

Efficient Quantum Gibbs Sampling with Local Circuits

Dominik Hahn^{1,2,*}, Ryan Sweke^{3,4,5,6}, Abhinav Deshpande³, and Oles Shtanko^{2,†}

¹*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*

²*IBM Quantum, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA*

³*IBM Quantum, IBM Research—Almaden, San Jose, California 95120, USA*

⁴*African Institute for Mathematical Sciences (AIMS), Cape Town, South Africa*

⁵*Department of Mathematical Sciences, Stellenbosch University, Stellenbosch 7600, South Africa*

⁶*National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa*



(Received 30 June 2025; revised 13 January 2026; accepted 17 March 2026; published 23 April 2026)

The problem of simulating the thermal behavior of quantum systems remains a central open challenge in quantum computing. Unlike well-established quantum algorithms for unitary dynamics, *provably efficient* algorithms for preparing thermal states—crucial for probing equilibrium behavior—became available only recently with breakthrough algorithms based on the simulation of well-designed dissipative processes, a quantum-analog to Markov chain Monte Carlo algorithms. We show a way to implement these algorithms avoiding expensive block encoding and relying only on dense local circuits, akin to Hamiltonian simulation. Specifically, our method leverages spatial truncation and Trotterization of exact quasilocal dissipative processes. We rigorously prove that the approximations we use have little effect on rapid mixing at high temperatures and allow convergence to the thermal state with small bounded error. Moreover, we accompany our analytical results with numerical simulations that show that this method, unlike previously thought, is within the reach of current generation of quantum hardware. These results provide the first provably efficient quantum thermalization protocol implementable on near-term quantum devices, offering a concrete path toward practical simulation of equilibrium quantum phenomena.

DOI: [10.1103/m5zb-z4nn](https://doi.org/10.1103/m5zb-z4nn)

I. INTRODUCTION

The exponential complexity of quantum many-body systems has long posed a formidable barrier to our understanding of condensed matter, chemistry, and high-energy physics. While classical methods have achieved notable success in simulating some quantum systems, they often break down in strongly correlated regimes—precisely where the most intriguing quantum phenomena emerge, such as high-temperature superconductivity or quantum phase transitions [1]. Quantum computers, originally envisioned by Feynman as simulators for these complex systems [2,3], are now becoming a viable platform for tackling such challenges as advances in quantum hardware accelerate rapidly [4].

Among the central goals in this domain is the preparation of thermal, or Gibbs, states—quantum analogs of classical equilibrium distributions that encode all information about a system at finite temperature [5]. These states are crucial for exploring quantum statistical mechanics [6] and serve as important computational primitives in quantum machine learning, inference, and optimization [7,8].

Despite their importance, the efficient preparation of Gibbs states remains a central open problem in quantum computing. Classically, Markov chain Monte Carlo (MCMC) techniques, such as the Metropolis sampling algorithm [9], have long been established as universal and efficient algorithms for simulating classical thermal states in many relevant cases [10]. However, while quantum algorithms for quantum thermal-state preparation—including phase estimation techniques [11], thermal bath emulation [12], quantum imaginary time evolution [13], variational methods [14–16], and quantum Metropolis algorithms [17,18]—have shown substantial progress, each comes with trade-offs in terms of scalability, provable guarantees, or compatibility with near-term devices.

A particularly promising direction involves the use of open-system dynamics inspired by classical MCMC,

*Contact author: dominik.hahn@ox.ac.uk

†Contact author: oles.shtanko@ibm.com

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.

wherein a quantum system evolves under dissipative evolution that converges to a desired thermal state [19,20]. Recent breakthroughs [21,22] have led to developments of quasilocal dissipative processes that converge to the Gibbs state with provable guarantees on mixing time, at least in certain regimes [23–25]. Using the direct analogy with classical MCMC, we will refer to this approach below as quantum Markov chain Monte Carlo (or qMCMC). While earlier schemes required a continuous family of jump operators in the corresponding Lindblad equation, it has recently been shown that qMCMC can be implemented with as few as one jump operator [26,27]. A major challenge that remains is the intrinsic quasilocality of these dissipative processes. As a result, current implementations [22,26] cannot exploit techniques from local Hamiltonian simulation and must instead rely on block-encoding methods, which are difficult to realize on noisy quantum hardware (but see Refs. [28–30] for alternative approaches).

In this work, we show how to address this remaining challenge. Specifically, we demonstrate that applying spatial truncation and Trotterization to qMCMC yields thermalization protocols that can be implemented directly with only *local* quantum circuits. Importantly, we provide rigorous theoretical bounds showing that our proposed protocol preserves logarithmic mixing time for high temperatures. Additionally, in order to facilitate the execution of this algorithm on near-term quantum devices, we show that variational compilation can be used to compile the local processes into circuits which can be executed on currently available quantum hardware, in a way that allows one to trade off accuracy with circuit depth. As such, our method opens a path toward thermal-state preparation that is both analytically grounded and experimentally feasible on near-term quantum devices.

More specifically, our implementation scheme consists of three main steps. The first addresses one of the central challenges in existing qMCMC algorithms: the quasilocal nature of the dissipation. Our approach is to provide truncated versions of these dissipative processes, resulting in a strictly local Lindbladian generator amenable to circuit-based quantum simulation. This process can be understood as constructing a Lindbladian satisfying a *local detailed-balance* condition: each jump operator satisfies detailed-balance condition with respect to a Hamiltonian truncated to the system subset.

We show that jump operators computed on different, overlapping Hamiltonian patches collectively generate a Gibbs state of the full (untruncated) Hamiltonian. To establish this, we extended the rapid mixing analysis from Rouzé *et al.* [31], originally applied to the scheme in Chen *et al.* [22], to the more recent construction proposed in Ding *et al.* [26]. This extension enables a more efficient Gibbs sampler using only a few jump operators per qubit. At sufficiently high temperatures, we show that the

output of the truncated dynamics converges to the true Gibbs distribution when the truncation radius scales at most logarithmically with system size. In contrast, accurate estimation of local observables requires only a truncation radius that is independent of system size. These results, formalized in Theorems 1–3 below, provide the first provably efficient construction of local dissipative preparation of Gibbs states.

Next, having obtained a truncated Lindbladian with provable thermalization guarantees, we employ a randomized Trotterization strategy [28] and demonstrate that its continuous-time evolution can be faithfully approximated by a circuit composed of small local quantum channels. This construction ensures that the convergence properties of the continuous generator are retained, while the Trotter error scales quadratically with the step size.

Finally, to bridge the gap between these finite-size channels and operations available on realistic gate-based quantum devices, we employ a variational compilation framework. In particular, we give compelling numerical evidence that variational compilation allows one to compile these local channels into circuits with a single-qubit reset operation, and a fixed desired number of two-qubit gates, respecting hardware connectivity. Additionally, we show that this framework allows one to trade-off circuit depth with accuracy, which facilitates more accurate simulations in step with yearly improvements in gate fidelity.

The proposed protocol provides a practical alternative to the amplitude-encoding and controlled-unitary frameworks introduced in earlier works [22,26]. In particular, it is more readily implementable on near-term and early fault-tolerant quantum devices, which often face stringent constraints on circuit depth and qubit connectivity.

We validate the full protocol through numerical simulations on one- and two-dimensional spin systems. The results confirm rapid convergence to thermal states with low-energy density and accurate local observables, even at modest truncation radii that include only few near neighbors. Moreover, the method exhibits some resilience to noise, underscoring its applicability to near-term quantum devices.

The remainder of the paper is organized as follows. In Sec. II, we introduce the theoretical framework underlying quantum thermalization via Lindbladian dynamics and review the construction of quantum Gibbs samplers based on the detailed-balance condition. Section III A presents our main results on the truncation of quasilocal jump operators, including formal error bounds and implications for local observables. In Sec. III B, we describe the randomized Trotterization scheme and derive its associated error estimates. Section IV details our variational compilation strategy for mapping the resulting local gadgets onto shallow quantum circuits. Finally, Secs. V and VI report numerical experiments on spin chains: first in an idealized setting, then incorporating realistic hardware constraints.

II. THEORETICAL BACKGROUND

This section introduces the necessary background for understanding how to prepare a Gibbs state using the qMCMC method based on Lindblad time evolution. A comprehensive list of all definitions and notations for the entire paper is available in Appendix A.

We consider a finite D -dimensional square lattice Λ consisting of n sites such that each site $a \in \Lambda$ is associated with a qubit. Let the total Hilbert space of the lattice be $\mathcal{H}_\Lambda$, the corresponding algebra of operators over this space be denoted by $\mathcal{M}_\Lambda$, and the set of Hermitian operators denoted by $\mathcal{S}_\Lambda$. Then, given $H \in \mathcal{S}_\Lambda$ and a non-negative inverse temperature $\beta \geq 0$, our goal is to prepare the Gibbs state defined as

$$\rho_\beta := \frac{e^{-\beta H}}{\text{tr} e^{-\beta H}}. \quad (1)$$

In the remainder of this paper, we focus on local Hamiltonians of the form

$$H = \sum_{X \subset \Lambda} H_X, \quad (2)$$

where each term H_X is strictly supported on a subset $X \subseteq B_{a_\mu}(r_\mu) \subset \Lambda$ involving a constant number of qubits inside the ball $B_{a_\mu}(r_\mu)$ of constant radius r_μ around qubit $a_\mu \in \Lambda$.

In the 1970s, Davies proposed that, in the weak-coupling limit, the dynamics of a system coupled to a large thermal bath at inverse temperature β is described by a Markovian master equation that has the Gibbs state for its Hamiltonian H as its unique fixed point [19,32]. The Davies process, when expressed in the Hamiltonian's eigenbasis, typically induces transitions between pairs of eigenstates, resulting in nonlocal dynamics that often involve all qubits in the system.

Recent works [21,22,26] have shown that the Davies process arises as a special case within a broader family of thermalizing Markovian dynamics that respects the detailed-balance condition [see Eq. (9)]. In particular, the foundational work of Chen *et al.* [22] demonstrated that it is possible to construct a process that converges exactly to the Gibbs state even when the transitions are modified to have finite energy resolution, rather than mapping directly from eigenstate to eigenstate. Building on this, the work by Rouzé *et al.* [31] showed that at sufficiently high temperatures, such algorithms exhibit rapid mixing, achieving convergence to the Gibbs state in logarithmic time in the number of qubits. In this work, we adopt a more recent construction proposed by Ding *et al.* [26], which, in contrast to earlier approaches, requires only a finite number of jump operators per qubit.

Specifically we consider dynamics that, acting on a density matrix ρ , can be described by a Gorini-Kossakowski-

Sudarshan-Lindblad equation

$$\frac{d}{dt}\rho = \mathcal{L}^\beta(\rho) = \sum_{a \in \Lambda} \sum_{\alpha} \mathcal{L}_{a,\alpha}^\beta(\rho), \quad (3)$$

where each index α takes discrete values. Each local Lindbladian term [33] $\mathcal{L}_{a,\alpha}^\beta$ takes the form

$$\begin{aligned} \mathcal{L}_{a,\alpha}^\beta(\rho) = & -i[G_{a,\alpha}^\beta, \rho] \\ & + L_{a,\alpha}^\beta \rho L_{a,\alpha}^{\beta\dagger} - \frac{1}{2}\{L_{a,\alpha}^{\beta\dagger} L_{a,\alpha}^\beta, \rho\}, \end{aligned} \quad (4)$$

where $G_{a,\alpha}^\beta \in \mathcal{S}_\Lambda$ are Hermitian operators that we define later, and the jump operators $L_{a,\alpha}^\beta \in \mathcal{M}_\Lambda$ are given by

$$L_{a,\alpha}^\beta = \int_{-\infty}^{\infty} dt f(t) e^{iHt} A^{a,\alpha} e^{-iHt}. \quad (5)$$

Here, we choose $A^{a,\alpha}$ to be single-qubit Hermitian operators supported on the qubit at lattice site $a \in \Lambda$, as per Eq. (16). Also, $f(t)$ is a filter function given by

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dv q(v) e^{-\beta v/4} e^{itv}, \quad (6)$$

where the envelope function $q(v)$ is symmetric, satisfying $q(-v) = q^*(v)$. It can be chosen such that the resulting filter function effectively suppresses contributions to the integral in Eq. (5) at large times t . For local Hamiltonians H that satisfy a Lieb-Robinson bound, this implies that the jump operators $L_{a,\alpha}^\beta$ are quasilocal; in other words, they are predominantly supported on qubits near site a .

To define the coherent term, consider the spectral decomposition of the Hamiltonian $H = \sum_i \lambda_i P_i$, where $\lambda_i \in \mathbb{R}$ are eigenvalues and $P_i \in \mathcal{S}_\Lambda$ are the corresponding spectral projectors. The coherent term $G_{a,\alpha}^\beta \in \mathcal{S}_\Lambda$ is explicitly given by

$$G_{a,\alpha}^\beta = -\frac{i}{2} \sum_{v \in \Omega(H)} \tanh\left(-\frac{\beta v}{4}\right) (L_{a,\alpha}^{\beta\dagger} L_{a,\alpha}^\beta)_v, \quad (7)$$

where $\Omega(H) = \{\lambda_i - \lambda_j \mid i, j\}$ denotes the set of Bohr frequencies of H , and for any operator $A \in \mathcal{M}_\Lambda$, the component $A_v \in \mathcal{M}_\Lambda$ is defined as

$$A_v := \sum_{i,j: \lambda_i - \lambda_j = v} P_i A P_j, \quad (8)$$

where the sum is taken over all pairs of energy levels separated by frequency v [34].

It can be demonstrated [26] that the chosen filter function and coherent term $G_{a,\alpha}^\beta$ guarantee that the Lindbladian

$\mathcal{L}^\beta$ satisfies the Kubo-Martin-Schwinger (KMS) detailed-balance condition

$$\mathcal{L}^{\beta\dagger}(\cdot) = \rho_\beta^{-\frac{1}{2}} \mathcal{L}^\beta \left(\rho_\beta^{\frac{1}{2}} \cdot \rho_\beta^{\frac{1}{2}} \right) \rho_\beta^{-\frac{1}{2}}, \quad (9)$$

where the adjoint superoperator $\mathcal{L}^{\beta\dagger}$ is defined with respect to the inner product $\langle A, B \rangle = 2^{-n} \text{Tr}(A^\dagger B)$ via the relation $\langle A, \mathcal{L}^\beta(B) \rangle = \langle \mathcal{L}^{\beta\dagger}(A), B \rangle$. The KMS condition, expressed in Eq. (9), can be viewed as an extension of the detailed-balance condition known from classical Markov processes [10] to the quantum setting [22]. Directly from Eq. (9), it follows that the Gibbs state ρ_β is a steady state of the Lindbladian [21], i.e.,

$$\mathcal{L}^\beta(\rho_\beta) = 0. \quad (10)$$

For concreteness, unless otherwise stated, we choose the envelope function $q(v)$ in Eq. (6) to be a Gaussian,

$$q(v) = \exp\left(-\frac{(\beta v)^2}{8}\right). \quad (11)$$

With this specific choice of $q(v)$, it is possible to give an explicit expression for $f(t)$ and $G_{a,\alpha}^\beta$ as integrals in the time domain. With the Gaussian choice in Eq. (11), $f(t)$ is given by (see Appendix C)

$$f(t) = \sqrt{\frac{2}{\pi\beta^2}} \exp\left(\frac{(\beta - 4it)^2}{8\beta^2}\right) \quad (12)$$

and

$$G_{a,\alpha}^\beta = \int_{-\infty}^{\infty} dt g_1(t) e^{-iHt} \quad (13)$$

$$\times \left(\int_{-\infty}^{\infty} dt' g_2(t') e^{iHt'} A^{a,\alpha\dagger} e^{-2iHt'} A^{a,\alpha} e^{iHt'} \right) e^{iHt}.$$

with

$$g_1(t) = \left[-\frac{1}{\pi\beta \cosh\left(\frac{2\pi t}{\beta}\right)} \right] *_t \left[\frac{\sqrt{2}}{\beta} e^{\frac{1}{4} - \frac{4t^2}{\beta^2}} \sin\left(\frac{2t}{\beta}\right) \right], \quad (14)$$

where $f(t) *_t g(t) := \int_{-\infty}^{\infty} ds f(s) g(t-s)$ denotes the convolution, and

$$g_2(t) = \frac{2\sqrt{2}}{\beta} \exp\left(\frac{(\beta - 4it)^2}{4\beta^2}\right). \quad (15)$$

Below, we consider operators in the form of single-qubit Pauli operators, specifically

$$A^{a,1} = X_a, \quad A^{a,2} = Y_a, \quad A^{a,3} = Z_a. \quad (16)$$

This specific choice of jump operators is motivated by theoretical convenience: in the limit $\beta \rightarrow 0$, the

resulting Lindbladian describes depolarizing noise (see Appendix C2). This simplification streamlines the ensuing theoretical analysis for deriving the error and performance-time bounds. Nevertheless, our framework can, in principle, accommodate a much broader class of initial operators—so long as they remain sufficiently generic (for example, noncommuting with the Hamiltonian). Also, both the compilation strategy described below and our numerical simulation methods permit arbitrary operator choices. By selecting these operators strategically, one can optimize algorithmic runtime or minimize circuit depth, tailoring them to a particular Hamiltonian or compilation scheme. A comprehensive exploration of these optimizations, however, lies beyond the scope of the present work.

