 5.6 of Ref. [62]).* Suppose that  $\Lambda_{\sigma,p}$  is the channel described by  $\Lambda_{\sigma,p}(X) = (1-q)X + q\text{Tr}(X)\sigma$ . For  $\sigma \succ 0$ , define  $\Gamma_\sigma$  to be the superoperator  $\Gamma_\sigma(X) = \sigma X \sigma^{-1}$ . Suppose that  $\mathcal{N}$  is a channel with  $\sigma$  as its fixed point; then if  $\|\Gamma_\sigma^{-1/2} \mathcal{N} \Lambda_{\sigma,p}^{-1} \Gamma_\sigma^{1/2}\|_{2 \rightarrow 2} \leq 1$ , for any  $N > 0$  and  $N$ -qubit density matrix  $\rho$ ,

$$D_2[\mathcal{N}^{\otimes N}(\rho)||\sigma^{\otimes N}] \leq \alpha D_2(\rho||\sigma^{\otimes N}),$$

where  $\alpha = 2^{2(1-\|\sigma^{-1}\|^{-1}) \log_2(1-p)/\log_2(\|\sigma^{-1}\|)}$ .

*Proof of Lemma 2.* We pick  $\mathcal{N}$  to be the unital noise channel in Eq. (B1), and  $\sigma = I/2$ . With this choice, we have  $\Lambda_{I/2,q}(X) = (1-q)X + q\text{Tr}(X)I/2$  and  $\Gamma_\sigma(X) = 4X$ . Note that  $\|\Gamma_{I/2}^{-1/2} \mathcal{N} \Lambda_{\sigma,p}^{-1} \Gamma_{I/2}^{1/2}\|_{2 \rightarrow 2} = \|\mathcal{N} \Lambda_{I/2,p}^{-1}\|_{2 \rightarrow 2}$  and that  $\Lambda_{I/2,p}^{-1}(X) = (1-p)^{-1}[X - p\text{Tr}(X)I/2]$ . Further analysis is simplified in the Pauli basis—since both  $\mathcal{N}$  and

$\Lambda_{I/2,q}^{-1}$  are unital superoperators, it follows that, if written as  $4 \times 4$  matrices in the Pauli basis, they have the forms

$$\Lambda_{I/2,q}^{-1} \cong \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\Lambda}_{I/2,q}^{-1} \end{bmatrix} \quad \text{and} \quad \mathcal{N} \cong \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\mathcal{N}} \end{bmatrix}.$$

Furthermore we can explicitly calculate  $\tilde{\Lambda}_{I/2,q}^{-1}$  and  $\tilde{\mathcal{N}}$  to obtain

$$\begin{aligned} \tilde{\Lambda}_{I/2,q}^{-1} &= \frac{1}{1-q}I \quad \text{and} \quad \tilde{\mathcal{N}} = (1-p)I + \tilde{U} \\ &\times \begin{bmatrix} p_x - p_y - p_z & 0 & 0 \\ 0 & p_y - p_x - p_z & 0 \\ 0 & 0 & p_z - p_x - p_y \end{bmatrix} \tilde{V}, \end{aligned}$$

where  $\tilde{U}$  is a  $3 \times 3$  unitary matrix with matrix elements given by  $\tilde{U}_{a,a'} = \text{Tr}(AUA'U^\dagger)$  for  $A, A' \in \{X, Y, Z\}$ , and  $\tilde{V}$  is defined similarly. Now,  $\|\mathcal{N} \Lambda_{I/2,q}^{-1}\|_{2 \rightarrow 2} = \max(1, \|\tilde{\Lambda}_{I/2,q}^{-1} \tilde{\mathcal{N}}\|)$  and

$$\begin{aligned} \|\tilde{\Lambda}_{I/2,q}^{-1} \tilde{\mathcal{N}}\| &= \frac{1}{1-q} \|\tilde{\mathcal{N}}\| \leq \frac{1}{1-q} [1 - p + \max(p_x - p_y \\ &\quad - p_z, p_y - p_x - p_z, p_z - p_x - p_y)]. \end{aligned}$$

Now, we clearly have  $\max(p_x - p_y - p_z, p_y - p_x - p_z, p_z - p_x - p_y) \leq p - \min(p_x, p_y, p_z)$  and thus  $\|\tilde{\Lambda}_{I/2,q}^{-1} \tilde{\mathcal{N}}\| \leq [1 - \min(p_x, p_y, p_z)]/(1-q)$ . Thus, choosing  $q = \min(p_x, p_y, p_z)$  yields  $\|\Gamma_{I/2}^{-1/2} \mathcal{N} \Lambda_{\sigma,p}^{-1} \Gamma_{I/2}^{1/2}\|_{2 \rightarrow 2} = \|\mathcal{N} \Lambda_{I/2,q}^{-1}\|_{2 \rightarrow 2} \leq 1$ ; thus, we can now apply Lemma 5 with  $\sigma = I/2$ ,  $q = \min(p_x, p_y, p_z)$ , which yields  $\alpha = 1 - \min(p_x, p_y, p_z)$ . In particular, if  $\rho$  is an  $N$ -qubit density matrix, we obtain, from Lemma 5,

$$D_2\left(\mathcal{N}^{\otimes N}(\rho) \left\| \frac{I^{\otimes N}}{2^N} \right.\right) \leq [1 - \min(p_x, p_y, p_z)] D_2\left(\rho \left\| \frac{I^{\otimes N}}{2^N} \right.\right).$$

