

Classically computing performance bounds on depolarized quantum circuits

Sattwik Deb Mishra*,¹ Miguel Frías-Pérez*,² and Rahul Trivedi†^{2,3}

¹*Ginzton Laboratory, Stanford University, 348 Via Pueblo Mall, Stanford, California 94305, USA*

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

³*Electrical and Computer Engineering, University of Washington, Seattle, Washington 98195, USA*

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.

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 meth-

ods 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 phenomena 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.

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, 29] and quantum optical devices [30, 31].

* These authors contributed equally to this work. † rtriv@uw.edu

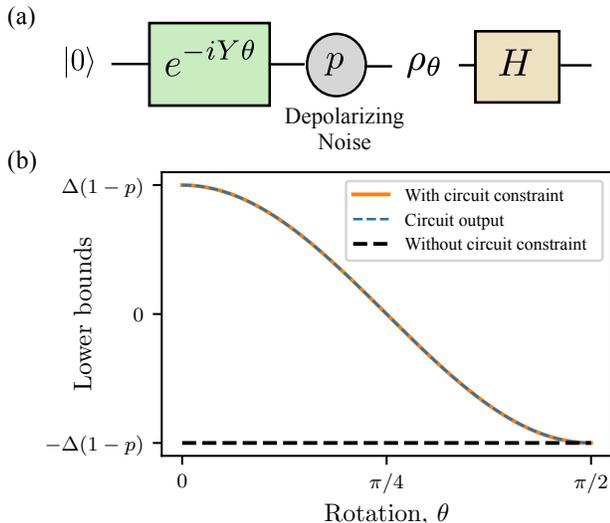


Figure 1. Comparison of bounds, with and without accounting for circuit constraints, on the minimum energy corresponding to the Hamiltonian $H = \Delta Z$ attainable by the output of the single qubit circuit shown in the schematic. The circuit consists of a Y-axis rotation followed by depolarizing noise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

II. 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 θ 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 Renyi Entropies [32]. For concreteness, we use the trace purity measure of a state ρ : $P(\rho) = \text{Tr}(\rho^2)$ — $P(\rho) = 1$ if and only if ρ is a pure state, else $P(\rho) < 1$. Now, since the state ρ_θ 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 ρ_θ is necessarily mixed, it cannot produce the pure ground state of the Hamiltonian H perfectly irrespective of the choice of θ — 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 following

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 [33]. 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, 34]. For this, we extend the problem in Eq. 1 by adding an additional constraint due to the circuit:

$$\begin{aligned} & \underset{\rho \in \mathcal{D}_1(\mathbb{C}^2)}{\text{minimize}} && \text{Tr}(H\rho) \\ & \text{subject to} && \rho = \mathcal{E}_\theta(\rho_0), \\ & && P(\rho) \leq P_0, \end{aligned} \quad (2)$$

where $\rho_0 = |0\rangle\langle 0|$ and \mathcal{E}_θ is the channel corresponding to the unitary $e^{-iY\theta}$ followed by the single-qubit depolarizing noise. To provide a lower bound on this optimization, we construct its Lagrangian $\mathcal{L}(\sigma, \lambda)$ by introducing Lagrange multipliers $\sigma \in \mathcal{M}(\mathbb{C}^2)$ [33], where $\mathcal{M}(\mathbb{C}^2)$ is the space of Hermitian single-qubit operators, and $\lambda \geq 0$,

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

$\mathcal{L}(\rho, \sigma, \lambda)$ can be considered to be a modified energy function which, 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 the state ρ . Minimizing the Lagrangian with respect to ρ , we obtain the *dual function*,

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

which is a function of σ, λ , the *dual variables*. It follows from the principle of Lagrange duality that for any σ 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 ρ , 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

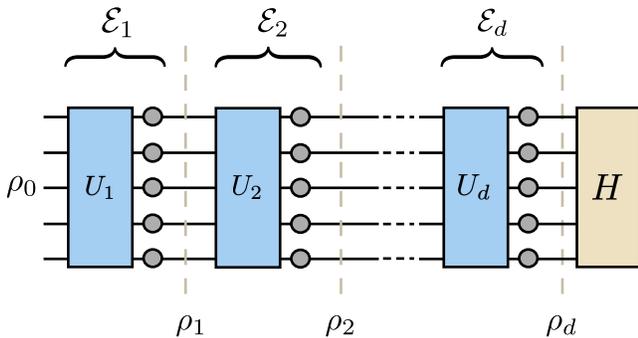


Figure 2. Schematic depiction of the problem setting considered in this paper. The unitaries (colored boxes) implement a quantum algorithm to prepare an approximation of the ground state of a target Hamiltonian H in the absence of noise. Each layer of unitary is followed by single-qubit depolarizing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

the best lower bound can be obtained by maximizing $g(\sigma, \lambda)$ with respect to σ, λ . 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 θ . 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 has been designed to approximate the ground state of a target Hamiltonian H of N qubits. In the presence of noise, the purity of the state decreases as it goes through the circuit — while computing the exact purity of the state generated by the circuit is generally difficult, an analytical upper bound on several purity measures can be obtained by making use of the fact that the noise is depolarizing. In particular, Ref. [23] establishes explicit upper bounds for two purity measures after t time-steps — the information content of the quantum state, as well as its trace purity.

Lemma 1 (Ref. [23]). *Suppose ρ_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.