The performance of quantum Gibbs samplers is characterized by their mixing time, defined as follows (e.g., see Proposition E.4 in Ref. [22]):

Definition 1. For a given Lindbladian $\mathcal{L}$, the mixing time is the smallest time t_{mix} satisfying

$$\forall \rho, \rho' \in \mathcal{D}_\Lambda : \quad \|e^{\mathcal{L}t_{\text{mix}}}(\rho - \rho')\|_1 \leq \frac{1}{2} \|\rho - \rho'\|_1. \quad (17)$$

The mixing time provides a natural way to define the runtime of the algorithm. Let ρ be the initial state and ρ' the fixed point of the channel $\mathcal{L}$. It follows by the definition of mixing time that the time t required for the system to reach ρ' within a trace distance of $\epsilon > 0$ is bounded by $t \leq t_{\text{mix}} \log_2(1/\epsilon)$. Therefore, estimating the mixing time is a central objective in our analysis.

Previously, it was demonstrated [31] that the quantum Gibbs sampler introduced in Ref. [22] exhibits a mixing time scaling linearly with the number of qubits n for temperatures below a threshold $\beta < \beta^*$, where β^* does not depend on system size. This result was obtained by establishing bounds on the Lindbladian's spectral gap, which, due to the KMS detailed-balance condition in Eq. (9), provides an upper bound for the mixing time. This result was improved in Ref. [23], which showed a logarithmic mixing time for high enough temperature, by bounding an oscillator norm. Using analogous techniques from Rouzé *et al.* [23], we demonstrate that the Lindbladian described by Ding *et al.* [26], namely Eq. (3), also achieves convergence to a Gibbs state with mixing time scaling logarithmically with the system size, and we extend this result to truncated dynamics that will be introduced later.

While we do not cover this in our work, one can also consider using similar methods that have shown that quantum Gibbs samplers effectively prepare thermal states in the Fermi-Hubbard model at arbitrary temperatures and weak interactions [24,25].

III. ANALYTICAL RESULTS

In this section, we outline the steps required for a hardware-efficient simulation of the Lindbladian described by Eq. (3). The implementation procedure consists of three main stages: (a) truncating the jump operators and the coherent term, (b) applying Trotterization to the resulting truncated Lindbladian, and (c) compiling the elementary quantum channels that implement Trotterized evolution for each Lindbladian component.

For the truncation step, we present formal results that establish rigorous error bounds on the truncation error. Specifically, we demonstrate that for sufficiently high temperatures, which are independent of the system size, the truncation error in evaluating both local observables and the trace distance can be controlled by increasing the truncation radius, as outlined in Corollaries 1 and 2 below. In particular, Gibbs sampling generally requires the truncation radius to scale logarithmically with the number of qubits, whereas estimating the expectation value of local observables requires only a constant radius.

Subsequently, we introduce a simulation protocol for the truncated Lindbladian based on the randomized compilation approach for open system dynamics [28]. This protocol allows us to approximate the truncated dynamics using a circuit consisting of local quantum channels, each of which implements the short time evolution of a truncated jump operator. The randomized approach reduces the depth of the circuit and makes it independent of the number of jump operators, at the cost of increased sampling overhead. The discussion of compilation of the circuit is deferred to Sec. IV.

A. Truncation of quasilocal Lindbladians

In this section, we analyze the error arising from replacing quasilocal operators with their locally truncated versions. Consider a Hamiltonian introduced in Eq. (2). Let us introduce a fixed truncation radius $r \geq 0$ and define the locally truncated Hamiltonian $H_{a,r}$ as the sum of all local terms whose supports are contained within the ball $B_a(r)$ of radius r centered around qubit a ,

$$H_{a,r} := \sum_{X \subseteq B_a(r)} H_X. \quad (18)$$

Using this construction, we define the truncated Lindbladian as

$$\mathcal{L}^{\beta,r} = \sum_{a \in \Lambda} \sum_{\alpha} \mathcal{L}_{a,\alpha}^{\beta,r}, \quad (19)$$

where

$$\begin{aligned} \mathcal{L}_{a,\alpha}^{\beta,r}(\rho) = & -i[G_{a,\alpha}^{\beta,r}, \rho] \\ & + L_{a,\alpha}^{\beta,r} \rho L_{a,\alpha}^{\beta,r\dagger} - \frac{1}{2} \{L_{a,\alpha}^{\beta,r\dagger} L_{a,\alpha}^{\beta,r}, \rho\}, \end{aligned} \quad (20)$$

expressed through the truncated coherent terms and jump operators

$$\begin{aligned} L_{a,\alpha}^{\beta,r} &= \sum_{v \in \Omega(H_{a,r})} q(v) (A^{a,\alpha})_{v[a,r]}, \\ G_{a,\alpha}^{\beta,r} &= -\frac{i}{2} \sum_{v \in \Omega(H_{a,r})} \tanh\left(-\frac{\beta v}{4}\right) (L_{a,\alpha}^{\beta,r\dagger} L_{a,\alpha}^{\beta,r})_{v[a,r]}. \end{aligned} \quad (21)$$

Here, the projector $A_{v[a,r]}$ is defined similarly to Eq. (8), but using the projectors and eigenstates of the truncated Hamiltonian $H_{a,r}$. This truncation procedure is summarized in Fig. 1. Consequently, the resulting truncated jump operators have compact support on $B_a(r)$.

The truncation introduces a controllable error in the resulting fixed point of the Lindbladian evolution, as established in the following theorem.

Theorem 1 (Truncation error). The truncated Lindbladian $\mathcal{L}^{\beta,r}$ in Eq. (19), with the choice of complete set of generators in Eq. (16), has a unique fixed point $\rho_{\beta,r} \in \mathcal{D}_\Lambda$, i.e.,

$$\mathcal{L}^{\beta,r}(\rho_{\beta,r}) = 0.$$

Moreover, there exist constants $\beta^* > 0$ and $J > 0$ such that, for any $\beta < \beta^*$

$$\|\rho_{\beta,r} - \rho_\beta\|_1 = O\left((\beta J)^{\frac{r}{2}} n \log n\right). \quad (22)$$

The proof of Theorem 1 is provided in Appendix D. To establish this result, we invoke Lemma 2 from Ref. [22], which provides an upper bound on the distance between the fixed points of two Lindbladians, which is at most four times of the product between: (a) mixing time t_{mix} of one of the Lindbladians and (b) the Schatten-1 induced norm $\|\mathcal{L}^\beta - \mathcal{L}^{\beta,r}\|_{1 \rightarrow 1}$, which remains bounded and controllable due to the (quasi)-locality of all terms appearing in each of the Lindbladians.

Consequently, to derive the bound on the mixing time, we extend the proof of rapid mixing from Rouzé *et al.* [23] devised for the process in the work by Chen *et al.* [22] to the algorithms we analyze, specifically the quantum Gibbs sampler in Eq. (3) and the truncated Lindbladian in Eq. (19). For the latter, we state the following result:

Theorem 2 (informal). Under conditions of Theorem 1, the Lindbladians $\mathcal{L}^\beta$ and $\mathcal{L}^{\beta,r}$, with $r \geq 0$ have mixing time

$$t_{\text{mix}} = O(\log n). \quad (23)$$

This proof builds on showing that the Lindbladian, similar to the one introduced in Chen *et al.* [22], for all $O \in \mathcal{M}_\Lambda$

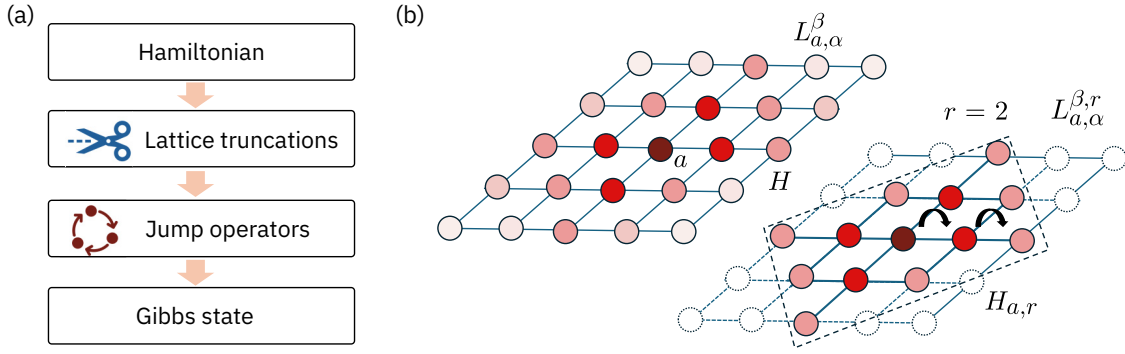


FIG. 1. Sketch of truncation-based dissipative Gibbs state preparation. (a) Starting from the target Hamiltonian H , we construct a family of local, truncated-lattice Hamiltonians $H_{a,r}$. From these truncated models, we derive the corresponding local Lindblad jump operators, which define global dissipative evolution that drives the system toward the desired Gibbs state. (b) Truncation schematic. *Left*: Jump operator derived from the full-system dynamics on Λ , chosen to be a square lattice; only a portion of the lattice is shown. The action of the jump operator decays exponentially with the graph (Manhattan) distance from the central qubit $a \in \Lambda$. *Right*: Jump operator constructed by restricting the dynamics to qubits within a graph radius r (here, $r = 2$), yielding strictly local support.

satisfies the inequality [23]

$$\|e^{t\mathcal{L}^{\beta\dagger}}(O)\| \leq e^{-(1-\kappa)t} \|O\| \quad (24)$$

for some $\kappa > 0$ with the oscillator norm defined as

$$\|O\| := \sum_{a \in \Lambda} \|O - \frac{1}{2}I_a \otimes \text{tr}_a(O)\|_\infty,$$

where $\text{tr}_a(\cdot)$ is partial trace over qubit a . From that statement, it follows (cf. Appendix D)

$$\|e^{t\mathcal{L}^\beta}(\rho) - \rho_\beta\|_1 \leq 2n\|\rho - \rho_\beta\|_1 e^{-(1-\kappa)t}. \quad (25)$$

It turns out that κ can be bounded by exploiting the quasilocality of the Lindbladian, provided the temperature is sufficiently high, i.e., $\beta < \beta^*$. Crucially, since the proof relies on the locality of the jump operators, the bound on the mixing time for the original Lindbladian $\mathcal{L}^\beta$ is always greater than or equal to that of the truncated Lindbladian $\mathcal{L}^{\beta,r}$. In particular, truncation does not worsen the mixing time bounds at high enough temperatures.

As a direct corollary of the established results so far, Theorem 1 provides an estimate for the required truncation radius r . In particular, it asserts that:

Corollary 1. For any fixed $\beta < \beta^*$, the truncated Lindbladian $\mathcal{L}^{\beta,r}$ has a fixed point $\rho_{\beta,r}$ that is ϵ -close to the Gibbs state ρ_β , provided that

$$r = \Omega\left(\log\left[\frac{n}{\epsilon}\right]\right). \quad (26)$$

The linear scaling with the number of qubits in Eq. (22) arises from the general subadditivity of the trace distance [3] and is unlikely to be improved if the goal is to sample

from the exact Gibbs distribution. However, in many physical applications, the primary objective is to compute the expectation values of local observables. For this task, the requirements on the truncation radius are more relaxed as one can use restrictions such as the Lieb-Robinson bound. The result is demonstrated by the following theorem.

Theorem 3 (Truncation error for local observables). Under the assumptions of Theorem 1, consider $\beta < \beta^*$ and let $O_X \in \mathbf{S}_X$ be supported on $X \subseteq \Lambda$. Then there exist constants $\gamma, \eta > 0$, depending only on the model parameters and β , such that $\rho_{\beta,r}$ satisfies

$$\begin{aligned} & |\text{tr}(\rho_{\beta,r} O_X) - \text{tr}(\rho_\beta O_X)| \\ & \leq \|O_X\| c(|X|) [e^{-\gamma r} + n e^{-\eta d_X}], \end{aligned} \quad (27)$$

where $c(|X|) = O(\text{poly}(|X|))$, $n \equiv |\Lambda|$ is the number of qubits, and

$$d_X = \min_{a \in X} \ell(a, \partial\Lambda)$$

is the minimal graph distance from any $a \in X$ to the boundary $\partial\Lambda$.

For a fixed region X , the distance d_X , which is distance between X and to the complement Λ^c of the lattice Λ and is thus the distance to the boundaries, increases with system size, which implies that the term $|\Lambda|v_\eta^{-1}(d_X)$ vanishes as the system becomes large. This term can therefore be interpreted as a finite-size correction [35]. The proof of Theorem 3 is detailed in Appendix E. This proof relies on a theorem concerning the stability under local perturbations of Lindbladians with logarithmic scaling of the mixing time, as derived in Ref. [35]. In fact, this stability theorem allows us to prove an even stronger statement. Specifically, consider the time evolution generated by $\mathcal{L}^\beta$, $\exp(t\mathcal{L}^\beta)$, and the time evolution $\exp(t\mathcal{L}^{\beta,r})$. We show that for any

observable O_X supported on X , the distance between $\exp(t\mathcal{L}^{\beta,\dagger})[O_X]$ and $\exp(t\mathcal{L}^{\beta,r,\dagger})[O_X]$ remains small for *all* times t . The result then follows from considering the limit $t \rightarrow \infty$. We emphasize that the support X does not necessarily have to be connected, thus this theorem also applies to correlation functions. Also, note that the bound of $c(|X|)$ by a polynomial is useful when $|X|$ is small, but it becomes useless for global observables.

A direct corollary of Theorem 3 is:

Corollary 2. For any $\beta < \beta^*$, the truncated Lindbladian $\mathcal{L}^{\beta,r}$ admits a fixed point for which the expectation values of observables whose support size is $O(1)$ and sufficiently distant from the lattice boundary are ϵ -close to those of the Gibbs state ρ_β , provided that

$$r = \Omega(\log(1/\epsilon)). \quad (28)$$

Notably, this requirement is independent of the system size and applies to observables located far from the lattice boundaries, where the last term in Eq. (27) can be neglected; that is, when $d_X = \Omega(\log n)$.

In addition to the analytical results above, we also present in Appendixes V and G 1 numerical experiments aimed at evaluating the error in expectation values of local observables in certain one-dimensional models. Our findings indicate that the error remains modest even at low temperatures (large inverse temperatures β), a regime where the aforementioned theoretical bounds may not strictly apply. Consequently, these numerical results suggest that the practical validity of truncation methods extends beyond the analytically predicted limits, at least for one-dimensional systems.

B. Trotterization of truncated dynamics

After deriving the truncated Lindblad operator and demonstrating its efficiency in preparing Gibbs states, the subsequent step toward practical implementation is to discretize the evolution using Trotterization. Consider an evolution over total time $t = \tau M$, where M is a large integer. Then, we can partition the continuous evolution into discrete time steps of length τ and approximate the dynamics by

$$\exp(\mathcal{L}^{\beta,r}t) \rightarrow \left[\prod_{a \in \Lambda} \exp\left(\sum_{\alpha} \mathcal{L}_{a,\alpha}^{\beta,r} \tau\right) \right]^M, \quad (29)$$

where $\prod_{a \in \Lambda}$ denotes a composition of channels applied in an arbitrary but fixed order, see Fig. 2(a). This method generalizes the Trotterization approach commonly used in Hamiltonian simulations and is known to introduce an error scaling quadratically with time t [36,37], as we will see later in this Section.