Now consider the setting where starting from a pure state  $\rho_0$ , a sequence of  $N$ -qubit unitaries  $U_1, U_2, \dots, U_d$  is applied interspersed with the noise channel  $\mathcal{N}$  acting on each qubit. Since  $D_2[U_i \rho U_i^\dagger || (I/2)^{\otimes N}] = D_2[\rho || (I/2)^{\otimes N}]$ , the final state  $\rho_d = \mathcal{N}^{\otimes N} U_d \mathcal{N}^{\otimes N} U_{d-1} \dots \mathcal{N}^{\otimes N} U_1(\rho_0)$  satisfies

$$\begin{aligned} D_2\left(\rho_d \left\| \frac{I^{\otimes N}}{2^N} \right.\right) &\leq [1 - \min(p_x, p_y, p_z)]^d D_2\left(\rho_0 \left\| \frac{I^{\otimes N}}{2^N} \right.\right) \\ &= N[1 - \min(p_x, p_y, p_z)]^d. \end{aligned}$$

This completes the proof. ■



## 2. Nonunital noise channels

Next, we consider nonunital noise channels on  $\mathbb{C}^d$  (where we are typically interested in  $d = 2$ )—we restrict ourselves to noise channels that have a full Kraus rank, i.e., the Kraus operators  $K_1, K_2, \dots, K_{d^2}$  span the entire space of operators on  $\mathbb{C}^d$ . Given a noise rate  $p$ , we assume that the noise channel  $\mathcal{N}$  is given by

$$\mathcal{N}(X) = (1-p)X + p \sum_{i=1}^{d^2} K_i X K_i^\dagger, \quad (\text{B2})$$

where  $\sum_{i=1}^{d^2} K_i^\dagger K_i = I$ . A common example of such a channel would be  $\mathcal{N}(X) = (1-p)X + p \text{Tr}(X)\tau$ , where  $\tau > 0$ , i.e., a channel that traces the qudit and replaces it with a, generally nonidentity, full rank state. Physically, this would be a good model for an environment that disentangles the qubits in the quantum circuit, and brings them to a finite temperature state.

For channels of the form of Eq. (B2), the following lemma straightforwardly follows.

*Lemma 6.* Suppose that  $\mathcal{N}$  is a channel of the form given in Eq. (B2) with unique fixed point  $\tau > 0$ , and that there exists another channel  $\mathcal{N}'$  with fixed point  $\tau$  such that  $\mathcal{N}(X) = \Lambda_{\tau,q} \mathcal{N}'(X)$ , where  $\Lambda_{\tau,q}(X) = (1-q)X + q \text{Tr}(X)\tau$  for some  $q \in (0, 1)$ . Furthermore, the largest such  $q$  can be computed by solving the semidefinite program

$$\max_q \quad q \quad (\text{B3a})$$

$$\text{such that } \Phi_{\mathcal{N}} - qI \otimes \tau \geq 0, \quad (\text{B3b})$$

$$q > 0, \quad q < 1, \quad (\text{B3c})$$

where  $\Phi_{\mathcal{N}}$  is the Choi state of  $\mathcal{N}$ .

*Proof.* Since  $\mathcal{N}$  has a full Kraus rank, we can find linearly independent operators  $L_1, L_2, \dots, L_{d^2}$  such that  $\sum_{i=1}^{d^2} L_i^\dagger L_i = I$  and  $\mathcal{N}(X) = \sum_{i=1}^{d^2} L_i X L_i^\dagger$ . Furthermore, since  $L_1, L_2, \dots, L_{d^2}$  are linearly independent, it follows that, for any  $M \in \mathbb{C}^{d \times d}$ ,  $\text{Tr}(L_i^\dagger M) = 0$  for all  $i \in \{1, 2, \dots, d^2\}$  implies that  $M = 0$ . Consequently, there exists  $\lambda_0 > 0$  such that

$$\sum_{i=1}^{d^2} |\text{Tr}(L_i^\dagger M)|^2 \geq \lambda_0 \text{Tr}(M^\dagger M) \quad \text{for all } M \in \mathbb{C}^{d \times d}. \quad (\text{B4})$$

Now, suppose that  $q \in (0, 1)$  and that  $\mathcal{N}' = [\mathcal{N} - q \text{Tr}(\cdot)\tau]/(1-q)$ . It is clear that  $\mathcal{N}'$  is trace preserving. We need to establish that  $\mathcal{N}'$  is also completely positive. For

that, consider the Choi state corresponding to  $\mathcal{N}'$ ,  $\Phi_{\mathcal{N}'}$ :

$$\begin{aligned} \Phi_{\mathcal{N}'} &= (\mathcal{N}' \otimes \text{id})(|\Phi\rangle\langle\Phi|), \\ &= \frac{1}{1-q} \left( \sum_{i=1}^{d^2} \sum_{j,k=1}^d L_i |j\rangle\langle k| L_i^\dagger \otimes |j\rangle\langle k| - qI \otimes \tau \right), \\ &= \frac{1}{1-q} \left( \sum_{i=1}^{d^2} \sum_{j,k,j',k'=1}^d (L_i)_{j',j} (L_i)_{k',k}^* |j'\rangle\langle k'| \right. \\ &\quad \left. \otimes |j\rangle\langle k| - qI \otimes \tau \right). \end{aligned}$$