In the remainder of this subsection, we denote by P_t an upper bound on the purity of the state at t^{th} time-step — we will formulate the results of this subsection for general convex purity measures, and specialize them to concrete purity measures (such as 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,

$$\begin{aligned} & \underset{\rho_1, \rho_2 \dots \rho_d \in \mathcal{S}}{\text{minimize}} && \text{Tr}(H\rho_d) \\ & \text{subject to} && \rho_t = \mathcal{E}_t(\rho_{t-1}), \quad t \in \{1, \dots, d\}, \\ & && P(\rho_t) \leq P_t, \quad t \in \{1, \dots, d\}, \end{aligned} \quad (5)$$

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 ρ_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:

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

The dual function is obtained by minimizing the Lagrangian with respect to $\rho_1, \rho_2 \dots \rho_d \in \mathcal{S}$,

$$g(\vec{\sigma}, \vec{\lambda}) = -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] + \sum_{t=1}^d \left(\mathcal{F}_{\mathcal{S}, P}(H_t, \lambda_t) - \lambda_t P_t \right), \quad (7a)$$

where $H_d = H + \sigma_d$, $H_t = \sigma_t - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1})$ for $t \in \{1, 2 \dots d-1\}$ and

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

Again, 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.

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

The function $\mathcal{F}_{\mathcal{S},P}(H, \lambda)$ can be interpreted as a generalized free energy corresponding to the Hamiltonian H at temperature λ which 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 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 the Hamiltonian H .

Proposition 1. *For the dual function defined in Eq. 7, it follows that*

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

and the maximum is attained at

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

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

Proof: The proof of this proposition follows simply by noting that, from definition,

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

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

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 σ_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 Section III, we will see that the dual function obtained on choosing the purity measure to be 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 of bond-dimension $\text{poly}(N)$. In Section IV, we will 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 $\mathcal{S}_\sigma \subset \mathcal{M}((\mathbb{C}^2)^{\otimes N})$, 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 $0 \in \mathcal{S}_\sigma$,

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

Now, we can note that

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

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

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

Furthermore, this convex problem trivially satisfies the Slater's conditions [34, 35]. This can be checked by noting that the 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.

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

which proves the proposition. \square

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 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}_σ 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.

III. TRACE PURITY-BASED LOWER BOUND

A. Formulation

In this section, we consider now 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 poly(N) bond dimension. We choose the purity measure to be 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 $\mathcal{F}_{\mathcal{S},P}(H, \lambda)$ defined in Eq. 7b evaluates to

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

and therefore, we obtain that

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

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

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

From the expression for $h(\vec{\sigma})$, we immediately notice 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 obtained $h(\vec{\sigma})$ 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) [36–40] on the Heisenberg evolution and compress the operators in each step into bond-dimension D i.e. at $\vec{\sigma}^{h,D} = \{\sigma_1^{h,D}, \sigma_2^{h,D}, \dots, \sigma_d^{h,D}\}$

$$\begin{aligned} \sigma_d^{h,D} &= -H \text{ and,} \\ \sigma_t^{h,D} &= \Pi_D \mathcal{E}_t^\dagger(\sigma_{t+1}^{h,D}) \text{ for } t \in \{1, 2 \dots d-1\}, \end{aligned} \quad (9)$$

where Π_D compresses an N -qubit operator to an operator with a bond-dimension D [41]. The bound $h(\vec{\sigma}^{h,D})$ can be related to the truncation errors made on performing TEBD. Note that from Eq. 9, the error incurred at step t of the Heisenberg evolution by the compression is $\|\mathcal{E}_t^\dagger(\sigma_{t+1}^{h,D}) - \Pi_D \mathcal{E}_t^\dagger(\sigma_{t+1}^{h,D})\| = \|H_t\|$. Importantly, while the error can be calculated with the norm chosen to be the operator norm, we use the Frobenius norm ($\|H_t\|_F = \sqrt{\text{Tr}(H_t^2)}$) since, unlike the operator norm, it can be computed efficiently for matrix product operators of low bond dimension. Using this error, the energy obtained from TEBD, $E_{\text{TEBD}} = -\text{Tr}(\rho_0 \mathcal{E}_1^\dagger(\sigma_1^{h,D}))$ and the exact energy $E = \text{Tr}(H \mathcal{E}_d \mathcal{E}_{d-1} \dots \mathcal{E}_1(\rho_0))$ deviate by at most

$$|E - E_{\text{TEBD}}| \leq \delta = \sum_{t=1}^d \sqrt{\text{Tr}(H_t^2)}. \quad (10)$$

We emphasize that δ can be computed while performing TEBD without explicitly simulating the circuit exactly. 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 depolarizing noise. In the following proposition, we show that $h(\vec{\sigma}^{h,D})$ can also be interpreted in terms of the TEBD compression errors, but with the added consideration of explicitly taking into account the purity decrease at each time step.