To further reduce circuit depth, we introduce a randomized compilation method tailored specifically for Lindblad dynamics at each discrete time step [28]. We define a simulation trajectory

$$\mathcal{E}_{t,\tau}^{\alpha} := \prod_{i=1}^M \prod_{a \in \Lambda} \exp\left(\mathcal{L}_{a,\alpha[a]}^{\beta,r} \tau\right), \quad (30)$$

where $\alpha[a]$ represents a single choice of operator over site a , $\alpha = \{\alpha[a] \mid a \in \Lambda\}$. Next, we consider the index $\alpha[a]$ to be selected independently and uniformly at random, thus inducing a uniform probability distribution $\alpha[a] \sim \Sigma$. Then we obtain that for sufficiently large M (see

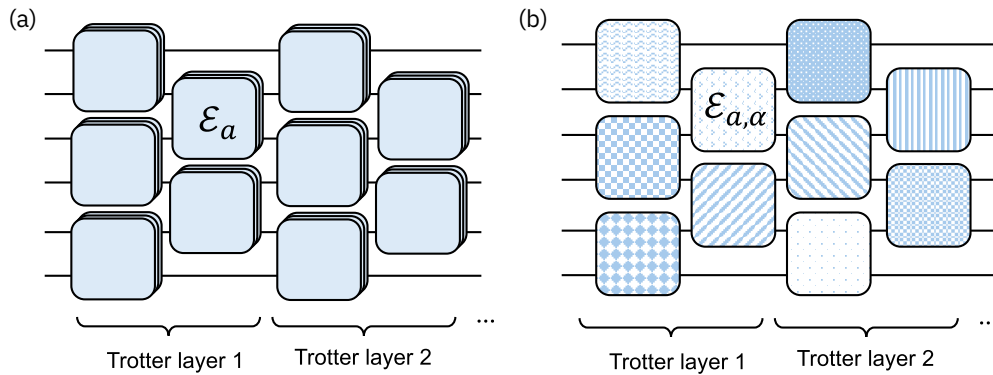


FIG. 2. Standard vs randomized Trotterization. (a) Standard Trotterization approximates continuous-time evolution by a sequence of discrete operations $\mathcal{E}_a := \exp\left(\sum_{\alpha} \mathcal{L}_{a,\alpha}^{\beta,r} \tau\right)$, where τ is the Trotter step size. Each operation is visualized as a stack of gates representing the sum in the exponent. (b) To reduce circuit depth, we employ a randomized compilation strategy: at each Trotter step and for each site a , we randomly select a single term $\exp\left(\mathcal{L}_{a,\alpha}^{\beta,r} \tau\right)$ from Eq. (30) by drawing α uniformly from the set of possible values. These randomized operations correspond to shallower circuits derived from the exponentiation in Eq. (31). Averaging over many such randomized circuit realizations recovers the full dynamics.

Theorem 4 below),

$$\exp(\mathcal{L}^{\beta,r}t) \approx \mathbb{E}_{\alpha \sim \Sigma^{\otimes n}} [\mathcal{E}_{t,\tau}^{\alpha}], \quad (31)$$

where the RHS represents the average of the randomized, discretized evolution with time step τ . Practically, this evolution is realized by randomly selecting Lindbladian terms indexed by different values of α at each circuit layer and averaging over the resulting quantum trajectories, as illustrated in Fig. 2(b).

The following results shows that for sufficiently large M , the output of these expectation values converges to the expectation value of the target evolution.

Theorem 4 (Theorem 4 from Ref. [28]). The error of the average channel $\mathcal{E}_{t,M}$ in Eq. (30) can be bounded by

$$\left\| \mathbb{E}_{\alpha \sim \Sigma^{\otimes n}} [\mathcal{E}_{t,\tau}^{\alpha}] - e^{t\mathcal{L}^{\beta}} \right\|_{\diamond} = O\left(\frac{n^2 t^2}{M}\right). \quad (32)$$

The proof of this theorem is in Ref. [28].

IV. CIRCUIT COMPILATION

The final step required for an implementation of the Trotterized dynamics that is suitable for near-term quantum hardware is a compilation of the local quantum channels into circuit-level gate operations. In principle, this can be done efficiently using dilation methods, followed by compilation into the desired gate set. According to standard results in the literature [3], for a fixed r , this compilation requires a number of gates that scales at most polylogarithmically with the inverse of the additive error.

Here we show that a single-ancilla qubit is in fact sufficient, and that variational compilation of the resulting dilated unitary provides a practically feasible method for yielding sufficiently short circuits for near-term implementation. More specifically, we provide a compilation procedure for each channel $\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})$, which represents the randomized evolution described in Eq. (30). This can be done by first defining the Hermitian operator

$$O_r^{a,\alpha} := \begin{pmatrix} \sqrt{\tau} G_r^{a,\alpha} & L_r^{a,\alpha\dagger} \\ L_r^{a,\alpha} & \sqrt{\tau} G_r^{a,\alpha} \end{pmatrix}, \quad (33)$$

where $L_r^{a,\alpha}$ and $G_r^{a,\alpha}$ are the jump operators and associated coherent terms, respectively, as was previously defined in Eq. (21).

Then we can approximate (see Lemma 1)

$$\exp(\mathcal{L}_{a,\alpha}^{\beta,r} \tau) \approx \mathcal{C}_{r,\tau}^{a,\alpha}(\rho), \quad (34)$$

where we introduced the quantum channel

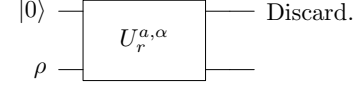
$$\mathcal{C}_{r,\tau}^{a,\alpha}(\rho) := \text{Tr}_{\text{anc}} (U_r^{a,\alpha}(\rho \otimes |0\rangle\langle 0|_{\text{anc}}) U_r^{a,\alpha\dagger}), \quad (35)$$

where $\text{Tr}_{\text{anc}}(\cdot)$ denotes the partial trace over the ancilla qubit and the unitary $U_r^{a,\alpha}$ describes the evolution over a

time interval $\sqrt{\tau}$ as

$$U_r^{a,\alpha} = \exp(-iO_r^{a,\alpha} \sqrt{\tau}). \quad (36)$$

The channel described by Eq. (34) can be implemented using the following procedure:



This subcircuit acts on a single-ancilla qubit initialized in the state $|0\rangle$ and the subsystem supporting the jump operators. In architectures with a limited number of ancilla qubits, reset operations can be utilized to obtain a fresh supply of ancilla qubits.

This single-ancilla based construction can be used to approximate the Trotter steps up to an additional error comparable to the Trotter error itself, as follows from the following lemma.

Lemma 1. The channel in Eq. (34) satisfies

$$\left\| \mathcal{C}_{r,\tau}^{a,\alpha} - \exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r}) \right\|_{\diamond} = O(\tau^2). \quad (37)$$

This result was previously shown without the coherent part in Ref. [28], while a detailed proof of the lemma is provided in Appendix F. The proof is based on a standard Taylor expansion argument [38]. Notably, the error in Eq. (37) can be further reduced to $O(\tau^k)$ using k ancilla qubits [39]. As a corollary, the compilation of channels $\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})$ reduces to compilation of the unitaries in Eq. (36).

Compiling the unitary involves designing a circuit for a small subsystem of a D -dimensional lattice, encompassing $\Theta(r^D)$ qubits. According to Corollary 1, constructing a circuit that exactly reproduces the Gibbs distribution requires polynomial classical resources in one dimension, but grows quasipolynomially with dimension for $D > 1$. In contrast, when the goal is to estimate local observables, Corollary 2 shows that local channels can be constructed using only a constant number of qubits. The latter results in a highly efficient compilation process from a complexity-theoretic standpoint.

However, from a practical perspective, the runtime of standard compilers may be prohibitive, and we may want even shorter circuits, tailored to a specific architecture and gate set. To achieve this, we can *variationally* compile the necessary local unitaries resulting from dilation. Specifically, we approximate the target unitary $U_r^{a,\alpha}$ with a parameterized template circuit $V(\theta_r^{a,\alpha})$, where $\theta_r^{a,\alpha}$ denotes the set of variational parameters defining the ansatz circuit [40,41]. An example of such a circuit for a one-dimensional system with $r = 1$ is shown in Fig. 3(a). This

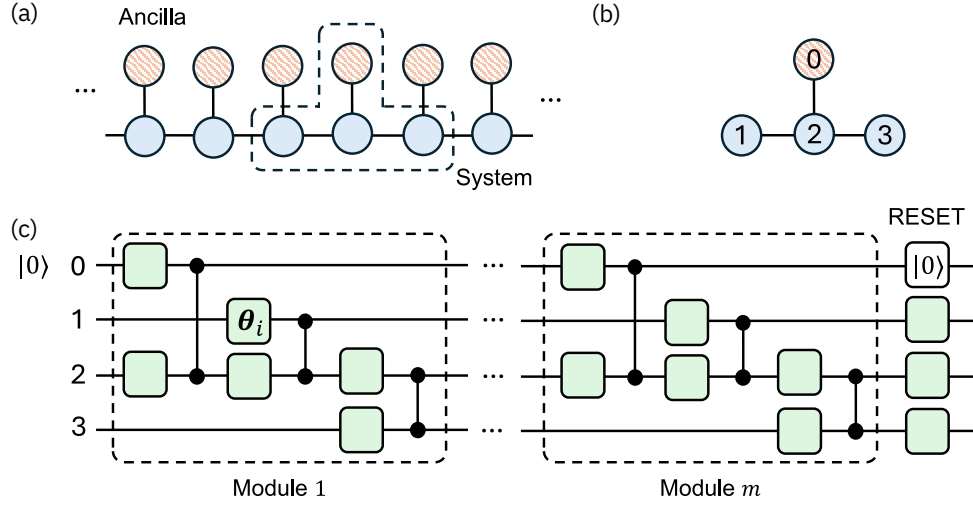


FIG. 3. Gadgets for one-dimensional Hamiltonians. (a) The “ladder” device architecture, shown as an example of a local layout for simulating thermal states of one-dimensional Hamiltonians. The lower row of qubits (solid blue) represents the system qubits, while the upper row (dashed red) represents the ancilla. (b) The gadget connectivity for a truncation radius of $r = 1$, which requires implementing a four-qubit unitary. (c) A template circuit example, consisting of m individually optimized modules followed by single-qubit gates. The ancilla is initialized in the state $|0\rangle$ before the gadget is applied and is reset afterward to ensure correct initialization for the next gadget. Each module consists of three controlled-Z gates arranged in an alternating pattern, interleaved with single-qubit gates parametrized as in Eq. (38).

ansatz employs entangling two-qubit CZ gates interleaved with parameterized single-qubit gates of the form

$$u(\theta, \phi, \lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda} \sin\left(\frac{\theta}{2}\right) \\ e^{i\phi} \sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)} \cos\left(\frac{\theta}{2}\right) \end{pmatrix}, \quad (38)$$

where θ , ϕ , and λ are free parameters comprising the vector $\theta_r^{a,\alpha}$. Further decomposition of these single-qubit gates into native hardware-specific operations can be performed straightforwardly.

To determine the optimal parameters, we use minimization

$$\theta_r^{a,\alpha} = \arg \min_{\theta} C_r^{a,\alpha}(\theta) \quad (39)$$

of the loss function $C_r^{a,\alpha}(\theta)$ that quantifies the discrepancy between the target and the implemented quantum channels. For systems with translational symmetry, it is sufficient to optimize the resulting operation for each unique local configuration, drastically reducing the overall number of independent optimization tasks. The choice of loss function is important to reduce classical simulation overhead.

In this work, we consider a modified squared-Frobenius norm for the relevant matrix elements,

$$C_r^{a,\alpha}(\theta) := \sum_i \sum_{j \leq 2^{k-1}} \left| [U_r^{a,\alpha}]_{ij} - [V(\theta)]_{ij} \right|^2, \quad (40)$$

where A_{ij} corresponds to the matrix element of matrix A and k is the total number of qubits in the gadget. Importantly, it is not necessary to match the entire unitary. Since

the ancilla qubit is always initialized in the state $|0\rangle$ and discarded at the end, only the corresponding subspace of the unitary contributes to the final outcome. We justify this choice of loss function more in Appendix G 4.

In general, representing an arbitrary quantum channel requires an exponential number of parameters in Eq. (39) as a function of the radius r , which could in principle make the optimization costly. However, one may note that the channels appearing in Eq. (37), for small finite evolution times τ , admit accurate approximations via a Magnus expansion in terms of jump operators in Eq. (21). Since each jump operator can be generated with a combination of linear-depth Hamiltonian simulation circuits by construction, each jump operator and the channel itself exhibit polynomial rather than exponential complexity in terms of radius r . We therefore expect that, in practice, optimization landscapes are substantially more benign than in the fully generic case. This expectation is corroborated by the results in the following section, which demonstrate that comparatively low-depth circuits—significantly shallower than those typically required for generic channels—are sufficient to capture the relevant dynamics.

V. NUMERICAL STUDIES

In this section, we evaluate the accuracy of the proposed Gibbs state preparation method using exact numerical simulations of two types of nonintegrable Hamiltonians: mixed-field Ising spin chain and two-dimensional transverse-field Ising model. Our results indicate that a

relatively small truncation radius ($r \leq 3$) and a moderate number of Trotter steps (up to 100) are sufficient to ensure rapid convergence of local observables, such as energy density and correlation functions, to their thermal expectation values as determined by exact diagonalization. These findings demonstrate that the method operates within the capabilities of what is commonly referred to as noisy intermediate-scale quantum devices. In the following section, we will examine how various noise levels affect these simulations.

The Hamiltonian of the mixed-field Ising spin chain is

$$H = \sum_{(i,j) \in \mathcal{S}_1} S_i^z S_j^z + g \sum_{i=1}^n S_i^x + h \sum_{i=1}^n S_i^z, \quad (41)$$

where $\mathcal{S}_1$ is a set of nearest neighbors on a one-dimensional ring with periodic boundary conditions, $S_i^x := \frac{1}{2}X_i$ and $S_i^z := \frac{1}{2}Z_i$ are spin-1/2 operators. We choose $g = (\sqrt{5} + 5)/8 \approx 0.9045$ and $h = (\sqrt{5} + 1)/4 \approx 0.809$. These parameters ensure that no single term dominates the energy spectrum, placing the system deep in the nonintegrable (chaotic) regime, with ballistic operator spreading and rapid entanglement growth [42]. Such strong transport dynamics aims to shorten the thermalization time. In Appendix G 1, we also show that our truncated Lindbladian method remains effective for examples of integrable models and systems with $U(1)$ symmetry.

The two-dimensional transverse-field Ising model is given by

$$H = \sum_{(i,j) \in \mathcal{S}_2} S_i^z S_j^z + g' \sum_{i=1}^n S_i^x, \quad (42)$$

where $\mathcal{S}_2$ contains the nearest neighbors on a square lattice with periodic boundary conditions. We choose $g' = 0.2$. The model is known to exhibit a thermal phase transition [43].

We choose the details of the dynamics as follows. First of all, we study quantum Gibbs state preparation starting from a random product state. This choice is equivalent to setting $\rho(0) = \rho_{\beta=0}$, i.e., as the maximally mixed, infinite-temperature state. Thus, one may consider it as an energy-reducing, “cooling” process. The nonunitary part of dynamics is generated by the transformed Pauli X , Y , and Z operators acting on each qubit [see Eq. (16)] constructed following Eq. (21) for a fixed truncation radius r . Unless stated otherwise, we use a Gaussian envelope function $q(v)$ defined in Eq. (11).

As the target of our numerical study, we analyze how different parameter choices (e.g., Trotter step or truncation radius) impact the performance of the algorithm. As the main precision metric, we use the Gibbs state energy

density and its deviation from the true value,

$$\begin{aligned} E(t) &:= \frac{1}{n} \text{tr}(\rho(t)H), \\ \Delta E(t) &:= \frac{1}{n} |\text{tr}(\rho(t)H) - \text{tr}(\rho_{\beta}H)|, \end{aligned} \quad (43)$$

where $\rho(t)$ is the result of the evolution

$$\rho(t) = \mathcal{E}_{t,\tau}(\rho(0)), \quad (44)$$

with $\mathcal{E}_{t,\tau}$ defined in Eq. (31) for a fixed Trotter step τ , and time taking discrete values $t = M\tau$.