For  $\mathcal{N}'$  to be completely positive, it is necessary and sufficient for  $\Phi_{\mathcal{N}'} \geq 0$ . To impose this condition, consider a state  $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ ; then

$$\begin{aligned} \langle\psi|\Phi_{\mathcal{N}'}|\psi\rangle &= \frac{1}{1-q} \left( \sum_{i=1}^{d^2} \sum_{j,k,j',k'=1}^d (L_i)_{j',j} (L_i)_{k',k}^* \psi_{j',j}^* \psi_{k',k} \right. \\ &\quad \left. - q \langle\psi|I \otimes \tau|\psi\rangle \right), \\ &= \frac{1}{1-q} \left( \sum_{i=1}^{d^2} |\text{Tr}(L_i^\dagger \Psi)|^2 - q \langle\psi|I \otimes \tau|\psi\rangle \right), \end{aligned}$$

where  $\Psi \in \mathbb{C}^{d \times d}$  is state  $|\psi\rangle$  reshaped as a matrix. Using Eq. (B4), we obtain

$$\begin{aligned} \langle\psi|\Phi_{\mathcal{N}'}|\psi\rangle &\geq \frac{1}{1-q} [\text{Tr}(\Psi^\dagger \Psi) \lambda_0 - q \|\tau\| \|\psi\|^2] \\ &= \frac{\lambda_0 - q \|\tau\|}{1-q} \|\psi\|^2. \end{aligned}$$

Thus, for  $q = \lambda_0 / \|\tau\|$ , we obtain  $\Phi_{\mathcal{N}'} \geq 0$  and hence  $\mathcal{N}'$  is completely positive. This establishes that there exists  $q > 0$  such that  $\mathcal{N} = (1-q)\mathcal{N}' + q \text{Tr}(\cdot)\tau$ . Furthermore, it trivially follows that  $\mathcal{N}'(\tau) = \tau$  from this definition of  $\mathcal{N}'$ . It therefore also follows that  $\mathcal{N} = \Lambda_{\tau,q} \mathcal{N}'$ . The optimization problem written for the calculation of  $q$  in Eq. (17) is simply a reformulation of the condition that the Choi state of  $\mathcal{N}'$  needs to be positive semidefinite. ■

Finally, we now consider the setting of  $N$  qubits that have unitaries  $U_1, U_2, \dots, U_d$  applied to them, along with noise  $\mathcal{N}$  acting on each qubit at every time step. As described in the main text, we assume that each unitary  $U_i$  is composed of two-qubit and single-qubit gates, with the gate set being chosen such that the two-qubit gates leave  $\sigma^{\otimes 2}$  invariant (i.e., the two-qubit gate is diagonal on the basis of eigenvectors of  $\sigma$ ).

*Proof of Lemma 3.* Suppose that  $\rho$  is an  $N$ -qubit density matrix, and consider the state  $\mathcal{N}^{\otimes N}(\rho)$ . We first establish that

$$D(\mathcal{N}^{\otimes N}(\rho)||\tau^{\otimes N}) \leq (1-q)D(\rho||\tau^{\otimes N}). \quad (\text{B5})$$

Let  $\mathcal{N} = \Lambda_{\tau,q}\mathcal{N}'$  and  $\mathcal{N}^{\otimes N} = \Lambda_{\tau,q}^{\otimes N}\mathcal{N}'^{\otimes N}$ . Suppose that  $\rho' = \mathcal{N}'^{\otimes N}(\rho)$  and then  $D[\mathcal{N}^{\otimes N}(\rho)||\tau^{\otimes N}] = D[\Lambda_{\tau,q}^{\otimes N}(\rho')||\tau^{\otimes N}]$ . We point out that  $\Lambda_{\tau,q}$  can be viewed as generated from the Lindbladian  $\mathcal{L}_\tau = \text{id} - \text{Tr}(\cdot)\tau$  with evolution time  $t = -\log(1-q)$ . It now follows from Theorem 19 of Ref. [64] that if  $\tilde{\mathcal{L}}_\tau := \sum_{i=1}^N \mathcal{L}_\tau^{(i)}$ , where  $\mathcal{L}_\tau^{(i)}$  is the Lindbladian  $\mathcal{L}_\tau$  acting on the  $i$ th qubit, satisfies  $D[e^{\tilde{\mathcal{L}}_\tau t}(\rho')||\tau^{\otimes N}] \leq e^{-t}D[\rho'||\tau^{\otimes N}]$ . Setting  $t = -\log(1-q)$  in this inequality, we obtain  $D[\Lambda_{\tau,q}^{\otimes N}(\rho')||\tau^{\otimes N}] \leq (1-q)D(\rho'||\tau^{\otimes N})$ . Now, this yields

$$\begin{aligned} D[\mathcal{N}^{\otimes N}(\rho)||\tau^{\otimes N}] &\leq (1-q)D[\mathcal{N}'^{\otimes N}(\rho)||\tau^{\otimes N}] \\ &= (1-q)D[\mathcal{N}'^{\otimes N}(\rho)||\mathcal{N}'(\tau^{\otimes N})] \\ &\leq (1-q)D(\rho||\tau^{\otimes N}). \end{aligned}$$