Proposition 3. *Suppose H is a Hamiltonian which can be exactly represented as a MPO of bond dimension D , then*

$$|E_{\text{TEBD}} - E| \leq E_{\text{TEBD}} - h(\vec{\sigma}^{h,D}) = \sum_{t=1}^d \sqrt{P_t \text{Tr}(H_t^2)}$$

Proof: We have that,

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

and

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

Defining $\rho_t = \mathcal{E}_t \mathcal{E}_{t-1} \dots \mathcal{E}_1(\rho_0)$, we then obtain that

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

where $H_t = \sigma_t^{h,D} - \mathcal{E}_{t+1}^\dagger(\sigma_{t+1}^{h,D})$. It then follows that

$$|E - E_{\text{TEBD}}| \leq \sum_{t=1}^{d-1} |\text{Tr}(\rho_t H_t)| \leq \sum_{t=1}^{d-1} \sqrt{\text{Tr}(\rho_t^2) \text{Tr}(H_t^2)}.$$

Using $\text{Tr}(\rho_t^2) \leq P_t$, we obtain the proposition. \square

In Fig. 3, we numerically exhibit the difference between the bound in Eq. 11 and $h(\vec{\sigma}^{h,D})$ for a 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 ℓ_D^{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),

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

to the lower bound ℓ^I analyzed in Ref. [24],

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

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 ℓ_D^{dual} , with $D = O(N)$ [33], scales super-exponentially with the depth of the circuit and thus captures the propagation of errors through the circuit, while ℓ^I scales at-most exponentially with the circuit depth.

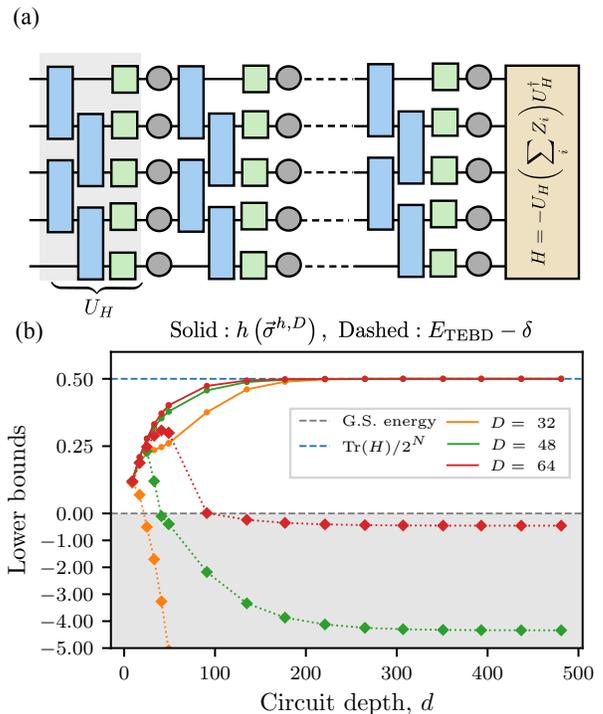


Figure 3. (a) Schematic of benchmark circuits considered for 1D spin systems: colored boxes indicate unitaries and grey circles depolarizing noise. Two-qubit unitaries are chosen to be $\exp(-i\theta X \otimes X)$ and single-qubit unitaries are independently Haar random. The Hamiltonian is chosen to 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 trace purity-based dual bound ($h(\vec{\sigma}^{h,D})$ in Eq. 8) (solid lines, circular markers) and 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 circuit depth d for a system of $N = 40$ spins, with two-qubit gate parameter $\theta = 0.05$, depolarizing noise rate of $p = 3\%$ and varying MPO ansatz bond dimensions D . Grey dashed line indicates G.S. energy, grey shaded area indicates region of trivial bounds (less than G.S. energy), blue dashed line indicates 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 4. *There exists a 1D circuit and a 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 $D \leq O(N)$.*