We also analyze the algorithm’s performance in reproducing other local observables. To test this, we consider the correlation function

$$\delta(\mathbf{a}, t) := \text{tr}[\rho(t)S_{a_1}^z S_{a_2}^z] - \text{tr}[\rho(t)S_{a_1}^z] \text{tr}[\rho(t)S_{a_2}^z], \quad (45)$$

where $\mathbf{a} = (a_1, a_2)$. This function characterizes the emergence of spin-spin correlations within the system. Since the one-dimensional mixed-field Ising model does not exhibit an equilibrium ferromagnetic phase, we expect the correlator to decay exponentially with distance, i.e., $\delta(\mathbf{a}, t) \sim \exp(-\ell(a_1, a_2)/l_0)$, where $\ell(a_1, a_2)$ is the spatial separation between a_1 and a_2 and $l_0 > 0$ denotes the correlation length.

In order to detect the phase transition in the two-dimensional model, we analyze the heat capacity

$$C_{\beta} := \beta^2 (\text{tr}[\rho(t)H^2] - \text{tr}[\rho(t)H]^2). \quad (46)$$

In Secs. VA–VC, we study the performance of the algorithm without accounting for compilation errors. In Secs. VIA and VIB, we further compile each gadget into two-qubit gates, assuming the architecture illustrated in Fig. 3. In all cases, we compare the results to numerical values obtained via exact diagonalization, i.e., the “ground truth.”

A. Effect of truncation

Figure 4 shows how our protocol cools the system from an infinite-temperature initial state to thermal states at two values of temperature using a Trotter step $\tau = 0.1$. At the higher temperature ($\beta = 1$), the internal energy already achieves best possible value with the relative error $\Delta E(t)/|E(t)| \sim 10^{-2}$ for the minimal truncation radius $r = 1$, and increasing r yields no noticeable improvement. By contrast, at the lower temperature ($\beta = 3$), the convergence depends strongly on r . We attribute this behavior to the enhanced spin-spin correlations in the mixed-field Ising model, which become more significant at lower energies. At the same time, the cooling rate, as measured by the slope of the exponential decay at early times, is largely insensitive to the truncation radius. A bigger picture across

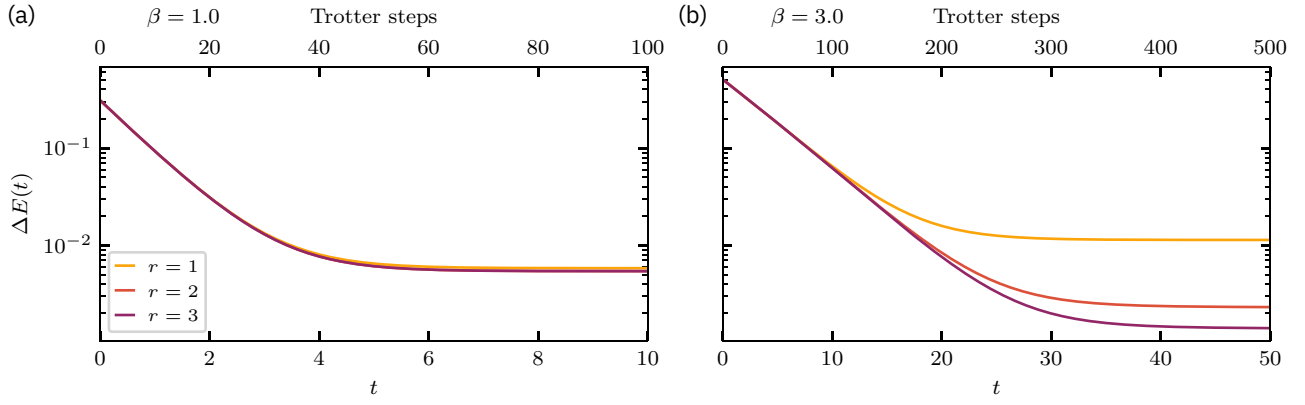


FIG. 4. Effect of truncation on time convergence. Absolute error in the internal energy density $\Delta E(t)$ in Eq. (43) for the mixed-field Ising Hamiltonian in Eq. (41) on a chain of $n = 12$ spins. Dynamics are generated by direct mixed-state simulations of the Trotterized channel (29) with Trotter step $\tau = 0.1$, for a system initialized in the maximally mixed state. The x -axis markers on the bottom denote the physical time t , while those at the top indicate the number of Trotter steps $M = t/\tau$. Curves correspond to truncation radii $r = 1$ (yellow), $r = 2$ (orange), and $r = 3$ (red). (a) Inverse temperature $\beta = 1$. (b) Inverse temperature $\beta = 3$.

temperatures is shown in Fig. 5. These plots illustrate that the energy convergence is nonmonotonic across temperatures but rapidly converges with respect to truncation radius at low temperatures.

The rapid convergence with respect to the truncation radius is further demonstrated by local observables other than energy density. In Fig. 6, we present the two-point correlation function at $\beta = 3$. While the minimal truncation radius ($r = 1$) yields noticeable deviations from the expected spatial profile, extending the radius to $r = 3$ suffices to accurately reproduce the correct space correlation profile.

These plots indicate that the radius r can be employed as a systematic control parameter, which is increased until the observable of interest converges to a stable value. The

results further show, as expected, that lower temperatures require larger values of r to achieve convergence.

B. Effect of Trotterization

A critical parameter in our protocol is the choice of the Trotter time step, τ . As shown in Fig. 7, for parameter regimes relevant to near-term quantum processors ($\beta = 1$, $r = 1$), the Trotterization error constitutes the dominant contribution to the total simulation error. By reducing τ one can suppress this error by multiple orders of magnitude; intriguingly, however, over the range of τ values we examined the error does not display clear convergence even up to $\tau \sim 10^{-3}$, implying that truncation artifacts remain subdominant. Unfortunately, a smaller τ

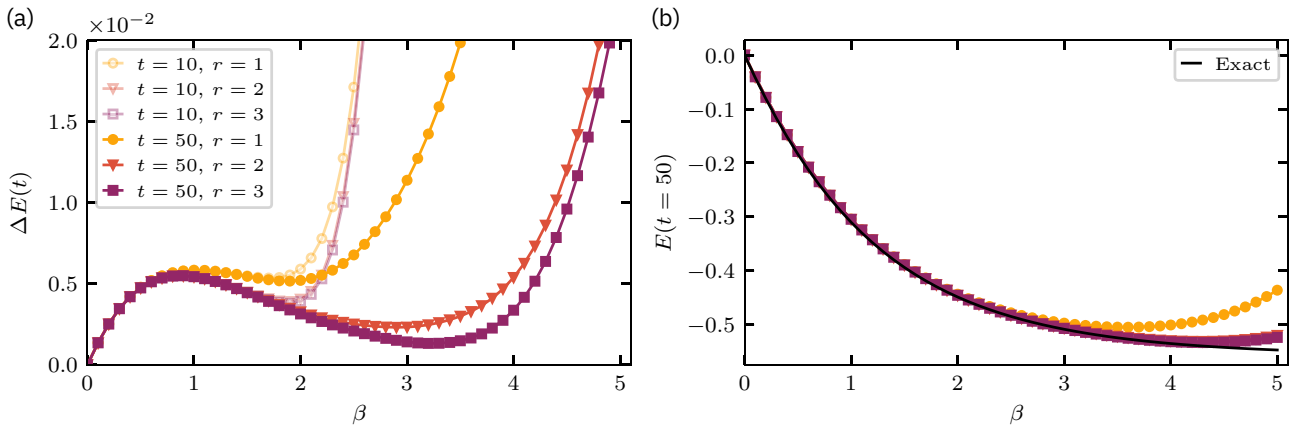


FIG. 5. Energy convergence across temperatures. Absolute error in the energy density $\Delta E(t)$ in Eq. (43) for the same simulation as in Fig. 4, plotted as a function of inverse temperature β . (a) Error $\Delta E(t)$ at $t = 10$ and $t = 50$ for truncation radii $r = 1$ (yellow), $r = 2$ (orange), and $r = 3$ (red). The pronounced rise at large β reflects mixing times that exceed $t_{\text{mix}} > 50$. (b) Absolute energy density for $t = 50$ and various truncation radii demonstrating convergence toward the true energy density (black curve) that reaches ground-state value $E_0/n = -0.557$ in the low-temperature (high- β) regime.

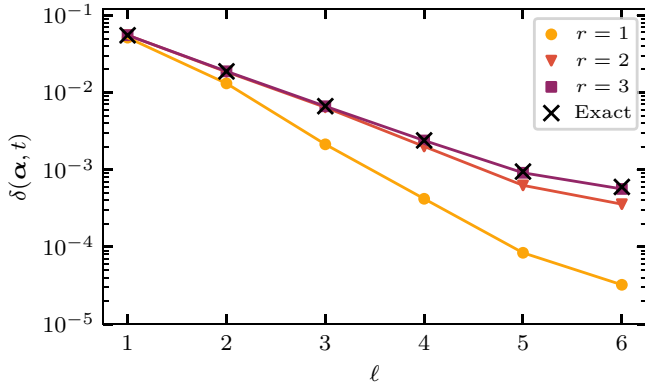


FIG. 6. Spatial profile of the two-spin correlators. Two-point correlator $\delta(\mathbf{a}, t)$ in Eq. (45) for spin pair $\mathbf{a} = (n/2, n/2 + \ell)$, as a function of spin separation ℓ for the mixed-field Ising chain in Eq. (41) at inverse temperature $\beta = 3$ and time $t = 10$. Curves correspond to truncation radii $r = 1$ (yellow), $r = 2$ (orange), and $r = 3$ (red), while black crosses denote the exact results. The accuracy of the exponential decay of correlations improves with truncation radius.

necessarily increases circuit depth, which is severely constrained on noisy hardware without full fault tolerance. We speculate that intermediate, early fault-tolerant architectures—where available qubit counts are the primary limitation, rather than circuit depth—will provide an ideal testbed for these Gibbs-state preparation methods.

C. Effect of the envelope function choice

The envelope function $q(v)$ in Eq. (6) provides an additional degree of freedom for tuning the precision of the Gibbs-state preparation protocol. Aside from the constraint $q(-v) = q^*(v)$, no further analytic restrictions apply, yet the choice of $q(v)$ markedly affects performance. Here, we

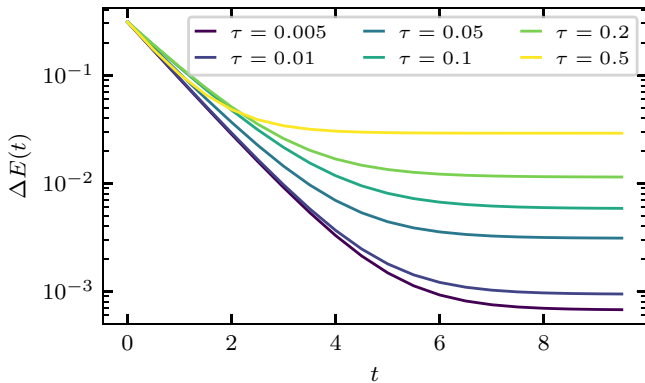


FIG. 7. Trotter step dependence. Error in the thermal-state preparation protocol for the mixed-field Ising Hamiltonian [Eq. (41)] at inverse temperature $\beta = 1$, plotted as a function of the Trotter time step τ for truncation radius $r = 1$. The rapid decrease in error with smaller τ confirms that Trotterization is the primary source of protocol inaccuracy.

compare the default Gaussian envelope of Eq. (6) against two alternatives,

$$q(v) = 1, \quad (47a)$$

$$q(v) = \exp\left(-\frac{\sqrt{1 + (\beta v)^2}}{4}\right), \quad (47b)$$

where Eq. (47b) represents a smoothed Metropolis-Hastings filter introduced in Ref. [26], and Eq. (47a) is a constant (flat) envelope. In each case, we additionally renormalize the jump operators to preserve their Frobenius norm relative to the Gaussian filter (see Appendix G 2 for details). Although a flat filter formally leads to a divergent filter function $f(t)$, the jump operators and coherent terms in Eq. (21) remain well defined operator on the support of $H_{a,r}$. The numerical tests on local observables show that it surprisingly yields the highest precision among the three envelopes.

In Fig. 8 we compare the cooling dynamics and residual errors: the flat envelope $q(v) = 1$ delivers the fastest initial cooling and the smallest asymptotic error, whereas the smoothed Metropolis-Hastings envelope exhibits the slowest convergence and largest errors, with this discrepancy growing at larger inverse temperatures, see Fig. 8(b). These findings indicate that the envelope choice is a powerful lever for optimizing Gibbs-state preparation and merits a systematic study in future work.

D. Two-dimensional system

Finally, we present results for the two-dimensional transverse-field Ising model defined in Eq. (42), demonstrating that our protocol is capable of reproducing thermal phase transitions in finite systems. Specifically, we compute the heat capacity C_β [cf. Eq. (46)] for simulations on 3×3 and 4×4 lattices. For the 4×4 system, we numerically implement the protocol by averaging over 16 384 stochastic trajectories. We fix the truncation radius to $r = 1$, choose a Trotter step size $\tau = 0.1$, and evolve the system up to a final time $t = 20$.

In these simulations, we empirically find that replacing the commonly used filter function in Eq. (11) with the temperature-independent Gaussian filter

$$q(v) = \exp\left(-\frac{v^2}{2}\right) \quad (48)$$

leads to significantly improved convergence toward the exact results at low temperatures. The underlying intuition is that the filter function must retain frequency components comparable to the Hamiltonian gap, which may lie outside the passband of standard temperature-dependent filters in the low-temperature regime.

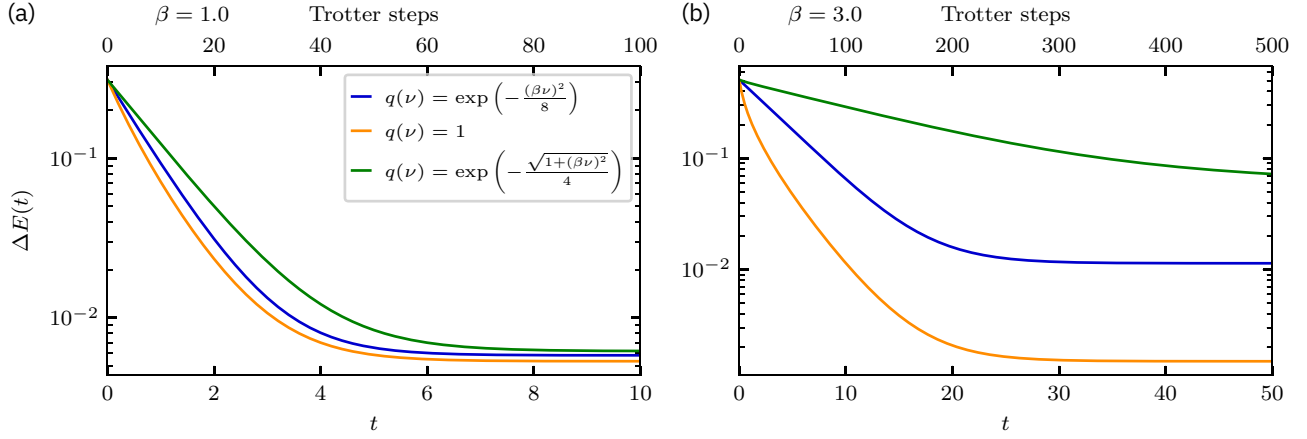


FIG. 8. Impact of the envelope function. Dynamics of our thermal-state preparation protocol with Trotter time step $\tau = 0.1$. For this plot only, jump operators and coherent terms in all cases are renormalized to Frobenius norm of the Gaussian filter for a consistent comparison, see Appendix G 2. (a) At inverse temperature $\beta = 1$, the flat envelope $q(\nu) = 1$ achieves the fastest convergence and lowest steady-state error. (b) At $\beta = 3$, the performance gap widens, highlighting that an optimal choice of $q(\nu)$ can dramatically accelerate the performance of the Gibbs sampler.

The resulting heat-capacity data are shown in Fig. 9. The presence of a broad peak in C_β signals a crossover between the ferromagnetic and paramagnetic phases. These results

demonstrate that the presence of a phase-transition boundary does not pose a fundamental limitation to the present algorithm.