Next, we consider one of the unitaries applied in the circuit,  $U_t$ . The two-qubit gates in this unitary, by assumption, leave  $\tau^{\otimes 2}$  unchanged. Thus, if  $V^{(2)}$  is a two-qubit unitary that satisfies this condition, we have

$$D(V^{(2)}\rho V^{(2)\dagger}||\tau^{\otimes N}) = D(\rho||V^{(2)\dagger}\tau^{\otimes N}V^{(2)}) = D(\rho||\tau^{\otimes N}).$$

Furthermore, if  $V^{(1)}$  is a single-qubit gate, from the data processed triangle inequality we obtain

$$D(V^{(1)}\rho V^{(1)\dagger}||\tau^{\otimes N}) \leq D(\rho||\tau^{\otimes N}) + D_\infty(V^{(1)}\tau V^{(1)\dagger}||\tau).$$

Suppose that the single qubits applied in  $U_t$  are  $V_{t,\alpha}^{(1)}$  for  $\alpha \in \{1, 2, 3, \dots\}$ ; then from the above two inequalities we obtain

$$D(U_t \rho U_t^\dagger || \tau^{\otimes N}) \leq D(\rho || \tau^{\otimes N}) + \sum_{\alpha} D_{\infty}(V_{t,\alpha}^{(1)} \tau V_{t,\alpha}^{(1)\dagger} || \tau). \quad (\text{B6})$$

Recurring Eqs. (B5) and (14), it follows that  $\rho_t = U_t U_{t-1} \dots U_1 \rho_0 U_1^\dagger \dots U_{t-1}^\dagger U_t^\dagger$  satisfies

$$\begin{aligned} D(\rho_t || \tau^{\otimes N}) &\leq (1-q)^t D(\rho_0 || \tau^{\otimes N}) \\ &\quad + \sum_{s=0}^{t-1} \sum_{\alpha} (1-q)^{t-s} D_{\infty}(V_{s+1,\alpha}^{(1)} \tau V_{s+1,\alpha}^{(1)\dagger} || \tau). \end{aligned}$$

Finally, this bound can be translated to a bound on  $\|\rho_t - \sigma^{\otimes N}\|_F$  by a standard application of Pinsker's inequality:

$$\|\rho_t - \tau^{\otimes N}\|_F^2 \leq 2\|\rho_t - \tau^{\otimes N}\|_1 \leq \sqrt{2D(\rho_t||\tau^{\otimes N})}.$$

This completes the proof.  $\blacksquare$

## APPENDIX C: FREE FERMION FORMULATION

In this appendix, we describe our approach to numerical studies of Gaussian fermionic systems. We describe the fermionic system model and how the energy of the state at the output of the noisy circuit and the information content-based dual function (23) can be efficiently calculated by utilizing covariance matrices and quadratic Hamiltonian transformations.

### 1. Model

We consider  $N$  fermions arranged in a lattice and define for each lattice site  $x$  the Majorana operators  $c_x^1, c_x^2$  as satisfying the anticommutation relations

$$\begin{aligned} \{c_x^1, c_{x'}^1\} &= \{c_x^2, c_{x'}^2\} = \delta_{xx'}, \\ \{c_x^1, c_{x'}^2\} &= 0 \quad \text{for all } x, x' \in \text{lattice}. \end{aligned} \quad (\text{C1})$$

We choose the target Hamiltonian  $H$  to be quadratic,

$$H = i \sum_{\alpha, \alpha', x, x'} h_{x, x'}^{\alpha, \alpha'} c_x^\alpha c_{x'}^{\alpha'}.$$

Defining the operator-valued vector

$$\vec{c} = (c_1^1, c_2^1, \dots, c_N^1, c_1^2, c_2^2, \dots, c_N^2)^T, \quad (\text{C2})$$

the target Hamiltonian can be written as  $H = i\vec{c}^T h \vec{c}$ , where matrix  $h$  specifying Hamiltonian  $H$  is real and antisymmetric. The ground-state energy of the target Hamiltonian can be computed by diagonalization of its compact representation  $h$  [65].

### 2. Calculation of the noisy circuit output

The output energy, with respect to the quadratic target Hamiltonian  $H$ , can be computed from the covariance matrix of the state at the end of the circuit. We define the covariance matrix of a state  $\rho$  as

$$\gamma_{x, x'}^{\alpha, \alpha'} = i\text{Tr}(\rho [c_x^\alpha, c_{x'}^{\alpha'}]), \quad (\text{C3})$$

where  $[\cdot, \cdot]$  denotes the commutator. The energy of a state  $\rho$  with respect to the target Hamiltonian  $H$  is then [65]

$$\text{Tr}(\rho H) = -\frac{\text{Tr}(h\gamma)}{2}. \quad (\text{C4})$$

The covariance matrix of the state after each instance of a Gaussian unitary or noise in the circuit can be calculated by evolving the covariance matrix  $\gamma_0$  of the vacuum ( $\gamma_0$  can be analytically calculated [65]) using the following transformations.