Proof sketch (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

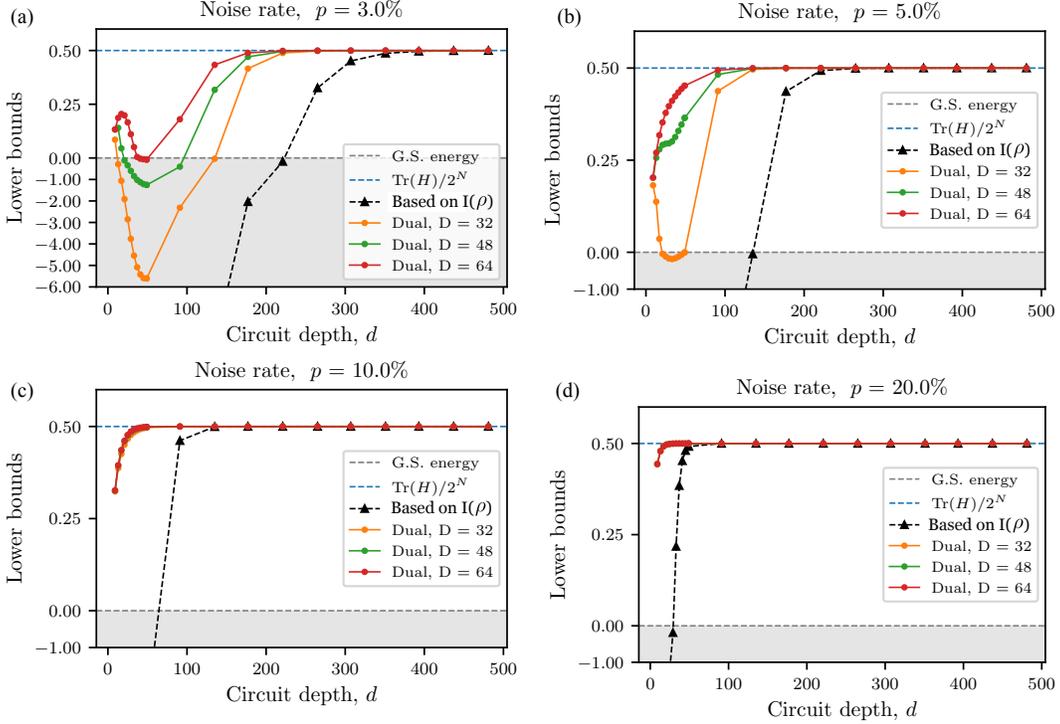


Figure 4. Comparison of the trace purity-based dual bound ($h(\bar{\sigma}^{h,D})$ in Eq. 8) and the bound based on just the information content of the output state ($\ell_{\chi_c}^I$ in Eq. 13) for 1D many-body spin systems. Both bounds are lower bounds on the ground state (G. S.) energy of the same target Hamiltonian as considered for the results in Fig. 3 (see description of Hamiltonian in 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) $p = 20\%$. 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. Grey dashed line indicates G.S. energy, grey shaded area indicates region of trivial bounds (less than G.S. energy), blue dashed line indicates energy of the completely mixed state $1/2^N$. 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 $[0, 1]$.

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 [42, 43]. Since a Pauli string is representable as an MPO of bond dimension 1, σ_t obtained from Heisenberg picture evolution will be a sum of N MPOs of bond-dimension 1 and will thus be a 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, 44]. 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 $\ell^I = -N(1-p)^{\Omega(d)}$.

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 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 $h(\bar{\sigma})$ at the dual variables obtained from TEBD on the Heisenberg picture as in Eq. 9. Fig. 4 shows numerical studies of the bounds $h(\bar{\sigma}^{h,D})$ 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 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 low-

est 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 λ . However, since H is generally a many-body Hamiltonian, an accurate evaluation of $G(H, \lambda)$ can only be guaranteed at sufficiently high temperatures [45]. Thus, instead of evaluating the bound ℓ^I in Eq. 12, we introduce a lower bound λ_c on the temperature λ 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, λ_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 λ_c as provided in Ref. [45], which provides an algorithm for evaluating $G(H, \lambda)$ at temperatures above λ_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 non-zero 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

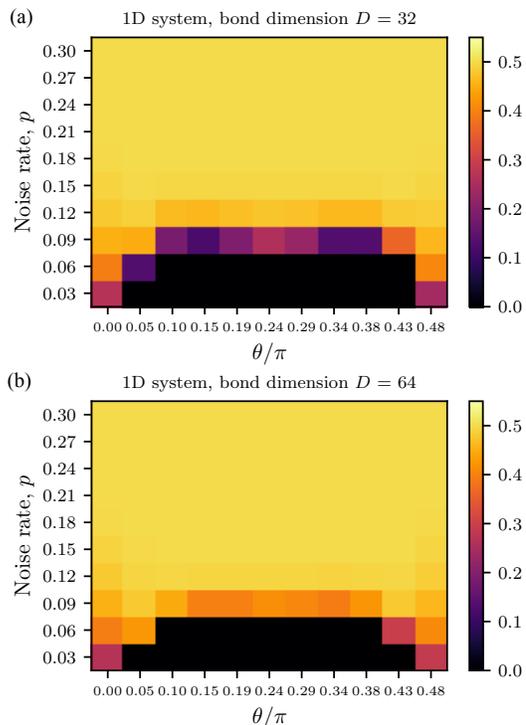


Figure 5. Trace purity-based dual bounds ($h(\bar{\sigma}^{h,D})$ in Eq. 8) for a 1D system of $N = 32$ spins, the same target Hamiltonian as considered for the results in Fig. 3 (see description of Hamiltonian in caption of Fig. 3), and brick-wall quantum circuits [Fig. 3(a)] where two-qubit unitaries 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 noise rate p and circuit parameter θ 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 $[0, 1]$.

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 θ which controls the entanglement being produced — for example, at $\theta = 0, \pi/2$, there is no entanglement at all. Fig. 5 shows the bounds as a function of the angular parameter θ and the noise rate p , for constant bond dimensions D and circuit depths d , for a 1D system of $N = 32$ spins and Fig. 6(d,e) show the same for a 2D system of $N = 36$ spins in a 6×6 lattice. For the 2D system, we consider the MPO ansatz to have a ‘snake-like’ bond structure on the 2D lattice [Fig. 6(c)] — such a snake-like 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 [46–48] 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