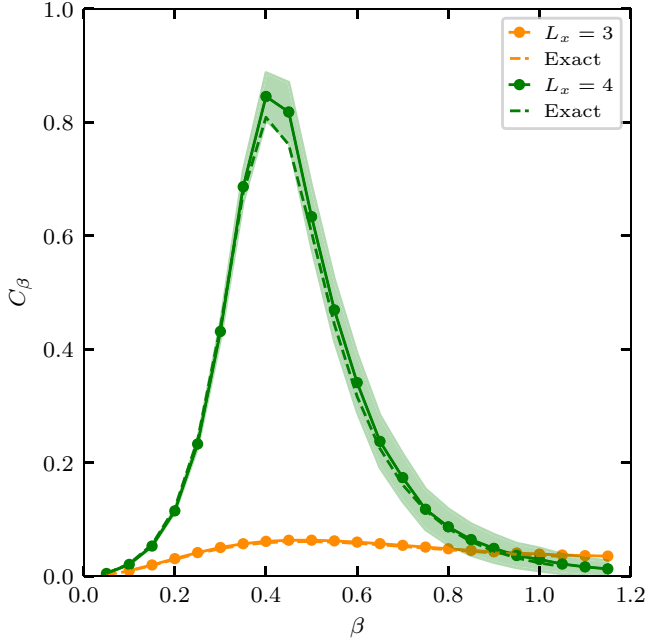


FIG. 9. Signatures of thermal phase transition in 2D transverse-field Ising model. The heat capacity C_β [cf. Eq. (46)] as a function of inverse temperature β for the 3×3 (orange) and 4×4 (green) square lattices with periodic boundary conditions. Comparison between exact data (dashed lines) and simulation (dots, solid lines). The results for the 4×4 lattice were obtained using sampling over 16 384 state-vector simulations, and the shaded region indicates one standard deviation. The results show that our protocol is able to reproduce the peak in the heat capacity for finite systems.

VI. HARDWARE-AWARE SIMULATIONS

In Sec. V we characterized the performance of the exact Trotterized channels, neglecting any overhead from compilation and hardware noise. In this section, we incorporate realistic circuit-compilation constraints and noise models to assess their impact on our Gibbs-state preparation protocol.

A. Effect of circuit compilation

Our objective is to show that thermalizing Lindbladian evolution can be approximated with a handful of circuits with a manageable two-qubit gate count compatible with the current noisy quantum devices.

The hardware compilation procedure proceeds as follows. We begin by approximating the Lindblad evolution via Eq. (31). Each of these circuits is constructed by substituting every elementary exponential in Eq. (30) with its approximate representation given in Eq. (34). The unitaries $U_r^{a,\alpha}$ are compiled using circuit templates illustrated in Fig. 3(b), where the parameters are optimized via the Adam algorithm [44] to minimize the loss function defined in Eq. (39). We investigate circuit architectures comprising $m = 2, 4, 6$, and 8 modules, which correspond to two-qubit gate depths of $d = 6, 12, 18$, and 24, respectively. For each configuration, the optimization is performed over 50 random parameter initializations, with 8000 optimization steps per instance. The parameters yielding the lowest final loss are selected for compilation. Further implementation details are provided in Appendix G 3.

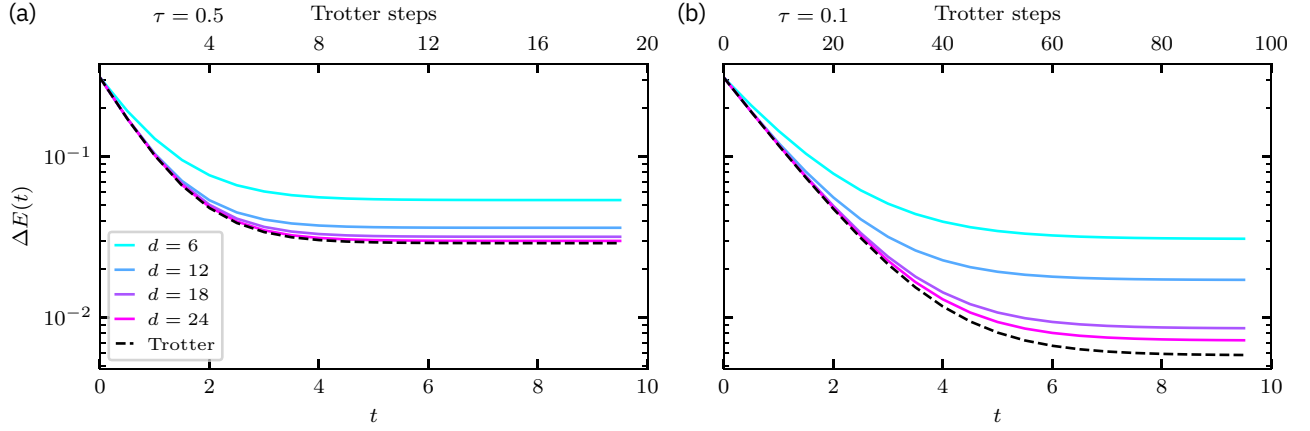


FIG. 10. Compilation error. Error in circuit implementation using the randomized compilation strategy described in Fig. 2(b), where each elementary operation is approximated using the template from Fig. 3(b). Results are shown for circuit depths $d = 6, 12, 18$, and 24 , corresponding to $m = 2, 4, 6$, and 8 modules, respectively. The results are compared to Trotterized evolution with no compilation error (dashed curve). Simulations are performed at inverse temperature $\beta = 1.0$ and final time $t = 10$. Panels show results for Trotter step sizes: (a) $\tau = 0.5$ and (b) $\tau = 0.1$. The results show that for $\tau = 0.5$, a depth of $d = 12$ suffices to render the compilation error negligible compared to the Trotterization error. For the smaller step size $\tau = 0.1$, a depth of $d = 18$ yields an energy density error of approximately $\Delta E(t) \sim 10^{-2}$ on a 12-qubit system.

To isolate the effects of gate compilation, we compare the time dynamics of the compiled circuits to those of the ideal Trotterized evolution, which employs the exact unitary $U_r^{a,\alpha}$ in Eq. (36). We assess the performance of the thermal state preparation protocol at inverse temperature $\beta = 1$ for two Trotter step sizes, $\tau = 0.5$ and $\tau = 0.1$. The results, shown in Fig. 10, indicate that for $\tau = 0.5$, a depth of $d = 12$ suffices to render the compilation error negligible compared to the Trotterization error. For the smaller step size $\tau = 0.1$, a depth of $d = 18$ yields an energy density error of approximately $\Delta E(t) \sim 10^{-2}$ on a 12-qubit system. These depths are well within the capabilities of current experimental platforms [4] and are

significantly lower than the estimated lower bound of $d = 64$ for generic four-qubit gate synthesis [45].

B. Effect of noise

Finally, we analyze the influence of noise on the accuracy of the simulations. To emulate realistic conditions, we assume that each k -qubit unitary gate is followed by a k -qubit depolarizing noise channel,

$$\mathcal{N}_k(\rho) := (1 - p_k)\rho + \frac{p_k}{4^k - 1} \sum_{i=1}^{4^k - 1} P_i^{(k)} \rho P_i^{(k)}, \quad (49)$$

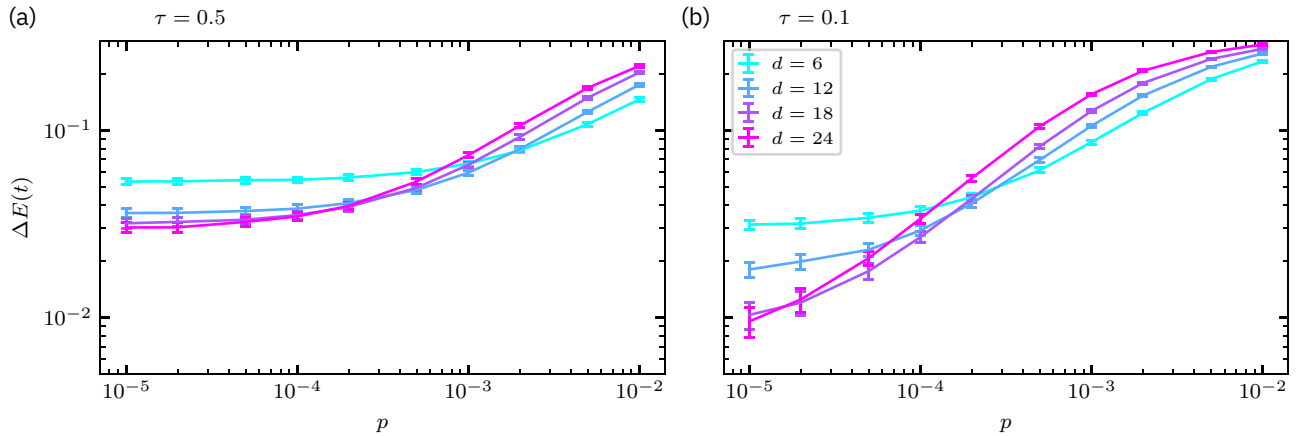


FIG. 11. Effect of noise. Accuracy of thermal-state preparation as a function of the depolarizing noise rate p , for inverse temperature $\beta = 1$ and evolution time $t = 10$. Error bars reflect the statistical error arising from 10^3 random circuits and 1024 shots per circuit. Results are shown for Trotter step sizes: (a) $\tau = 0.5$ and (b) $\tau = 0.1$. For target noise rates on the order of $p \sim 10^{-4}$, energy density errors on the order of $\sim 10^{-2}$ are achievable.

where $P_i^{(k)}$ denote the $4^k - 1$ nonidentity Pauli operators acting on k qubits, and p_k is the uniform k -qubit noise rate. In our simulations, we consider only one- and two-qubit gates, and set $p_1 = 0.1p$ and $p_2 = p$, where p is a global noise parameter. This reflects the empirical observation that two-qubit gates are typically the dominant source of error in current quantum hardware, whereas single-qubit gates have an order of magnitude less error.

Each data point in our analysis includes statistical uncertainty arising from limited number of 10^3 independently generated circuits corresponding to random trajectories $\alpha \sim \Sigma^{\otimes n}$ and a finite number of measurement shots, with 1024 shots per circuit. This amounts to approximately 10^6 shots per data point run on a hypothetical quantum processor.

The simulation results are presented in Fig. 11. We observe a distinct crossover in the accuracy as a function of the noise rate p , with the crossover point p_0 dependent on the Trotter step size τ . For noise levels $p > p_0$, the error is dominated by noise, and increasing the circuit depth d leads to worse performance. Conversely, in the regime $p < p_0$, the primary contribution to the error originates from imperfect compilation, which improves with increasing depth. This crossover behavior is transient: as the circuit depth increases, p_0 shifts toward lower values, consistent with the expectation that compilation error vanishes in the large-depth limit, leaving noise as the sole limiting factor. Our results further suggest that for target noise rates on the order of $p \sim 10^{-4}$, energy density errors on the order of $\sim 10^{-2}$ are achievable.

VII. DISCUSSION

This work introduces a framework for quantum Gibbs-state preparation using local quantum circuits, extending the qMCMC framework developed in Refs. [22,26]. By integrating operator truncation, Trotterization, and variational circuit compilation, we demonstrate that the protocol can be adapted for implementation on near-term quantum hardware. Our results show that the steady state of the truncated Lindbladian closely approximates the true Gibbs state, both in trace distance and in expectation values of local observables, with errors that can be systematically reduced by increasing the truncation radius. In addition, the randomized Trotterization strategy ensures low circuit depth while preserving theoretical convergence guarantees.

A key strength of our approach lies in its flexibility. The truncation scheme and circuit ansatz can be adapted to arbitrary lattice geometries and hardware topologies, making the method broadly applicable beyond the one-dimensional systems studied numerically in this work. Moreover, the use of variationally optimized circuit gadgets enables efficient compilation even under restricted gate sets and limited qubit connectivity. This makes the

protocol more practical for experimental implementation compared to earlier proposals [22,26], especially in settings where only local operators and a small number of ancilla qubits are available.

Several promising directions for future investigation emerge from this study. One natural question concerns the scaling of the required truncation radius at low temperatures, particularly near phase transitions. While our current focus is on the high-temperature regime, the protocol may be extensible to weakly interacting fermionic systems at arbitrary temperatures, in light of recent theoretical advances [24,25]. Another important consideration is noise: although our numerical results suggest robustness to moderate levels of depolarizing noise, further improvements may be possible by incorporating noise models directly into the gadget optimization process.

Additionally, our investigation of different filter functions reveals that algorithmic performance can be substantially improved through appropriate parameter tuning. A systematic understanding of how performance depends on the choice of filter remains an open question. Recent work on quasiparticle cooling [46] suggests that algorithmic efficiency may depend sensitively on the structure of low-energy excitations. Our results pave the way for studying this dependence in greater depth and may ultimately offer new insights into the nature of thermalization processes.

Finally, an important open question is whether this approach can provide a viable route to quantum advantage on near-term or early fault-tolerant setting. Specifically, future work must establish whether low-noise, finite-size quantum processors can achieve sample complexity that surpasses both established classical algorithms for simulation of quantum thermal states [5] and leading tensor-network methods of the relevant Lindbladian dynamics, e.g., proposed in Ref. [47].

ACKNOWLEDGMENTS

We thank Alvaro Alhambra, Anirban Chowdhury, Alexander Miessen, Bibek Pokharel, Kunal Sharma, and Derek Wang for useful discussions. R.S. is grateful to the Alexander von Humboldt Foundation for support, via the German Research Chair program. D.H. acknowledges support from the UK EPSRC under Grant No. EP/X030881/1, and from a Leverhulme Trust International Professorship grant (Award No. LIP-2020-014, for a Leverhulme-Peierls Fellowship at Oxford).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms

of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: LIST OF NOTATIONS

In this Appendix, we collect the main symbols, operators, and definitions used throughout the paper. We use lowercase Latin and Greek letters for numbers, numerical functions, labels, and indices (e.g., n or β), uppercase Greek letters for sets (e.g., Λ), uppercase Latin letters for operators (e.g., H), except for density matrices, which are traditionally denoted by ρ . Calligraphic uppercase letters represent superoperators (e.g., $\mathcal{L}$), while sans-serif letters denote linear spaces and algebras (e.g., $\mathbf{H}$).

1. General notation

- (a) $n \in \mathbb{N}$ —Number of qubits (system size). This number does *not* include ancilla qubits.
- (b) $\Lambda = \{a_1, \dots, a_n\}$ —A finite D -dimensional square lattice with n sites equipped with a graph structure.
- (c) $\ell(a, b)$ —The graph distance defined by

$$\ell(a, b) := \min_{\Pi = \{a \rightarrow c_1, \dots, c_k \rightarrow b\}} |\Pi|,$$

where the minimum is taken over all paths Π between $a \in \Lambda$ and $b \in \Lambda$.

- (d) $B_a(r) = \{b \mid \ell(a, b) \leq r\}$ —The subset of lattice sites located within a ball of radius $r \geq 0$ centered at site a .
- (e) $\mathbf{H}_{\{a\}} \cong \mathbb{C}^2$ —Local Hilbert space associated with qubit at site $a \in \Lambda$.
- (f) $\mathbf{H}_X = \bigotimes_{a \in X} \mathbf{H}_{\{a\}} \cong (\mathbb{C}^2)^{\otimes |X|}$ —Hilbert space of the qubits subset $X \subseteq \Lambda$.
- (g) $\mathbf{M}_X = \text{End}(\mathbf{H}_X)$ —Algebra of linear operators acting on $\mathbf{H}_X$ for $X \subseteq \Lambda$.
- (h) $\mathbf{S}_X = \{O \in \mathbf{M}_X \mid O = O^\dagger\}$ —The set of self-adjoint (Hermitian) operators.
- (i) $\mathbf{D}_X = \{\rho \in \mathbf{S}_X \mid \rho \geq 0, \text{tr} \rho = 1\}$ —The set of density matrices representing mixed quantum states of the qubits.
- (j) $H \in \mathbf{S}_\Lambda$ —Local Hamiltonian on the lattice Λ ; defined in Eq. (2).
- (k) $\lambda_i \in \mathbb{R}$ —Eigenvalues (energies) of the Hamiltonian H , indexed by $i \in \{1, \dots, 2^n\}$.
- (l) $P_i \in \mathbf{S}_\Lambda$ —Projector onto the energy eigenspace of eigenvalue λ_i , satisfying $P_i H P_i = \lambda_i P_i$.
- (m) $\beta \in \mathbb{R}_{\geq 0}$ —Inverse temperature.
- (n) $\rho_\beta \in \mathbf{D}_\Lambda$ —Gibbs (thermal) state associated with H and inverse temperature β , see Eq. (1).
- (o) $\mathbf{F}_\Lambda^{(k,l)} \subset \mathbf{S}_\Lambda$ —Family of (k, l) -local Hamiltonians on lattice Λ being a sum of terms that has no support on more than k sites and each site appears on at most l terms; see Definition B1.