- (1) *Unitary.* Any Gaussian unitary can be expressed as  $U = \exp(-iH_U)$ , where  $H_U = i\vec{c}^T h_U \vec{c}$  is the generating quadratic Hamiltonian, and it transforms the covariance matrix of the state as [65]

$$\gamma \rightarrow e^{2h_U} \gamma e^{-2h_U}. \quad (\text{C5})$$

- (2) *Noise.* Depolarizing noise with probability  $p$  acting on the fermion at lattice site  $x$  transforms the covariance matrix as [65,66]

$$\gamma \rightarrow p\tilde{\gamma} + (1-p)\gamma, \quad (\text{C6})$$

where matrix  $\tilde{\gamma}$  is obtained from the input covariance matrix  $\gamma$  by zeroing out the rows and columns corresponding to the fermion at site  $x$ , i.e.,

$$\tilde{\gamma}_{x',x''}^{\alpha,\alpha'} = \begin{cases} 0 & \text{if } x' = x \text{ or } x'' = x, \\ \gamma_{x',x''}^{\alpha,\alpha'} & \text{otherwise.} \end{cases} \quad (\text{C7})$$

### 3. Calculation of the dual function

We choose the dual variables  $\sigma_1, \sigma_2, \dots, \sigma_d$  to be quadratic Hamiltonians, i.e., we consider the ansatz

$$\sigma_t = i \sum_{\substack{\alpha,\alpha',x,x' \\ d(x,x') \leq r}} s_{x,x';t}^{\alpha,\alpha'} c_x^\alpha c_{x'}^{\alpha'} = i\vec{c}^T s_t \vec{c}, \quad t \in \{1, \dots, d\},$$

where  $d(x, x')$  is a distance measure between two lattice sites and the coefficients  $s_{x,x';t}^{\alpha,\alpha'}$  are real and can be arranged in an antisymmetric matrix  $s_t$ . The  $\sigma_t$  are thus local Hamiltonians with interaction range  $r$ .

In order to evaluate the circuit dual function given in Eq. (23), we calculate the effect of the circuit channels on the dual variables first, i.e.,  $\mathcal{E}_t^\dagger(\sigma_t)$ . Each such channel is the composition of a Gaussian unitary channel and the depolarizing channel, and both the unitary and noise channels map quadratic operators to quadratic operators. Hence, the channel's action can be expressed as a transformation of the compact Majorana representation  $s_t$  of the dual variable  $\sigma_t$  using the following transformation rules.

- (1) *Unitary.* Any Gaussian unitary can be expressed as  $U = \exp(-iH_U)$ , where  $H_U = i\vec{c}^T h_U \vec{c}$  is the generating quadratic Hamiltonian. Conjugation of the dual variable  $\sigma_t = i\vec{c}^T s_t \vec{c}$  by  $U$  transforms it into another quadratic operator  $\tilde{\sigma}_t = U\sigma_t U^\dagger = i\vec{c}^T \tilde{s}_t \vec{c}$  with the transformed Majorana representation

$$\tilde{s}_t = e^{2h_U} s_t e^{-2h_U}. \quad (\text{C8})$$

This transformation can be derived by expanding the conjugation using the Baker-Campbell-Hausdorff

formula  $e^B A e^{-B} = A + [B, A] + [B, [B, A]]/2! + [B, [B, [B, A]]]/3! + \dots$ , and using the commutation relation

$$\begin{aligned} [c_{x_1}^{\alpha_1} c_{x_2}^{\alpha_2}, c_{x_3}^{\alpha_3} c_{x_4}^{\alpha_4}] &= -c_{x_1}^{\alpha_1} c_{x_3}^{\alpha_3} \delta_{\alpha_2, \alpha_4} \delta_{x_2, x_4} \\ &\quad + c_{x_1}^{\alpha_1} c_{x_4}^{\alpha_4} \delta_{\alpha_2, \alpha_3} \delta_{x_2, x_3} \\ &\quad - c_{x_3}^{\alpha_3} c_{x_2}^{\alpha_2} \delta_{\alpha_1, \alpha_4} \delta_{x_1, x_4} \\ &\quad + c_{x_4}^{\alpha_4} c_{x_2}^{\alpha_2} \delta_{\alpha_1, \alpha_3} \delta_{x_1, x_3}. \end{aligned} \quad (\text{C9})$$

- (2) *Noise.* The noise channel  $\mathcal{N}_x$  corresponding to depolarizing noise on the fermion at site  $x$  with probability  $p$  transforms the dual variable  $\sigma_t$  as

$$\mathcal{N}_x(\sigma_t) = (1-p)\sigma_t + p\text{Tr}_x(\sigma_t) \otimes \frac{\mathbb{1}}{2}, \quad (\text{C10})$$

where  $\text{Tr}_x(\cdot)$  denotes the partial trace over the fermion at site  $x$ . Expanding out the partial trace term using linearity,

$$\begin{aligned} \text{Tr}_x(\sigma_t) \otimes \frac{\mathbb{1}}{2} &= \sum_{\alpha', \alpha'', x', x''} i s_{x', x''; t}^{\alpha', \alpha''} \text{Tr}_x(c_{x'}^{\alpha'} c_{x''}^{\alpha''}) \otimes \frac{\mathbb{1}}{2} \\ &= \sum_{\alpha', \alpha'', x', x''} i s_{x', x''; t}^{\alpha', \alpha''} (1 - \delta_{x, x'}) (1 - \delta_{x, x''}) \\ &\quad c_{x'}^{\alpha'} c_{x''}^{\alpha''}, \end{aligned} \quad (\text{C11})$$

where we obtain the second line by using  $\text{Tr}(c_{x'}^{\alpha'}) = 0$  and  $s_{x', x''; t}^{\alpha', \alpha''} = 0$  for all  $(x', \alpha')$ . Thus, the noise channel transforms the dual variable into another quadratic operator  $\tilde{\sigma}_t = i\vec{c}^T \tilde{s}_t \vec{c}$  with the transformed Majorana representation

$$\tilde{s}_t = p s_t + (1-p) z_t, \quad (\text{C12})$$

where, using Eq. (C11), matrix  $z_t$  is obtained from  $s_t$  by zeroing out the rows and columns corresponding to the fermion at site  $x$ , i.e.,

$$z_{x', x''; t}^{\alpha', \alpha''} = \begin{cases} 0 & \text{if } x' = x \text{ or } x'' = x, \\ s_{x', x''; t}^{\alpha', \alpha''} & \text{otherwise.} \end{cases} \quad (\text{C13})$$