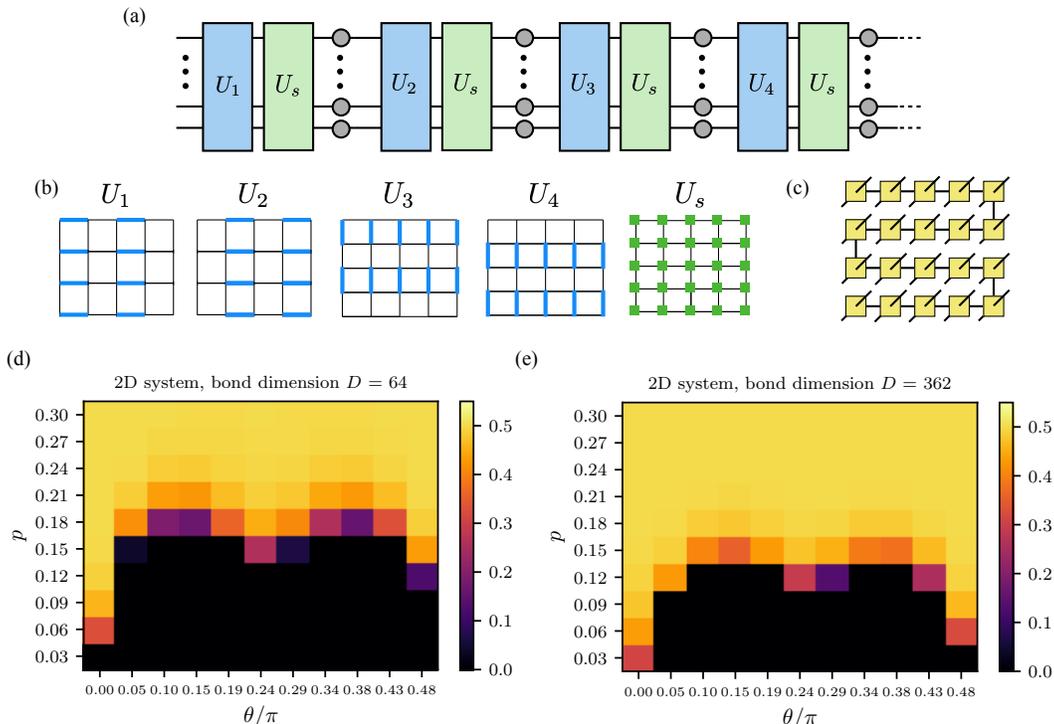


Figure 6. (a) Schematic of quantum circuits considered for 2D spin systems on a square lattice: colored boxes indicate unitaries and grey circles depolarizing noise. Each unitary layer consists of two-qubit unitaries $\exp(-i\theta X \otimes X)$ (blue boxes) followed by independently Haar random single-qubit 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 matrix product operator (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 noise rate p and circuit parameter θ for a system of $N = 36$ spins arranged in a 6×6 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 $[0, 1]$.

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 non-trivial bounds for small noise rates $p \approx 6\%$. For values of θ away from these limits, the region of triviality is larger but non-trivial dual bounds can still be obtained for higher noise rates.

IV. INFORMATION CONTENT-BASED LOWER BOUNDS

The trace purity-based dual function Eq. 8 contains terms with Frobenius norms $\sqrt{\text{Tr}(H_t^2)}$ which, 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. Alternatively, we can choose the purity function to be the information content $I(\rho) = N + \text{Tr}(\rho \log_2(\rho))$. In this case, we can obtain an analytical expression for the dual function by restricting the domain \mathcal{S} to be the space of all N -qubit density operators, i.e. Hermitian, positive semi-definite operators with unity trace. However, classically computing the dual function arising from considering the information content is difficult — we show below that it requires calculation of the Gibbs free energy of general Hamiltonians at arbitrary temperatures.

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

$$\mathcal{F}_{S,P}(H, \lambda) = \inf_{\rho \geq 0, \text{Tr}(\rho)=1} \left(\text{Tr}[H\rho] + \lambda I(\rho) \right) \quad (14)$$

$$= N\lambda - \lambda \log \text{Tr} \exp(-H/\lambda), \quad (15)$$

leading to the following dual function,

$$\tilde{g}(\vec{\sigma}, \vec{\lambda}) = -\text{Tr}[\rho_0 \mathcal{E}_1^\dagger(\sigma_1)] + \sum_{t=1}^d \left(-\lambda_t \log \text{Tr} \exp(-H_t/\lambda_t) + \lambda_t(N - I_t) \right), \quad (16)$$

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 Eq. 16 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 $h_{x, x'}^{\alpha, \alpha'}$ 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 σ_t which themselves are quadratic Hermitian operators. This motivates the following ansatz for σ_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 σ_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 σ_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 (L-BFGS-B), starting from the initial point where σ_t are chosen to be the Heisenberg picture evolution of $-H$, but projected on to 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 7 shows a numerical study of the bounds that we obtain — we consider systems with ~ 50 fermions