2. Norms and distances

- (a) $\|\cdot\|_1$ —Schatten 1-norm (trace norm) on operators. For $O \in \mathbf{M}_\Lambda$, it is defined by

$$\|O\|_1 := \text{Tr} \sqrt{O^\dagger O}. \quad (\text{A1})$$

- (b) $\|\cdot\|_\infty$ — ∞ -norm (operator or spectral norm) on operators. For $O \in \mathbf{M}_\Lambda$, it is defined by

$$\|O\|_\infty := \sup_{\|v\|=1} \|Ov\|. \quad (\text{A2})$$

Here $v \in \mathbf{H}_\Lambda$ and $\|v\| = \sqrt{\langle v|v \rangle}$.

- (c) $\|\cdot\|_\diamond$ —Diamond norm on superoperators. For $\mathcal{E} : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ it is defined as

$$\|\mathcal{E}\|_\diamond := \max_{X: \|X\|_1 \leq 1} \|(\mathcal{E} \otimes \mathcal{I})X\|_1,$$

where $\mathcal{I} : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ is the identity operator, and the maximum is taken over $X \in \mathbf{M}_\Lambda^{\otimes 2}$.

- (d) $\|X\|$ —oscillator norm

$$\|X\| := \sum_{a \in \Lambda} \|\delta_a(X)\|_\infty,$$

where $\delta_a(X) := X - \frac{1}{2}I_a \otimes \text{tr}_a(X)$.

- (e) $\|\mathcal{L}\|_{p \rightarrow p}$ —Induced $p - p$ norm on a superoperator $\mathcal{L}$,

$$\|\mathcal{L}\|_{p \rightarrow p} = \sup_O \frac{\|\mathcal{L}[O]\|_p}{\|O\|_p}.$$

3. Dissipative dynamics

- (a) $t \in \mathbb{R}_{\geq 0}$ —Continuous time.
- (b) $\mathcal{L}^\beta : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ —Lindbladian superoperator governing the dissipative dynamics that prepares the Gibbs state ρ_β ; see Eq. (3).
- (c) $A^{a,\alpha} \in \mathbf{M}_\Lambda$ —Single-qubit operator acting on site a ; chosen to be Pauli operators; see Eq. (5).
- (d) $\alpha \in \mathbb{N}$ or $\mathbb{R}$ —Discrete index labeling jump operators defined within the neighborhood of qubit $a \in \Lambda$.
- (e) $f(t), q(v) : \mathbb{R} \rightarrow \mathbb{C}$ —Filter and envelope functions, respectively; see Eq. (6).
- (f) $H_{a,r} \in \mathbf{S}_\Lambda$ —Truncated Hamiltonian for the neighborhood of qubit a with radius r ; see Eq. (18).
- (g) $\mathcal{L}^{\beta,r} : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ —Truncated Lindbladian by the radius r ; see Eq. (19).
- (h) $\rho_{\beta,r}$ —steady state of truncated Lindbladian $\mathcal{L}^{\beta,r}$, introduced in Theorem 1.
- (i) $t_{\text{mix}} \in \mathbb{R}_+$ —Mixing time, i.e., minimum time after which the distance between any two states contracts by at least a certain fixed fraction; see Definition 1.

4. Trotterization and compilation

- (a) $\tau \in \mathbb{R}_+$ —Trotter step size (time discretization). Introduced in Sec. III B.
- (b) $M \in \mathbb{N}$ —Number of Trotter steps, such that $t = M\tau$. Introduced in Sec. III B.
- (c) $\mathcal{C}_{r,\tau}^{a,\alpha} : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ —Quantum channel that approximates $\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})$ using one ancilla qubit and a unitary operation. Introduced in Eq. (34).
- (d) $\mathcal{E}_{t,\tau} : \mathbf{M}_\Lambda \rightarrow \mathbf{M}_\Lambda$ —Randomized Trotterization channel for evolution time t compiled as a sequence of $M = t/\tau$ Trotterization gadgets with Trotter step τ . Introduced in Eq. (30).
- (e) $m \in \mathbb{N}$ —Number of repeating depth-3 modules in the gadget compilation circuit. Introduced in Fig. 3.

5. List of functions

- (a) $f(t), q(v) : \mathbb{R} \rightarrow \mathbb{C}$ —Filter and envelope functions, respectively; introduced in Eq. (6).
- (b) $g_1(t), g_2(t) : \mathbb{R} \rightarrow \mathbb{C}$ —Kernel functions in time domain, introduced in Eq. (13).
- (c) $g(v_1, v_2) : \mathbb{R} \rightarrow \mathbb{C}$ —Combined kernel function for the coherent term in frequency domain, introduced in Eq. (C4).
- (d) $\Delta(r_0)$ function to bound the quasilocality of the Lindbladian, introduced in Eq. (D12).
- (e) $\eta(\beta)$ function to bound the distance of the Lindbladian $\mathcal{L}^\beta$ to the Lindbladian $\mathcal{L}^{\beta=0}$, introduced in Eq. (D13). The analogous function $\eta'(\beta)$ for the truncated Lindbladian $\mathcal{L}^{\beta,r}$ is introduced in Eq. (D60).
- (f) $k(r_0, \beta)$, function to bound the decay of the oscillator norm for the Lindbladian $\mathcal{L}^\beta$, introduced in Eq. (D14). The modified version to characterize the oscillator norm decay for the truncated Lindbladian $\mathcal{L}^{\beta,r}$, $k'(r_0, \beta)$, is introduced in Eq. (D60).
- (g) $u(r_0)$, function appearing in definition of $k(r_0, \beta)$, defined in Eq. (D15). The analogous function $u'(r_0)$ for the truncated Lindbladian $\mathcal{L}^{\beta,r}$ is introduced in Eq. (D60).
- (h) $\gamma(x) = \sqrt{x}/(1 + \sqrt{x})$, appearing in bounds for fixed point and local observable errors, first appearance above Eq. (D31).
- (i) $\xi_1(r, \beta J)$ and $\xi_2(r, \beta J)$ functions appearing in bounds of $\Delta(r_0)$, first defined in Eqs. (D33) and (D43).
- (j) $m(|X|) : \mathbb{R} \rightarrow \mathbb{R}$: quantifies how the truncation-error bound for an operator's expectation value on $X \subset \Lambda$ grows with the region size $|X|$ (see Theorem 3).
- (k) $b(r) : \mathbb{R} \rightarrow \mathbb{R}$: characterizes the decay of that error bound as the truncation radius r increases (see Theorem 3).

APPENDIX B: LIEB-ROBINSON BOUNDS

The error bounds for truncated Lindbladians arise from the locality structure of the underlying lattice, which enables the use of Lieb-Robinson bounds. This Appendix provides a very brief overview of the Lieb-Robinson bounds; for a more detailed introduction, see Ref. [31].

Definition B1. We define $\mathbf{F}_\Lambda^{(k,l)} \subset \mathbf{S}_\Lambda$ to be the family of (k, l) -local Hamiltonians on lattice Λ such that any $H \in \mathbf{F}_\Lambda^{(k,l)}$ has the form

$$H = \sum_{X \subseteq \Lambda} h_X,$$

where each term $h_X \in \mathbf{S}_X$ acts on $X \subseteq \Lambda$ with $|X| \leq k$, and such that each site $a \in \Lambda$ appears in the support of at most l terms in the sum. That is, for all $a \in \Lambda$, the number of subsets X with $a \in X$ and $h_X \neq 0$ is at most l .

For any Hamiltonian $H \in \mathbf{F}_\Lambda^{(k,l)}$, we define

$$h := \max_X \|h_X\|_\infty$$

as the uniform upper bound on the norm of the local terms in the Hamiltonian, where $\|\cdot\|_\infty$ is spectral norm, see Eq. (A2).

Assume that the parameters h, k , and l are independent of the total number of sites n . Under these assumptions, the Lieb-Robinson velocity is defined by

$$J := \max_{a \in \Lambda} \sum_{X \ni a} |X| \|h_X\|_\infty = O(hkl). \quad (\text{B1})$$

Given a radius r , the growth of an operator $A^a \in \mathbf{M}_a$ supported on site $a \in \Lambda$ is bounded by [48–52]

$$\|e^{-iHt} A^a e^{iHt} - e^{-iH_{a,r}t} A^a e^{iH_{a,r}t}\|_\infty \leq \|A^a\|_\infty \frac{(2J|t|)^r}{r!}, \quad (\text{B2})$$

where $H_{a,r}$ is a reduction of the Hamiltonian to the ball $B_a(r)$ of radius r centered at site a , as defined in Eq. (18).

APPENDIX C: DISSIPATIVE DYNAMICS STRUCTURE

This Appendix provides explicit time-domain expressions for the jump operators and the coherent term of the Lindbladian in Eq. (4), which are necessary to provide the proofs in Appendixes D and E.

1. The Lindbladian in time domain

Let us focus our attention on the Gaussian envelope function

$$q(v) = \exp\left(-\frac{(\beta v)^2}{8}\right). \quad (C1)$$

After taking the Fourier transform, the corresponding filter function from Eq. (6) becomes

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dv \exp\left(-\frac{(\beta v)^2}{8}\right) \exp\left(-\frac{\beta v}{4}\right) e^{ivt} \\ &= \sqrt{\frac{2}{\pi\beta^2}} \exp\left(\frac{(\beta - 4it)^2}{8\beta^2}\right). \end{aligned} \quad (C2)$$

The jump operators then have the form

$$\begin{aligned} L_{a,\alpha} &= \sum_{v \in \Omega(H)} q(v) e^{-\frac{\beta v}{4}} A_v^{a,\alpha} \\ &= \sum_{v \in \Omega(H)} \exp\left(-\frac{(\beta v + 1)^2 - 1}{8}\right) A_v^{a,\alpha}, \end{aligned} \quad (C3)$$

where $A_v^{a,\alpha}$ are spectral components of the operator $A^{a,\alpha}$ as defined in Eq. (8), and the coherent term in Eq. (7) is

$$\begin{aligned} G_{a,\alpha}^\beta &= -\frac{i}{2} \sum_{v \in \Omega(H)} \tanh\left(-\frac{\beta v}{4}\right) (L_{a,\alpha}^{\beta\dagger} L_{a,\alpha}^\beta)_v \\ &= \sum_{v_1, v_2 \in \Omega(H)} g(v_1, v_2) A_{v_2}^{a,\alpha\dagger} A_{v_1}^{a,\alpha}, \end{aligned} \quad (C4)$$

where the prefactor function is given by

$$\begin{aligned} g(v_1, v_2) &:= -\frac{i}{2} \tanh\left(-\frac{\beta(v_1 - v_2)}{4}\right) \\ &\quad \times \exp\left(-\frac{(\beta v_1 + 1)^2 - 1}{8}\right) \\ &\quad \times \exp\left(-\frac{(\beta v_2 + 1)^2 - 1}{8}\right). \end{aligned} \quad (C5)$$

Introducing the shorthand $v_\pm = v_1 \pm v_2$, this simplifies to

$$\begin{aligned} g(v_1, v_2) &= -\frac{i}{2} \tanh\left(-\frac{\beta v_-}{4}\right) \exp\left(-\frac{\beta^2 v_+^2 + 4\beta v_+}{16}\right) \\ &\quad \times \exp\left(-\frac{\beta^2 v_-^2}{16}\right). \end{aligned} \quad (C6)$$

The splitting into contributions in v_+ and v_- allows for conversion into a time representation. Using the results in

Appendix A of Ref. [22], $G_{a,\alpha}^\beta$ can be written as

$$\begin{aligned} G_{a,\alpha}^\beta &= \int_{-\infty}^{\infty} g_1(t) e^{-iHt} \\ &\quad \times \left(\int_{-\infty}^{\infty} g_2(t') e^{iHt'} A^{a,\alpha\dagger} e^{-2iHt'} A^{a,\alpha} e^{iHt'} dt' \right) e^{iHt} dt, \end{aligned} \quad (C7)$$

with the kernel functions $g_1(t)$ and $g_2(t)$ defined as follows Appendix A of Ref. [22]. The first function is

$$\begin{aligned} g_1(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dv_- e^{iv_- t} \frac{1}{2\pi} \frac{1}{2i} \tanh\left(-\frac{\beta v_-}{4}\right) e^{-\frac{\beta^2 v_-^2}{16}} \\ &= \left[\frac{-1}{\pi\beta \cosh\left(\frac{2\pi t}{\beta}\right)} \right] *_t \left[\frac{\sqrt{2}}{\beta} e^{\frac{1}{4} - \frac{4t^2}{\beta^2}} \sin\left(\frac{2t}{\beta}\right) \right], \end{aligned} \quad (C8)$$

where $f(t) *_t g(t) = \int_{-\infty}^{\infty} f(s) g(t-s) ds$ denotes the convolution. The second function is

$$\begin{aligned} g_2(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dv_+ e^{iv_+ t} \exp\left(-\frac{\beta^2 v_+^2 + 4\beta v_+}{16}\right) \\ &= \frac{2\sqrt{2}}{\beta} \exp\left(\frac{(\beta - 4it)^2}{4\beta^2}\right). \end{aligned} \quad (C9)$$

For future proofs, it would be useful to provide some upper bounds for these functions. Using Young's inequality and taking into account that $|\sin(2t/\beta)| \leq 1$, we get

$$\begin{aligned} \int_{-\infty}^{\infty} dt |g_1(t)| &\leq \int_{-\infty}^{\infty} dt \frac{1}{\pi\beta \cosh\left(\frac{2\pi t}{\beta}\right)} \times \int_{-\infty}^{\infty} ds \frac{\sqrt{2}}{\beta} e^{\frac{1}{4} - \frac{4s^2}{\beta^2}} \\ &= \frac{1}{2\pi} \times \frac{\sqrt{2\pi} e^{1/4}}{2} = \frac{e^{1/4}}{2\sqrt{2\pi}}. \end{aligned} \quad (C10)$$

Also, we can show that

$$\int_{-\infty}^{\infty} dt |g_2(t)| = \int_{-\infty}^{\infty} dt \frac{2\sqrt{2}}{\beta} e^{\frac{1}{4}} e^{-\frac{4t^2}{\beta^2}} = e^{\frac{1}{4}} \sqrt{2\pi}. \quad (C11)$$

Next, we prove the following results for the tails of these functions.

Lemma C1. For any $t_0 \geq 0$, the functions $g_1(t)$ and $g_2(t)$ satisfy

$$\int_{t_0}^{\infty} dt |g_1(t)| \leq \frac{\sqrt{2}}{2\pi^{3/2}} e^{\frac{1}{4}} \left(e^{\frac{\pi^2}{4}} e^{-\frac{2\pi t_0}{\beta}} + \frac{1}{2} \operatorname{erfc}\left(\frac{2t_0}{\beta}\right) \right), \quad (C12)$$

$$\int_{t_0}^{\infty} dt |g_2(t)| \leq \sqrt{\frac{\pi}{2}} e^{\frac{1}{4}} \operatorname{erfc}\left(\frac{2t_0}{\beta}\right). \quad (C13)$$

Proof. Consider first $g_1(t)$. We have for $t_0 > 0$

$$\begin{aligned} & \int_{t_0}^{\infty} dt |g_1(t)| \\ & \leq \int_{t_0}^{\infty} dt \int_{-\infty}^{\infty} ds \frac{1}{\pi \beta \cosh\left(\frac{2\pi s}{\beta}\right)} \frac{\sqrt{2}}{\beta} e^{\frac{1}{4} - \frac{4(t-s)^2}{\beta^2}} \\ & \leq \frac{2\sqrt{2}}{\pi \beta^2} e^{\frac{1}{4}} \int_{t_0}^{\infty} dt \int_0^{\infty} ds e^{-\frac{2\pi s}{\beta}} \left(e^{-\frac{4(t-s)^2}{\beta^2}} + e^{-\frac{4(t+s)^2}{\beta^2}} \right), \end{aligned} \quad (C14)$$

where we used $1/\cosh(x) \leq 2e^{-|x|}$ in the last line and split the integral over s into two halves $\int_0^{\infty}$ and $\int_{-\infty}^0$.