To calculate the dual function given in Eq. (23), we need to calculate the Gibbs free energy of  $H_t = \sigma_t - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1})$ , where  $H_t$  is a quadratic operator itself because, as mentioned earlier, the channel  $\mathcal{E}_{t+1}^\dagger$  maps quadratic operators to quadratic operators. Therefore,  $H_t = i\vec{c}^T h_t \vec{c}$ , where the Majorana representation  $h_t$  can be obtained using the transformations described above. We calculate the free energy

by diagonalizing  $H_t$ . This can be achieved by diagonalizing  $h_t$ —as  $h_t$  is antisymmetric, it is possible [65] to find an orthogonal transformation  $O$  such that

$$O^T h_t O = \bigoplus_{\alpha, x} \begin{pmatrix} 0 & \epsilon_{x;t} \\ -\epsilon_{x;t} & 0 \end{pmatrix}.$$

We define the operator-valued vector  $\vec{c} = O^T \vec{c} = (\tilde{c}_1^1, \tilde{c}_1^2, \dots, \tilde{c}_N^1, \tilde{c}_1^2, \tilde{c}_2^2, \dots, \tilde{c}_N^2)^T$  and, due to the orthogonality of  $O$ , the operators  $\tilde{c}_x^\alpha$  satisfy the Majorana anticommutation relations given in Eq. (C1). Using the newly defined Majorana operators,  $H_t$  can be written in a diagonal form as

$$H_t = i \sum_x \epsilon_{x;t} (\tilde{c}_x^1 \tilde{c}_x^2 - \tilde{c}_x^2 \tilde{c}_x^1).$$

Using the diagonal form, the Gibbs free energy for  $H_t$  is

$$\begin{aligned} \mathcal{F}(H_t, \lambda_t) &= \log \text{Tr} \exp \left( -\frac{i}{\lambda_t} \sum_x \epsilon_{x;t} (\tilde{c}_x^1 \tilde{c}_x^2 - \tilde{c}_x^2 \tilde{c}_x^1) \right) \\ &= \log \prod_x \text{Tr} \exp \left( -\frac{i \epsilon_{x;t}}{\lambda_t} (\tilde{c}_x^1 \tilde{c}_x^2 - \tilde{c}_x^2 \tilde{c}_x^1) \right). \end{aligned}$$

The operator exponentials above can be represented in the Fock space basis of the fermion at site  $x$  as

$$\exp \left( -\frac{i \epsilon_{x;t}}{\lambda_t} (\tilde{c}_x^1 \tilde{c}_x^2 - \tilde{c}_x^2 \tilde{c}_x^1) \right) = \begin{pmatrix} e^{\epsilon_{x;t}/\lambda_t} & 0 \\ 0 & e^{-\epsilon_{x;t}/\lambda_t} \end{pmatrix}.$$

The free energy then simplifies to

$$\mathcal{F}(H_t, \lambda_t) = \sum_x \log(e^{\epsilon_{x;t}/\lambda_t} + e^{-\epsilon_{x;t}/\lambda_t}),$$

and therefore it can be calculated from just the diagonalization of the  $(2N \times 2N)$ -dimensional matrix  $h_t$ .