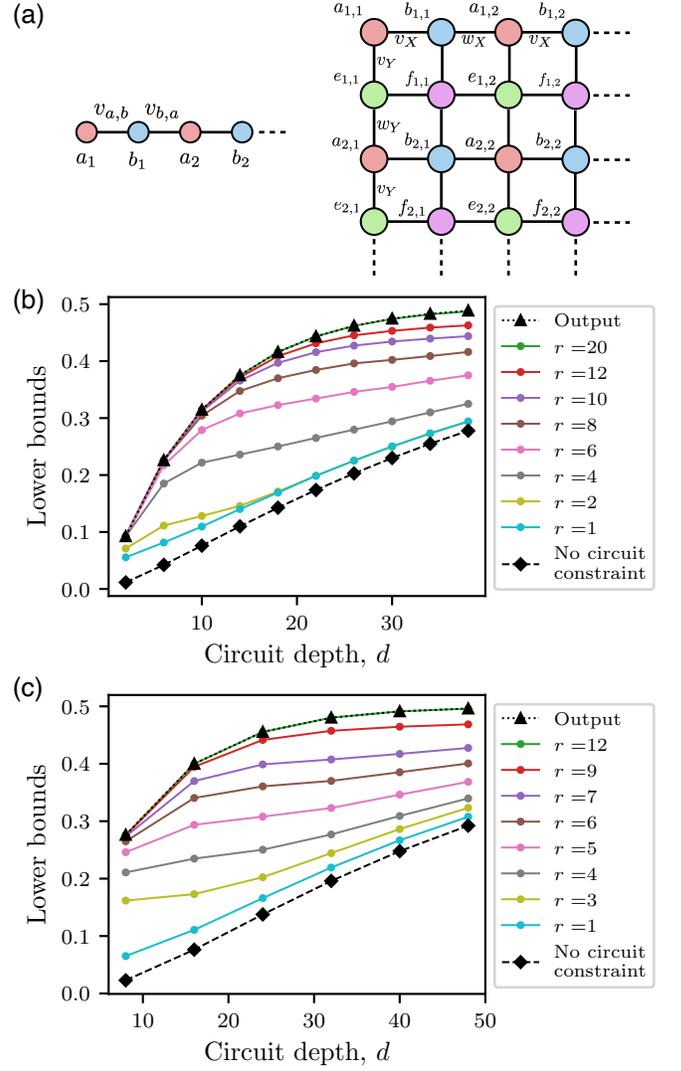


Figure 7. (a) Schematic of SSH model quadratic fermionic Hamiltonians with alternating hopping strengths in 1D and 2D. (b,c) Comparison of information content-based dual bounds with and without circuit constraints, and the output energies of noisy Gaussian circuits for systems consisting of (b) $N = 48$ fermions arranged in a 1D lattice, (c) $N = 49$ fermions arranged in a 7×7 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 $[0, 1]$.

arranged both on 1D [Fig. 7(b)] and 2D lattices [Fig. 7(c)] and experiencing depolarizing noise at a rate of 5%. H is chosen to be a SSH model, nearest neighbor Hamiltonian with alternating hopping strengths [Fig. 7(a)]. For the

1D benchmarks, we choose

$$H = \sum_x \left(v_{a,b} a_x^\dagger b_x + v_{b,a} b_x^\dagger a_{x+1} + \text{h.c.} \right), \quad (17a)$$

and for 2D benchmarks, we choose

$$H = \sum_{x,y} \left(\sum_{\substack{p,q \in \\ \{(a,b),(e,f)\}}} \left(v_{p,q} p_{x,y}^\dagger q_{x,y} + v_{q,p} q_{x,y}^\dagger p_{x+1,y} + \text{h.c.} \right) \right. \\ \left. + \sum_{\substack{p,q \in \\ \{(a,e),(b,f)\}}} \left(v_{p,q} p_{x,y}^\dagger q_{x,y} + v_{q,p} q_{x,y}^\dagger p_{x,y+1} + \text{h.c.} \right) \right), \quad (17b)$$

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. 7, 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. 7, 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.

V. CONCLUSION

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 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. Larger-scale tensor-network numerics can allow us to also study higher dimensional circuits, with a number of qubits reaching state-of-the-art experiments. Numerical optimization algorithms can be explored for obtaining better 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 in understanding resilience of quantum circuit architectures to noise.

ACKNOWLEDGEMENTS

We thank J. Ignacio Cirac, Mari Carmen Bañuls, and Alex White 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 Ref. [26] which analyzed error propagation in a family of random quantum circuits. Specifically, they considered brick-wall quantum circuits of depth d (assumed to be even) with unitaries $U_1, U_2 \dots U_d$, where

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

The unitaries $U_1, U_2 \dots U_{d/2}$ are chosen randomly depending on the circuit architecture. We will specifically consider the 1D case, where $U_1, U_3, U_5 \dots$ are formed by applying random 2-qubit gates between qubits $(1, 2), (3, 4), (5, 6) \dots$ and $U_2, U_4, U_6 \dots$ are formed by applying random 2-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. Ref. [26] establishes 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 two-local Hamiltonians, we will only consider the Hamiltonian $H = -\sum_{i=1}^N Z_i$.

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

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

Proof of proposition 4: Consider the Hamiltonian $H = -\sum_{i=1}^N Z_i$, and choose the two qubit gates to be a Haar-random Clifford gate — since Haar-random Clifford gates form a 2-design [49], we can use lemma A.1. 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 $-H$,

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

Note that for Clifford circuits, Z_i will be mapped to a single Pauli string [43], which is expressible as an MPO with bond dimension 1, and consequently σ_t will have a bond-dimension of at-most N . Since the Heisenberg picture evolution can be captured exactly with a MPO ansatz of bond dimension $O(N)$, the dual bound ℓ_D^{dual} with $D \leq O(N)$ is exactly equal to the energy E of the output of the noisy quantum circuit. Now, from lemma A.1, 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 ℓ^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 ℓ^I which converges to $\text{Tr}(H/2^N) = 0$ exponentially with d can be obtained by computing $\text{Tr}(H\rho)$ at

$$\rho = \left(p_d |0\rangle\langle 0| + (1-p_d) \frac{I}{2} \right)^{\otimes N} \quad \text{where } p_d = (1-p)^d.$$

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

Appendix B: Free fermion formulation

In this section, 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 (16) 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 following anti-commutation relations,

$$\begin{aligned} \{c_x^1, c_{x'}^1\} &= \{c_x^2, c_{x'}^2\} = \delta_{xx'} \\ \{c_x^1, c_{x'}^2\} &= 0, \quad \forall x, x' \in \text{lattice}. \end{aligned} \quad (\text{B1})$$

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,

$$\vec{c} = (c_1^1, c_2^1, \dots, c_N^1, c_1^2, c_2^2, \dots, c_N^2)^T, \quad (\text{B2})$$

the target Hamiltonian can be written as $H = i\vec{c}^T h \vec{c}$ where the matrix h specifying the Hamiltonian H is real and anti-symmetric. The ground state energy of the target Hamiltonian can be computed by diagonalization of its compact representation h [50].