Let us consider the first integral in the form

$$I_1 = \int_{t_0}^{\infty} dt \int_0^{\infty} ds e^{-\frac{2\pi s}{\beta}} e^{-\frac{4(t-s)^2}{\beta^2}},$$

one finds by completing the square in s that

$$\int_0^{\infty} ds e^{-\frac{2\pi s}{\beta}} e^{-\frac{4(t-s)^2}{\beta^2}} = \frac{\beta\sqrt{\pi}}{4} e^{\frac{\pi^2}{4}} e^{-\frac{2\pi t}{\beta}} \operatorname{erfc}\left(\frac{\pi}{2} - \frac{2t}{\beta}\right).$$

Hence, we get

$$I_1 = \frac{\beta\sqrt{\pi}}{4} e^{\frac{\pi^2}{4}} \int_{t_0}^{\infty} dt e^{-\frac{2\pi t}{\beta}} \operatorname{erfc}\left(\frac{\pi}{2} - \frac{2t}{\beta}\right).$$

Since $\operatorname{erfc}(x) \leq 2$ for all x , we get the bound

$$I_1 \leq \frac{\beta\sqrt{\pi}}{2} e^{\frac{\pi^2}{4}} \int_{t_0}^{\infty} dt e^{-\frac{2\pi t}{\beta}} = \frac{\beta^2}{4\sqrt{\pi}} e^{\frac{\pi^2}{4}} e^{-\frac{2\pi t_0}{\beta}}.$$

The other term can be bounded as

$$\begin{aligned} I_2 &= \int_{t_0}^{\infty} dt \int_0^{\infty} ds e^{-\frac{2\pi s}{\beta}} e^{-\frac{4(t+s)^2}{\beta^2}} \\ &\leq \int_{t_0}^{\infty} dt \int_0^{\infty} ds e^{-\frac{2\pi s}{\beta}} e^{-\frac{4t^2}{\beta^2}} = \frac{\beta^2}{8\sqrt{\pi}} \operatorname{erfc}\left(-\frac{2t_0}{\beta}\right). \end{aligned} \quad (C15)$$

This leads us to the expression in Eq. (C12). The expression in Eq. (C13) follows from bounding an Gaussian integral. ■

2. Zero-temperature Lindblad operator

In the infinite-temperature limit, i.e., as $\beta \rightarrow 0$, the function $g(v_1, v_2)$ in Eq. (C5) vanishes, implying that $G_{a,\alpha}^{\beta=0} = 0$. Moreover, from Eq. (C2), it follows that

$$\lim_{\beta \rightarrow 0} f(t) = \delta(t), \quad (C16)$$

where $\delta(t)$ denotes the Dirac delta function.

Given this limiting behavior, and choosing the operator set as in Eq. (16), the Lindblad generator in the zero-temperature case simplifies to

$$\begin{aligned} \mathcal{L}^{\beta=0}(\rho) &= \sum_{a \in \Lambda} \sum_{\alpha} \left(A^{a,\alpha} \rho A^{a,\alpha\dagger} - \frac{1}{2} \{A^{a,\alpha\dagger} A^{a,\alpha}, \rho\} \right) \\ &= \sum_{a \in \Lambda} \left(\frac{1}{2} \operatorname{tr}_a(\rho) \otimes I_a - \rho \right), \end{aligned} \quad (C17)$$

where tr_a is the partial trace over the single-qubit space $\mathcal{H}_a$ and $I_a \in \mathcal{H}_a \rightarrow \mathcal{H}_a$ is the identity operator on qubit a . This expression corresponds to the generator of the depolarizing noise channel, which drives the system toward the maximally mixed state by uniformly averaging over local Pauli operators [23].

APPENDIX D: BOUNDS ON THE TRACE DISTANCE

This Appendix gives the proof for Theorems 1 and 2.

Theorem D1 (Truncation error, Theorem 1 restated). Let H be a local Hamiltonian on a finite lattice Λ with n qubits, ρ_{β} the Gibbs state, and $\mathcal{L}^{\beta,r}$ the truncated Lindbladian in Eq. (19). Then, there exists a constant $\beta^* = 1/500^D J$ such that for any $\beta < \beta^*$, there exists a unique $\rho_{\beta,r} \in \mathcal{M}_{\Lambda}$ satisfying $\mathcal{L}^{\beta,r}(\rho_{\beta,r}) = 0$ and $J > 0$, such that

$$\|\rho_{\beta,r} - \rho_{\beta}\|_1 = O((\beta J)^{r/2} n \log n), \quad (D1)$$

where $\|\cdot\|_1$ denotes the trace norm.

Theorem D2 (Theorem 2 restated). Let H be a (k, l) -local Hamiltonian with Lieb-Robinson velocity $J > 0$ acting on n qubits arranged on a D -dimensional lattice, and let $\beta^* = 1/2020^D J$. Then, for any $\beta < \beta^*$, state ρ and $\epsilon > 0$, the generator $\mathcal{L}^{\beta}$ in Eq. (3) with the choice $q(v) = e^{-(\beta v)^2/8}$ and its fixed point ρ_{β} satisfy

$$\|e^{t\mathcal{L}}(\rho) - \rho_{\beta}\|_1 \leq \epsilon \|\rho - \rho_{\beta}\|_1 \quad \text{for all } t = \Omega(\log(n/\epsilon)). \quad (D2)$$

From this statement, it follows that for $\beta < \beta^*$, the mixing time of the Lindbladian $\mathcal{L}^{\beta}$ scales as $t_{\text{mix}} = \mathcal{O}(\log n)$. This statement is also true for the truncated Lindbladian $\mathcal{L}^{\beta,r}$ and its fixed point $\rho_{\beta,r}$, respectively.

Note that the results for the nontruncated Lindbladian generators presented in Ref. [26] are analogous to those for the generators introduced in Ref. [22]. In particular, for the latter generators, a logarithmic scaling of the mixing time with system size at inverse temperatures $\beta < \beta^*$ was established in Ref. [23]. By leveraging the structure of their proof, we extend this result to the Lindblad generator defined in Eq. (3) of Ref. [26] for the specific choice

$q^{a,\alpha}(\nu) = \exp(-(\beta\nu)^2/8)$. Crucially, the argument relies only on the quasilocality of the jump operators and hence applies equally to any other choice of $q_{a,\alpha}(\nu)$ for which the corresponding envelope function $f(t)$ in Eq. (6) decays sufficiently rapidly in t .

Furthermore, quasilocality of the jump operators is preserved under truncation. Consequently, the mixing time remains logarithmic in the system size even when one considers the truncated Lindbladian $\mathcal{L}^{\beta,r}$, albeit now with respect to the perturbed fixed point $\rho_{\beta,r}$. It therefore suffices to prove Theorem D2 for $\mathcal{L}^{\beta,r}$, and hence to establish Theorem D1.

The rest of the Appendix section is structured as follows. In Appendix D1, we state the necessary propositions and lemmas to set up the proofs. These preliminaries allow us to construct the proof of Theorem D2 in Appendix D2, which in turn relies on Lemma D2, proved in Appendix D3. Finally, we complete the present Appendix section by proving Theorem D1 in Appendix D4.

1. Preliminaries

This subsection collects the necessary lemmas to provide the proofs for Theorems D1 and D2. The proof of Theorem D1 is built on the following lemma that connects the fixed points of a pair of close Lindbladians:

Lemma D1 (Lemma II.1 of Ref. [21]). Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two generators of Lindbladian evolution with unique fixed points $\rho_{\text{fix}}(\mathcal{L}_1)$ and $\rho_{\text{fix}}(\mathcal{L}_2)$, respectively. Then, the trace norm difference between these fixed points satisfies

$$\|\rho_{\text{fix}}(\mathcal{L}_1) - \rho_{\text{fix}}(\mathcal{L}_2)\|_1 \leq 4 \|\mathcal{L}_1 - \mathcal{L}_2\|_{1 \rightarrow 1} \cdot t_{\text{mix}}(\mathcal{L}_1), \quad (\text{D3})$$

where $t_{\text{mix}}(\mathcal{L}_1)$ denotes the mixing time of the generator $\mathcal{L}_1$.

This lemma establishes a bound relating the distance between the fixed points of two Lindbladian generators, the superoperator difference norm $\|\mathcal{L}_1 - \mathcal{L}_2\|_{1 \rightarrow 1}$, and the mixing time t_{mix} . The proof can be found in Ref. [21].

In order to prove Theorem D2 and a fast mixing time, we make use of the following lemma:

Lemma D2. Consider the oscillator norm

$$\|X\| := \sum_{a \in \Lambda} \|\delta_a(X)\|_\infty,$$

where $\delta_a(X) := X - \frac{1}{2}I_a \otimes \text{tr}_a(X)$. Under the same assumptions as for Theorem D2, for $\beta < \beta^*$ there exists a $\kappa \equiv \kappa(\beta)$ with $0 \leq \kappa < 1$ such that, for any initial

observable X ,

$$\|e^{t\mathcal{L}^{\beta\dagger}}(X)\| \leq e^{-(1-\kappa)t} \|X\|. \quad (\text{D4})$$

Additionally, for $\beta < \beta^*$, there exists a $\kappa'(\beta)$ with $0 \leq \kappa' < 1$ such that, for any initial observable X ,

$$\|e^{t\mathcal{L}^{\beta,r\dagger}}(X)\| \leq e^{-(1-\kappa')t} \|X\|. \quad (\text{D5})$$

We provide the proof of this lemma, which relies on the quasilocality of the Lindbladian, in Appendix D3. It follows the steps of Ref. [23]. Furthermore, the proof gives as a by-product a bound on $\|\mathcal{L}^\beta - \mathcal{L}^{\beta,r}\|_{1 \rightarrow 1}$ which we use in combination with Lemma D1 and Theorem D2 to obtain the proof of Theorem D1.

2. Proof of Theorem D2

The proof follows the same steps as the proof for the logarithmic mixing time in Ref. [23] for the specific instance of quantum Gibbs sampler in Ref. [22].

The uniqueness of the fixed point follows because the operators $A^{a,\alpha}$ in the Lindbladian in Eq. (19) form a complete set of generators and thus the Lindbladian is irreducible [26,31]. Consider

$$\begin{aligned} \|e^{t\mathcal{L}^\beta}(\rho) - \rho_\beta\|_1 &= \sup_{\|X\|_\infty \leq 1} \text{tr}(X(e^{t\mathcal{L}^\beta}(\rho) - \rho_\beta)) \\ &= \sup_{\|X\|_\infty \leq 1} \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)(\rho - \rho_\beta)). \end{aligned} \quad (\text{D6})$$

Here we used the definition of the adjoint operator $\text{tr}(Xe^{t\mathcal{L}^\beta}(\rho)) = \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)\rho)$, and the invariance of the steady state under dissipative evolution $e^{t\mathcal{L}^\beta}(\rho_\beta) = \rho_\beta$. Observe now, that $\text{tr}(\rho - \rho_\beta) = 0$ and thus

$$\begin{aligned} &\text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)(\rho - \rho_\beta)) \\ &= \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)(\rho - \rho_\beta)) - 2^{-n} \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)) \text{tr}(\rho - \rho_\beta) \\ &= \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X)(\rho - \rho_\beta)) - \text{tr}(2^{-n} \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X))I)(\rho - \rho_\beta) \\ &= \text{tr}\left(\left(e^{t\mathcal{L}^{\beta\dagger}}(X) - 2^{-n} \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X))I\right)(\rho - \rho_\beta)\right). \end{aligned} \quad (\text{D7})$$

We can use this expression in Eq. (D6) and obtain together with Lemma D2 that

$$\begin{aligned} &\|e^{t\mathcal{L}^\beta}(\rho) - \rho_\beta\|_1 \\ &= \sup_{\|X\|_\infty \leq 1} \|e^{t\mathcal{L}^{\beta\dagger}}(X) - 2^{-n} \text{tr}(e^{t\mathcal{L}^{\beta\dagger}}(X))I\|_\infty \|\rho - \rho_\beta\|_1 \\ &\leq \sup_{\|X\|_\infty \leq 1} \|e^{t\mathcal{L}^{\beta\dagger}}(X)\| \|\rho - \rho_\beta\|_1 \\ &\leq e^{-(1-\kappa)t} \sup_{\|X\|_\infty \leq 1} \|X\| \|\rho - \rho_\beta\|_1 \\ &\leq 2n \|\rho - \rho_\beta\|_1 e^{-(1-\kappa)t}. \end{aligned} \quad (\text{D8})$$

In the first line, we have used the fact that the trace over many sites can be written sequentially as a composition of trace over single sites. In the last line, we used

$$\|\delta_a(X)\|_\infty \leq \|X\|_\infty + \frac{1}{2}\|I_a \otimes \text{tr}_a(X)\|_\infty \leq 2\|X\|_\infty, \quad (\text{D9})$$

since the operator norm is not increased by partial trace, and thus $\|X\| \leq 2n$. From this it follows that

$$\|e^{t\mathcal{L}}(\rho) - \rho_\beta\|_1 \leq \epsilon\|\rho - \rho_\beta\|_1, \quad (\text{D10})$$

with

$$t = \frac{1}{1-\kappa} \log\left(\frac{2n}{\epsilon}\right). \quad (\text{D11})$$

This concludes the proof of Theorem D2.

3. Proof of Lemma D2

In this Appendix subsection, we give the proof of Lemma D2, following the same steps as Ref. [23].

Proof of Lemma D2. We define for any $r_0 \in \mathbb{N}$ and β the quantities

$$\Delta(r_0) = \sum_{r \geq r_0} \|\mathcal{L}^{\beta,r\dagger} - \mathcal{L}^{\beta,(r-1)\dagger}\|_{\infty \rightarrow \infty}, \quad (\text{D12})$$

$$\eta(\beta) = \|\mathcal{L}^{\beta\dagger} - \mathcal{L}^{0\dagger}\|_{\infty \rightarrow \infty}. \quad (\text{D13})$$

Each term appearing in the oscillator norm can be bounded locally [23,53]. These local bounds allow bounding the decay of the oscillator norm, as is shown in Ref. [53]. More specifically, consider the following quantity [Eq. (A35) of Ref. [23]], which arises from geometrical considerations of the D -dimensional lattice and the behavior of the oscillator norm,

$$k(r_0, \beta) = 4(2r_0 + 1)^{2D} \eta(\beta) + u(r_0), \quad (\text{D14})$$

where $u(r_0)$ is defined as

$$\begin{aligned} u(r_0) &= 5(2r_0 + 1)^{2D} \Delta(r_0) + (5 + 2r_0 + 2(2r_0 + 1)^D) \\ &\quad \times \sum_{l \geq r_0} (2l + 1)^{2D-1} \Delta(l) \\ &\quad + 2 \sum_{l' \geq r_0} \sum_{l \geq l'} (2l + 1)^{2D-2} \Delta(l). \end{aligned} \quad (\text{D15})$$

Concretely, if there exists an $r_0 \in \mathbb{N}$ and β such that $k(r_0, \beta) < 1$, then the inequality equation (D4) is satisfied by the choice $\kappa = k(r_0, \beta)$. The remaining strategy for the proof is thus the following: First, we provide bounds on $\Delta(r)$ and $\eta(\beta)$ for the Lindbladian $\mathcal{L}^\beta$. This allows us to bound $u(r_0)$ for $\beta < \beta^*$, and a properly chosen r_0 such that $k(r_0, \beta) < 1$, which will complete the proof. ■

a. Bounding $\Delta(r)$

In the following, the steps in (Ref. [23], Appendix B) are adapted to the Lindbladian $\mathcal{L}^\beta$ in Eq. (3).

As the first step, the Lindbladian equation (3) can be rewritten as a telescopic sum

$$\mathcal{L}_{a,\alpha}^\beta = \mathcal{L}_{a,\alpha}^{\beta,0} + \sum_{r=0}^{\infty} (\mathcal{L}_{a,\alpha}^{\beta,r+1} - \mathcal{L}_{a,\alpha}^{\beta,r}) \equiv \mathcal{L}_{a,\alpha}^{\beta,0} + \sum_{r=0}^{\infty} \mathcal{E}_{a,\alpha}^{\beta,r}, \quad (\text{D16})$$

with $\mathcal{E}_{a,\alpha}^{\beta,r} := \mathcal{L}_{a,\alpha}^{\beta,r+1} - \mathcal{L}_{a,\alpha}^{\beta,r}$ the finite difference operator supported on $B_a(r+1)$ [54].