- 
- [1] Luigi Amico, Rosario Fazio, Andreas Osterloh, and Vlatko Vedral, Entanglement in many-body systems, *Rev. Mod. Phys.* **80**, 517 (2008).
  - [2] Sevag Gharibian, Yichen Huang, Zeph Landau, and Seung Woo Shin, Quantum Hamiltonian complexity, *Found. Trends Theor. Comput. Sci.* **10**, 159 (2015).
  - [3] Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac, Quantum computation and quantum-state engineering driven by dissipation, *Nat. Phys.* **5**, 633 (2009).
  - [4] Tameem Albash and Daniel A. Lidar, Adiabatic quantum computation, *Rev. Mod. Phys.* **90**, 015002 (2018).
  - [5] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O'Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, *Nat. Phys.* **16**, 205 (2020).
  - [6] Han-Sen Zhong, *et al.*, Quantum computational advantage using photons, *Science* **370**, 1460 (2020).
  - [7] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout van den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, and Abhinav Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
  - [8] Yulin Wu, *et al.*, Strong quantum computational advantage using a superconducting quantum processor, *Phys. Rev. Lett.* **127**, 180501 (2021).
  - [9] Yimin Ge, Jordi Tura, and J. Ignacio Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits, *J. Math. Phys.* **60**, 022202 (2019).
  - [10] Yimin Ge, András Molnár, and J. Ignacio Cirac, Rapid adiabatic preparation of injective projected entangled pair states and Gibbs states, *Phys. Rev. Lett.* **116**, 080503 (2016).
  - [11] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
  - [12] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, A quantum approximate optimization algorithm, *ArXiv:1411.4028*.
  - [13] Dave Wecker, Matthew B. Hastings, and Matthias Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
  - [14] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
  - [15] Michel X. Goemans and David P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* **42**, 1115 (1995).
  - [16] Miguel F. Anjos, On semidefinite programming relaxations for the satisfiability problem, *Math. Methods Oper. Res.* **60**, 349 (2004).
  - [17] Roberto Oliveira and Barbara M. Terhal, The complexity of quantum spin systems on a two-dimensional square lattice, *ArXiv:quant-ph/0504050*.
  - [18] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023), p. 945.
  - [19] Rahul Trivedi and J. Ignacio Cirac, Transitions in computational complexity of continuous-time local open quantum dynamics, *Phys. Rev. Lett.* **129**, 260405 (2022).
  - [20] Yiqing Zhou, E. Miles Stoudenmire, and Xavier Waintal, What limits the simulation of quantum computers? *Phys. Rev. X* **10**, 041038 (2020).
  - [21] Feng Pan and Pan Zhang, Simulation of quantum circuits using the big-batch tensor network method, *Phys. Rev. Lett.* **128**, 030501 (2022).
  - [22] Tibor Rakovszky, C. W. von Keyserlingk, and Frank Pollmann, Dissipation-assisted operator evolution method for capturing hydrodynamic transport, *Phys. Rev. B* **105**, 075131 (2022).

- [23] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, Limitations of noisy reversible computation, [ArXiv:quant-ph/9611028](https://arxiv.org/abs/quant-ph/9611028).
- [24] Daniel Stilck França and Raul García-Patrón, Limitations of optimization algorithms on noisy quantum devices, *Nat. Phys.* **17**, 1221 (2021).
- [25] Giacomo De Palma, Milad Marvian, Cambyse Rouzé, and Daniel Stilck França, Limitations of variational quantum algorithms: A quantum optimal transport approach, *PRX Quantum* **4**, 010309 (2023).
- [26] Guillermo González-García, Rahul Trivedi, and J. Ignacio Cirac, Error propagation in NISQ devices for solving classical optimization problems, *PRX Quantum* **3**, 040326 (2022).
- [27] Abhinav Deshpande, Pradeep Niroula, Oles Shtanko, Alexey V. Gorshkov, Bill Fefferman, and Michael J. Gullans, Tight bounds on the convergence of noisy random circuits to the uniform distribution, *PRX Quantum* **3**, 040329 (2022).
- [28] Rahul Trivedi, Guillermo Angeris, Logan Su, Stephen Boyd, Shanhui Fan, and Jelena Vučković, Bounds for scattering from absorptionless electromagnetic structures, *Phys. Rev. Appl.* **14**, 014025 (2020).
- [29] Guillermo Angeris, Jelena Vuckovic, and Stephen P. Boyd, Computational bounds for photonic design, *ACS Photonics* **6**, 1232 (2019).
- [30] Kurt Schab, Austin Rothschild, Kristi Nguyen, Miloslav Capek, Lukas Jelinek, and Mats Gustafsson, Trade-offs in absorption and scattering by nanophotonic structures, *Opt. Express* **28**, 36584 (2020).
- [31] Pengning Chao, Benjamin Streckha, Rodrick Kuate Defo, Sean Molesky, and Alejandro W. Rodriguez, Physical limits in electromagnetism *Nat. Rev. Phys.* **4**, 543 (2022).
- [32] Sean Molesky, Pengning Chao, Weiliang Jin, and Alejandro W. Rodriguez, Global  $\mathbb{T}$  operator bounds on electromagnetic scattering: Upper bounds on far-field cross sections, *Phys. Rev. Res.* **2**, 033172 (2020).
- [33] Sean Molesky, Pengning Chao, and Alejandro W. Rodriguez, Hierarchical mean-field  $\mathbb{T}$  operator bounds on electromagnetic scattering: Upper bounds on near-field radiative Purcell enhancement, *Phys. Rev. Res.* **2**, 043398 (2020).
- [34] Hanwen Zhang, Zeyu Kuang, Shruti Puri, and Owen D. Miller, Conservation-law-based global bounds to quantum optimal control, *Phys. Rev. Lett.* **127**, 110506 (2021).
- [35] Sattwik Deb Mishra, Rahul Trivedi, Amir H. Safavi-Naeini, and Jelena Vučković, Control design for inhomogeneous-broadening compensation in single-photon transducers, *Phys. Rev. Appl.* **16**, 044025 (2021).
- [36] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel, On quantum Rényi entropies: A new generalization and some properties, *J. Math. Phys.* **54**, 122203 (2013).
- [37] Stephen Boyd and Lieven Vandenbergh, Convex optimization, <https://www.cambridge.org/highereducation/books/convex-optimization/17D2FAA54F641A2F62C7C-CD01DFA97C4>.
- [38] Morton Slater, in *Traces and Emergence of Nonlinear Programming*, edited by Giorgio Giorgi and Tinne Hoff Kjeldsen (Springer Basel, Basel, 2014), p. 293.
- [39] Guifré Vidal, Efficient simulation of one-dimensional quantum many-body systems, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [40] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [41] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces, *J. Stat. Mech.* **2004**, P04005 (2004).
- [42] Steven R. White and Adrian E. Feiguin, Real-time evolution using the density matrix renormalization group, *Phys. Rev. Lett.* **93**, 076401 (2004).
- [43] Guifré Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
- [44] Ulrich Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys. (N. Y.)* **326**, 96 (2011).
- [45] Given a finite-dimensional Hilbert space  $\mathcal{H}$ , we use  $\mathcal{D}_1(\mathcal{H})$  to denote the set of all density matrices on  $\mathcal{H}$ , and  $\mathcal{M}(\mathcal{H})$  to denote the set of Hermitian linear operators on  $\mathcal{H}$ . We use the computer-science big-O notation for function asymptotics. In particular, given two functions  $f, g: [0, \infty) \rightarrow [0, \infty)$ ,  $f = O(g)$  if, for some  $c > 0$ ,  $f(x) \leq cg(x)$  as  $x \rightarrow \infty$  and  $f = \Omega(g)$  if, for some  $c > 0$ ,  $f(x) \geq cg(x)$  as  $x \rightarrow \infty$ .
- [46] Daniel Gottesman, The Heisenberg representation of quantum computers, [ArXiv:quant-ph/9807006](https://arxiv.org/abs/quant-ph/9807006).
- [47] Scott Aaronson and Daniel Gottesman, Improved simulation of stabilizer circuits, *Phys. Rev. A* **70**, 052328 (2004).
- [48] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert, Exponentially tighter bounds on limitations of quantum error mitigation, [ArXiv:2210.11505](https://arxiv.org/abs/2210.11505).
- [49] Tomotaka Kuwahara, Kohtaro Kato, and Fernando G. S. L. Brandão, Clustering of conditional mutual information for quantum Gibbs states above a threshold temperature, *Phys. Rev. Lett.* **124**, 220601 (2020).
- [50] Feng Pan, Pengfei Zhou, Sujie Li, and Pan Zhang, Contracting arbitrary tensor networks: General approximate algorithm and applications in graphical models and quantum circuit simulations, *Phys. Rev. Lett.* **125**, 060503 (2020).
- [51] Igor L. Markov and Yaoyun Shi, Simulating quantum computation by contracting tensor networks, *SIAM J. Comput.* **38**, 963 (2008).
- [52] Michael Lubasch, J. Ignacio Cirac, and Mari-Carmen Bañuls, Algorithms for finite projected entangled pair states, *Phys. Rev. B* **90**, 064425 (2014).
- [53] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac, Computational complexity of projected entangled pair states, *Phys. Rev. Lett.* **98**, 140506 (2007).
- [54] Luke Causser, Mari Carmen Bañuls, and Juan P. Garrahan, Optimal sampling of dynamical large deviations in two dimensions via tensor networks, *Phys. Rev. Lett.* **130**, 147401 (2023).
- [55] Daniel Robaina, Mari Carmen Bañuls, and J. Ignacio Cirac, Simulating  $2 + 1D$   $Z_3$  lattice gauge theory with an infinite