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 ρ as,

$$\gamma_{x, x'}^{\alpha, \alpha'} = i \text{Tr}(\rho [c_x^\alpha, c_{x'}^{\alpha'}]), \quad (\text{B3})$$

where $[\cdot, \cdot]$ denotes the commutator. The energy of a state ρ with respect to the target Hamiltonian H is then [50],

$$\text{Tr}(\rho H) = -\frac{\text{Tr}(h\gamma)}{2}. \quad (\text{B4})$$

The covariance matrix of the state after each instance of Gaussian unitary or noise in the circuit can be calculated by evolving the covariance matrix γ_0 of the vacuum (γ_0 can be analytically calculated [50]) 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 [50],

$$\gamma \rightarrow e^{2h_U} \gamma e^{-2h_U}. \quad (\text{B5})$$

2. *Noise*: Depolarizing noise with probability p acting on the fermion at lattice site x transforms the covariance matrix in the following manner [50, 51],

$$\gamma \rightarrow p\tilde{\gamma} + (1-p)\gamma, \quad (\text{B6})$$

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

$$\tilde{\gamma}_{x', x''}^{\alpha, \alpha'} = \begin{cases} 0, & \text{if } x' = x \text{ or } x'' = x \\ \gamma_{x', x''}^{\alpha, \alpha'}, & \text{else.} \end{cases} \quad (\text{B7})$$

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 following 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 into an anti-symmetric matrix s_t . σ_t are thus local Hamiltonians with interaction range r .

In order to evaluate the circuit dual function (16), 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 σ_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 ,

$$\tilde{s}_t = e^{2h_U} s_t e^{-2h_U}. \quad (\text{B8})$$

This transformation can be derived by expanding the conjugation using the Baker-Campbell-Hausdorff formula $e^B A e^{-B} = A + [B, A] + \frac{1}{2!}[B, [B, A]] + \frac{1}{3!}[B, [B, [B, A]]] + \dots$, and using the commutation relation,

$$\begin{aligned} [c_{x_1}^{\alpha_1} c_{x_2}^{\alpha_2}, c_{x_3}^{\alpha_3} c_{x_4}^{\alpha_4}] &= -c_{x_1}^{\alpha_1} c_{x_3}^{\alpha_3} \delta_{\alpha_2, \alpha_4} \delta_{x_2, x_4} + c_{x_1}^{\alpha_1} c_{x_4}^{\alpha_4} \delta_{\alpha_2, \alpha_3} \delta_{x_2, x_3} \\ &\quad - c_{x_3}^{\alpha_3} c_{x_2}^{\alpha_2} \delta_{\alpha_1, \alpha_4} \delta_{x_1, x_4} + c_{x_4}^{\alpha_4} c_{x_2}^{\alpha_2} \delta_{\alpha_1, \alpha_3} \delta_{x_1, x_3}. \end{aligned} \quad (\text{B9})$$

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 σ_t as,

$$\mathcal{N}_x(\sigma_t) = (1-p)\sigma_t + p\text{Tr}_x(\sigma_t) \otimes \frac{\mathbb{1}}{2}, \quad (\text{B10})$$

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

$$\begin{aligned} \text{Tr}_x(\sigma_t) \otimes \frac{\mathbb{1}}{2} &= \sum_{\alpha', \alpha'', x', x''} i s_{x', x''; t}^{\alpha', \alpha''} \text{Tr}_x(c_{x'}^{\alpha'} c_{x''}^{\alpha''}) \otimes \frac{\mathbb{1}}{2}, \\ &= \sum_{\alpha', \alpha'', x', x''} i s_{x', x''; t}^{\alpha', \alpha''} (1 - \delta_{x, x'}) (1 - \delta_{x, x''}) c_{x'}^{\alpha'} c_{x''}^{\alpha''}, \end{aligned} \quad (\text{B11})$$

where we obtain the second line by using $\text{Tr}(c_{x'}^{\alpha'}) = 0$ and $s_{x', x'; t}^{\alpha', \alpha'} = 0 \forall (x', \alpha')$. 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,

$$\tilde{s}_t = p s_t + (1-p) z_t, \quad (\text{B12})$$

where, using (B11), the matrix z_t is obtained from s_t by zeroing out the rows and columns corresponding to the fermion at site x , i.e.,

$$z_{x', x''; t}^{\alpha', \alpha''} = \begin{cases} 0, & \text{if } x' = x \text{ or } x'' = x \\ s_{x', x''; t}^{\alpha', \alpha''}, & \text{else.} \end{cases} \quad (\text{B13})$$

To calculate the dual function (16), 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 anti-symmetric, it is possible [50] 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 $\vec{c} = O^T \vec{c} = (\tilde{c}_1^1, \tilde{c}_2^1, \dots, \tilde{c}_N^1, \tilde{c}_1^2, \tilde{c}_2^2, \dots, \tilde{c}_N^2)^T$ and, due to the orthogonality of O , the operators \tilde{c}_x^α satisfy the Majorana anti-commutation relations (B1). 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 ,