- projected entangled-pair state, *Phys. Rev. Lett.* **126**, 050401 (2021).
- [56] Michael Lubasch, J. Ignacio Cirac, and Mari-Carmen Banuls, Algorithms for finite projected entangled pair states, *Phys. Rev. B* **90**, 064425 (2014).
- [57] Mari-Carmen Banuls, David Pérez-García, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac, Sequentially generated states for the study of two-dimensional systems, *Phys. Rev. A* **77**, 052306 (2008).
- [58] Fabio Mezzacapo, Norbert Schuch, Massimo Boninsegni, and J. Ignacio Cirac, Ground-state properties of quantum many-body systems: Entangled-plaquette states and variational Monte Carlo, *New J. Phys.* **11**, 083026 (2009).
- [59] Hitesh J. Changlani, Jesse M. Kinder, Cyrus J. Umrigar, and Garnet Kin-Lic Chan, Approximating strongly correlated wave functions with correlator product states, *Phys. Rev. B* **80**, 245116 (2009).
- [60] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac, Simulation of quantum many-body systems with strings of operators and Monte Carlo tensor contractions, *Phys. Rev. Lett.* **100**, 040501 (2008).
- [61] Alessandro Sfondrini, Javier Cerrillo, Norbert Schuch, and J. Ignacio Cirac, Simulating two- and three-dimensional frustrated quantum systems with string-bond states, *Phys. Rev. B* **81**, 214426 (2010).
- [62] Christoph Hirche, Cambyse Rouzé, and Daniel Stilck França, On contraction coefficients, partial orders and approximation of capacities for quantum channels, *Quantum* **6**, 862 (2022).
- [63] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, *Phys. Rev. A* **80**, 012304 (2009).
- [64] Salman Beigi, Nilanjana Datta, and Cambyse Rouzé, Quantum reverse hypercontractivity: Its tensorization and application to strong converses, *Commun. Math. Phys.* **376**, 753 (2020).
- [65] Jacopo Surace and Luca Tagliacozzo, Fermionic Gaussian states: An introduction to numerical approaches, *SciPost Phys. Lect. Notes* **54** (2022).
- [66] Norbert Schuch and Bela Bauer, Matrix product state algorithms for Gaussian fermionic states, *Phys. Rev. B* **100**, 245121 (2019).