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

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 just from 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,” *Reviews of modern physics* **80**, 517 (2008).
 - [2] Sevag Gharibian, Yichen Huang, Zeph Landau, Seung Woo Shin, *et al.*, “Quantum hamiltonian complexity,” *Foundations and Trends® in Theoretical Computer Science* **10**, 159–282 (2015).
 - [3] Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac, “Quantum computation and quantum-state engineering driven by dissipation,” *Nature Phys* **5**, 633–636 (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, *et al.*, “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution,” *Nat. Phys.* **16**, 205–210 (2020).
 - [6] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, *et al.*, “Quantum computational advantage using photons,” *Science* **370**, 1460–1463 (2020).
 - [7] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, *et al.*, “Evidence for the utility of quantum computing before fault tolerance,” *Nature* **618**, 500–505 (2023).
 - [8] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, *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,” *Journal of Mathematical Physics* **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, *et al.*, “Variational quantum algorithms,” *Nat Rev Phys* **3**, 625–644 (2021).
 - [12] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, “A Quantum Approximate Optimization Algorithm,” (2014), [arxiv:1411.4028 \[quant-ph\]](https://arxiv.org/abs/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, *et al.*, “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**, 1115–1145 (1995).
- [16] Miguel F Anjos *et al.*, “On semidefinite programming relaxations for the satisfiability problem,” *Math. Methods Oper. Res.* **60**, 349–367 (2004).
- [17] Roberto Oliveira and Barbara M Terhal, “The complexity of quantum spin systems on a two-dimensional square lattice,” arXiv preprint quant-ph/0504050 (2005).
- [18] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani, “A Polynomial-Time Classical Algorithm for Noisy Random Circuit Sampling,” in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023) pp. 945–957.
- [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,” (1996), arxiv:quant-ph/9611028.
- [24] Daniel Stilck França and Raul García-Patrón, “Limitations of optimization algorithms on noisy quantum devices,” *Nat. Phys.* **17**, 1221–1227 (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**, 1232–1239 (2019).
- [30] 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).
- [31] Sattwik Deb Mishra, Rahul Trivedi, Amir H. Safavi-Naeini, and Jelena Vučković, “Control Design for Inhomogeneous-Broadening Compensation in Single-Photon Transducers,” *Phys. Rev. Applied* **16**, 044025 (2021).
- [32] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel, “On quantum rényi entropies: A new generalization and some properties,” *Journal of Mathematical Physics* **54**, 122203 (2013).
- [33] 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 $x \rightarrow \infty$.
- [34] Stephen Boyd and Lieven Vandenbergh, “*Convex Optimization*,” <https://www.cambridge.org/highereducation/books/convex-optimization/17D2FAA54F641A2F62C7CCD01DFA97C4> (2004).
- [35] Morton Slater, “Lagrange Multipliers Revisited,” in *Traces and Emergence of Nonlinear Programming*, edited by Giorgio Giorgi and Tinne Hoff Kjeldsen (Springer Basel, Basel, 2014) pp. 293–306.
- [36] Guifré Vidal, “Efficient Simulation of One-Dimensional Quantum Many-Body Systems,” *Phys. Rev. Lett.* **93**, 040502 (2004).
- [37] 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).
- [38] 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).
- [39] Steven R. White and Adrian E. Feiguin, “Real-Time Evolution Using the Density Matrix Renormalization Group,” *Phys. Rev. Lett.* **93**, 076401 (2004).
- [40] Guifré Vidal, “Efficient Classical Simulation of Slightly Entangled Quantum Computations,” *Phys. Rev. Lett.* **91**, 147902 (2003).
- [41] Ulrich Schollwöck, “The density-matrix renormalization group in the age of matrix product states,” *Annals of physics* **326**, 96–192 (2011).
- [42] Daniel Gottesman, “The Heisenberg Representation of Quantum Computers,” (1998), arxiv:quant-ph/9807006.
- [43] Scott Aaronson and Daniel Gottesman, “Improved Simulation of Stabilizer Circuits,” *Phys. Rev. A* **70**, 052328 (2004), arxiv:quant-ph/0406196.
- [44] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert, “Exponentially tighter bounds on limitations of quantum error mitigation,” (2023), arxiv:2210.11505 [math-ph, physics:quant-ph].

- [45] 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 \(2020\)](#).
- [46] 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 \(2020\)](#).
- [47] Igor L. Markov and Yaoyun Shi, “Simulating Quantum Computation by Contracting Tensor Networks,” [SIAM J. Comput. **38**, 963–981 \(2008\)](#).
- [48] Michael Lubasch, J. Ignacio Cirac, and Mari-Carmen Bañuls, “Algorithms for finite projected entangled pair states,” [Phys. Rev. B **90**, 064425 \(2014\)](#).
- [49] 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 \(2009\)](#).
- [50] Jacopo Surace and Luca Tagliacozzo, “Fermionic Gaussian states: An introduction to numerical approaches,” [SciPost Phys. Lect. Notes , 54 \(2022\)](#).
- [51] Norbert Schuch and Bela Bauer, “Matrix product state algorithms for Gaussian fermionic states,” [Phys. Rev. B **100**, 245121 \(2019\)](#).