In this representation, the difference between $\mathcal{L}_{a,\alpha}^\beta$ and $\mathcal{L}_{a,\alpha}^{\beta,r}$ is given by

$$\mathcal{L}_{a,\alpha}^\beta - \mathcal{L}_{a,\alpha}^{\beta,r} = \sum_{r'=r}^{\infty} \mathcal{E}_{a,\alpha}^{\beta,r'}. \quad (\text{D17})$$

It remains to be shown that the terms $\mathcal{E}_{a,\alpha}^{\beta,r}$ decay sufficiently fast with the radius r .

We aim to show that there is a sequence of Lindbladians with jump operators with support on a region of radius r such that, for a fast-decaying function $\zeta(r)$,

$$\|\mathcal{E}_{a,\alpha}^{\beta,r}\|_{\infty \rightarrow \infty} \leq \zeta(r). \quad (\text{D18})$$

Recall the jump operators and coherent term of the Lindbladian $\mathcal{L}_{a,\alpha}^{\beta,r}$:

$$L_{a,\alpha}^{\beta,r} := \int_{-\infty}^{\infty} e^{iH_r t} A^{a,\alpha} e^{-iH_r t} f(t) dt, \quad (\text{D19})$$

and coherent part

$$\begin{aligned} G_{a,\alpha}^{\beta,r} &\equiv \int_{-\infty}^{\infty} g_1(t) e^{-iH_r t} \int_{-\infty}^{\infty} g_2(t') e^{iH_r t'} A^{a,\alpha\dagger} e^{-2iH_r t'} \\ &\quad \times A^{a,\alpha} e^{iH_r t'} dt' e^{iH_r t} dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(t) g_2(t') A_r^{a,\alpha\dagger}(t-t') A_r^{a,\alpha}(t+t') dt' dt. \end{aligned} \quad (\text{D20})$$

(D21)

The explicit expressions for $f(t)$, $g_1(t)$ and $g_2(t)$ with the choice of $q(v)$ in Eq. (11) are given by [cf. Appendix C 1]

$$f(t) = \sqrt{\frac{2}{\pi\beta^2}} e^{\frac{(\beta-4it)^2}{8\beta^2}}, \quad (\text{D22})$$

$$g_1(t) = \left[\frac{-1}{\pi\beta \cosh\left(\frac{2\pi t}{\beta}\right)} \right] *_t \left[\frac{\sqrt{2}}{\beta} e^{\frac{1}{4} - \frac{4t^2}{\beta^2}} \sin\left(\frac{2t}{\beta}\right) \right], \quad (\text{D23})$$

$$g_2(t) = \frac{2\sqrt{2}}{\beta} e^{\frac{(\beta-4it)^2}{4\beta^2}}. \quad (\text{D24})$$

In the following, we consider the dissipative part of the Lindbladian $\mathcal{D}_{a,\alpha}^{\beta,r}$ defined as

$$\forall X : \quad \mathcal{D}_{a,\alpha}^{\beta,r}(X) := L_{a,\alpha}^{\beta,r} X L_{a,\alpha}^{\beta,r\dagger} - \frac{1}{2} \{L_{a,\alpha}^{\beta,r\dagger} L_{a,\alpha}^{\beta,r}, X\}. \quad (\text{D25})$$

To proceed, we note that using the triangle inequality, we have

$$\begin{aligned} & \|L_{a,\alpha}^{\beta,(r+1)} X L_{a,\alpha}^{\beta,(r+1)\dagger} - L_{a,\alpha}^{\beta,r} X L_{a,\alpha}^{\beta,r\dagger}\|_{\infty} \\ &= \|(L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}) X L_{a,\alpha}^{\beta,(r+1)\dagger} - L_{a,\alpha}^{\beta,r} X (L_{a,\alpha}^{\beta,r\dagger} - L_{a,\alpha}^{\beta,(r+1)\dagger})\|_{\infty} \\ &\leq \|(L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}) X L_{a,\alpha}^{\beta,(r+1)\dagger}\|_{\infty} + \|L_{a,\alpha}^{\beta,r} X (L_{a,\alpha}^{\beta,r\dagger} - L_{a,\alpha}^{\beta,(r+1)\dagger})\|_{\infty}. \end{aligned} \quad (\text{D26})$$

We can further bound the terms using Hoelder's inequality,

$$\|(L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}) X L_{a,\alpha}^{\beta,(r+1)\dagger}\|_{\infty} \leq \|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_{\infty} \|X\|_{\infty} \|L_{a,\alpha}^{\beta,(r+1)\dagger}\|_{\infty}. \quad (\text{D27})$$

Using the same steps, the same result is obtained for $\frac{1}{2} \{L_{a,\alpha}^{\beta,r\dagger} L_{a,\alpha}^{\beta,r}, X\}$. Combining Eqs. (D26) and (D27), this gives

$$\|\mathcal{D}_{a,\alpha}^{\beta,(r+1)} - \mathcal{D}_{a,\alpha}^{\beta,r}\|_{\infty \rightarrow \infty} = \sup_X \frac{\|\mathcal{D}_{a,\alpha}^{\beta,(r+1)}(X) - \mathcal{D}_{a,\alpha}^{\beta,r}(X)\|_{\infty}}{\|X\|_{\infty}} \leq 2 \left(\|L_{a,\alpha}^{\beta,r}\|_{\infty} + \|L_{a,\alpha}^{\beta,(r+1)}\|_{\infty} \right) \|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_{\infty}. \quad (\text{D28})$$

Consider the last factor,

$$\|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_{\infty} = \left\| \int_{-\infty}^{\infty} dt f(t) e^{-iH_{r+1}t} A^{a,\alpha} e^{iH_{r+1}t} - e^{-iH_r t} A^{a,\alpha} e^{iH_r t} \right\|_{\infty}. \quad (\text{D29})$$

In order to bound the integral, we split the integration range in a short time interval $t \in [-t_0, t_0]$ and long-time tails $|t| > t_0$. The choice of t_0 will be fixed below such that the resulting bound shows explicit decay with r and it is explicit that it vanishes in the limit $\beta \rightarrow 0$.

For the short time interval $t \in [-t_0, t_0]$, we use the Lieb-Robinson bound (B2). For the long-time tails, we use the trivial upper bound

$$\|e^{-iH_{r+1}t} A^{a,\alpha} e^{iH_{r+1}t} - e^{-iH_r t} A^{a,\alpha} e^{iH_r t}\|_{\infty} \leq \|e^{-iH_{r+1}t} A^{a,\alpha} e^{iH_{r+1}t}\|_{\infty} + \|e^{-iH_r t} A^{a,\alpha} e^{iH_r t}\|_{\infty} \leq 2 \|A^{a,\alpha}\|_{\infty}.$$

The last factor in Eq. (D28) can then be bounded by

$$\begin{aligned} \|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_{\infty} &\leq \|A^{a,\alpha}\|_{\infty} \int_{-t_0}^{t_0} |f(t)| \frac{(2J|t|)^r}{r!} dt + 4 \|A^{a,\alpha}\|_{\infty} \int_{t_0}^{\infty} f(t) dt \\ &\leq \|A^{a,\alpha}\|_{\infty} \left(\frac{(2Je|t_0|)^r}{r^r} \int_{-\infty}^{\infty} |f(t)| dt + 4 \int_{t_0}^{\infty} |f(t)| dt \right) \\ &\leq \|A^{a,\alpha}\|_{\infty} e^{\frac{1}{8}} \left(\frac{(2Je|t_0|)^r}{r^r} + 2e^{\frac{-2t_0^2}{\beta^2}} \frac{\beta}{\sqrt{2\pi t_0}} \right). \end{aligned} \quad (\text{D30})$$

In the second line, we used the estimate $r! \geq r^r / e^{r-1}$ and in the last line the upper bound on the complementary error function $\text{erfc}(x) \leq e^{-x^2} / \sqrt{\pi} x$ to evaluate the integrals.

With the choice $t_0 = r\gamma(\beta J)/2Je$ with $\gamma(x) = \sqrt{x}/(1 + \sqrt{x})$, this gives

$$\|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_\infty \leq e^{\frac{1}{8}} \|A^{a,\alpha}\|_\infty \left(\gamma(\beta J)^r + \frac{4\beta Je}{\sqrt{2\pi}r\gamma(\beta J)} e^{-\frac{r^2\gamma(\beta J)^2}{2(\beta Je)^2}} \right). \quad (\text{D31})$$

Furthermore, we have

$$\|L_{a,\alpha}^{\beta,(r+1)}\|_\infty \leq e^{1/8} \|A^{a,\alpha}\|_\infty. \quad (\text{D32})$$

Thus we have

$$\|\mathcal{D}_{a,\alpha}^{\beta,(r+1)} - \mathcal{D}_{a,\alpha}^{\beta,r}\|_{\infty \rightarrow \infty} \leq 4 \|A^{a,\alpha}\|_\infty^2 e^{\frac{1}{4}} \left(\gamma(\beta J)^r + \frac{4\beta Je}{\sqrt{2\pi}r\gamma(\beta J)} e^{-\frac{r^2\gamma(\beta J)^2}{2(\beta Je)^2}} \right) \equiv \|A^{a,\alpha}\|_\infty^2 \xi_1(r, \beta J). \quad (\text{D33})$$

For the coherent part, similar to our derivation of Eqs. (D26) and (D27), it can be shown that

$$\| -i[G_{a,\alpha}^{\beta,r} - G_{a,\alpha}^{\beta,(r+1)}, \cdot] \|_{\infty \rightarrow \infty} \leq 2 \|G_{a,\alpha}^{\beta,r} - G_{a,\alpha}^{\beta,(r+1)}\|_\infty \quad (\text{D34})$$

and

$$\|G_{a,\alpha}^{\beta,r} - G_{a,\alpha}^{\beta,(r+1)}\|_\infty \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(t)g_2(t')| \left\| A_r^{a,\alpha}(t-t')^\dagger A_r^{a,\alpha}(t+t') - A_{r+1}^{a,\alpha}(t-t')^\dagger A_{r+1}^{a,\alpha}(t+t') \right\|_\infty dt' dt. \quad (\text{D35})$$

We can split the interval into five regions as follows, with a different cutoff t_0 :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots = \int_{-t_0}^{t_0} \int_{-t_0}^{t_0} + \int_{t_0}^{\infty} \int_{-t_0}^{t_0} + \int_{-\infty}^{\infty} \int_{t_0}^{\infty} + \int_{-\infty}^{-t_0} \int_{-t_0}^{t_0} + \int_{-\infty}^{\infty} \int_{-\infty}^{-t_0}. \quad (\text{D36})$$

The integral can be bounded as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\cdots| \leq \int_{-t_0}^{t_0} \int_{-t_0}^{t_0} + \int_{t_0}^{\infty} \int_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \int_{t_0}^{\infty} + \int_{-\infty}^{-t_0} \int_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \int_{-\infty}^{-t_0}. \quad (\text{D37})$$

Consider the first integral. With a similar argument as Eqs. (D26)–(D28), we obtain

$$\begin{aligned} \int_{-t_0}^{t_0} \int_{-t_0}^{t_0} |g_1(t)g_2(t')| \left\| A_r^{a,\alpha}(t-t')^\dagger A_r^{a,\alpha}(t+t') - A_{r+1}^{a,\alpha}(t-t')^\dagger A_{r+1}^{a,\alpha}(t+t') \right\|_\infty dt' dt &\leq e^{1/2} \frac{(4Je|t_0|)^r}{2r^r} \\ &\times \|A^{a,\alpha} A^{a,\alpha}\|_\infty, \end{aligned} \quad (\text{D38})$$

where we used the Lieb-Robinson bound and that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(t)g_2(t')| dt' dt \leq e^{1/2}/2$ (cf. Appendix C 1).

For the other long-time contributions, we simply bound

$$\left\| A_r^{a,\alpha}(t-t')^\dagger A_r^{a,\alpha}(t+t') - A_{r+1}^{a,\alpha}(t-t')^\dagger A_{r+1}^{a,\alpha}(t+t') \right\|_\infty \leq 2 \|A^{a,\alpha} A^{a,\alpha}\|_\infty = 2 \|A^{a,\alpha}\|_\infty^2. \quad (\text{D39})$$

Given the tail bounds [cf. Eqs. (C12) and (C13)]

$$\int_{t_0}^{\infty} |g_1(t)| dt \leq \frac{\sqrt{2}}{2\pi^{3/2}} e^{\frac{1}{4}} \left(e^{\frac{\pi^2}{4}} e^{-\frac{2\pi t_0}{\beta}} + \frac{1}{2} \operatorname{erfc}\left(\frac{2t_0}{\beta}\right) \right), \quad (\text{D40})$$

$$\int_{t_0}^{\infty} |g_2(t)| dt \leq \sqrt{\frac{\pi}{2}} e^{\frac{1}{4}} \operatorname{erfc}\left(\frac{2t_0}{\beta}\right), \quad (\text{D41})$$

the contribution of all the other terms in the integral in Eq. (D35) can be bounded as

$$2 \left(\int_{-\infty}^{\infty} \int_{t_0}^{\infty} + \dots \right) |g_1(t)g_2(t')| dt' dt \|A^{a,\alpha}\|_{\infty}^2 \leq 2 \left(\frac{2e^{\frac{1}{2}}}{\pi} e^{\frac{\pi^2}{4}} e^{-\frac{2\pi t_0}{\beta}} + e^{\frac{1}{2}} \frac{\beta e^{-\frac{2t_0^2}{\beta^2}}}{2t_0\sqrt{\pi}} \right) \|A^{a,\alpha}\|_{\infty}^2. \quad (\text{D42})$$

With the choice $t_0 = \gamma(\beta J)r/(4Je)$ we obtain

$$\| -i[G_{a,\alpha}^{\beta,r} - G_{a,\alpha}^{\beta,(r+1)}, \cdot] \|_{\infty \rightarrow \infty} \leq e^{1/2} \|A^{a,\alpha}\|_{\infty}^2 \left(\gamma(\beta J)^r + \frac{8}{\pi} e^{\frac{\pi^2}{4} - \frac{\pi\gamma(\beta J)}{2\beta Je}r} + \frac{8\beta Je}{\sqrt{\pi}r\gamma(\beta J)} e^{-\frac{(r\gamma(\beta J))^2}{8(\beta Je)^2}} \right) \quad (\text{D43})$$

$$\equiv \|A^{a,\alpha}\|_{\infty}^2 \xi_2(r, \beta J). \quad (\text{D44})$$

With the assumption $\|A^{a,\alpha}\|_{\infty} \leq 1$, we can take $\zeta(r)$ in Eq. (D18) to be

$$\zeta(r) = \xi_1(r, \beta J) + \xi_2(r, \beta J). \quad (\text{D45})$$

$\zeta(r)$ is exponentially decaying in r and vanishes as $\beta \rightarrow 0$. With the definition $v := \beta J/\gamma(\beta J) = \sqrt{\beta J}(1 + \sqrt{\beta J})$ and taking $r \geq 1$, a crude bound for $\zeta(r)$ is given by

$$\zeta(r) \leq 7\gamma(\beta J)^r + 44ve^{-\frac{r^2}{2e^2v^2}} + 50e^{-\frac{\pi r}{2ev}}. \quad (\text{D46})$$

For $\beta J \leq 1/100$, it can be verified numerically (cf. Fig. 12) that the term $7\gamma(\beta J)^r$ in Eq. (D46) is dominant [55], therefore the other terms contribute at most $7\gamma(\beta J)^r$. Thus for $\beta J \leq \frac{1}{100}$, we obtain

$$\|\mathcal{E}_{a,\alpha}^{\beta,r}\|_{\infty \rightarrow \infty} \leq \zeta(r) \leq 14\gamma(\beta J)^r, \quad (\text{D47})$$

and therefore

$$\Delta(r_0) = \sum_{r=r_0}^{\infty} \|\mathcal{E}_{a,\alpha}^{\beta,r}\|_{\infty \rightarrow \infty} = \sum_{r \geq r_0} \zeta(r) \leq \frac{14\gamma(\beta J)^{r_0}}{1 - \gamma(\beta J)}. \quad (\text{D48})$$

The proper r_0 will be chosen below in Appendix D 3 c.

b. Bounding $\eta(\beta)$

As a next step, we have to bound $\eta(\beta) = \|\mathcal{L}^{\beta\dagger} - \mathcal{L}^{0\dagger}\|_{\infty \rightarrow \infty}$.

Consider $\mathcal{L}^0$. It can be shown (cf. Sec. C 2) that $\mathcal{L}^{0\dagger}(X) := \sum_{a \in \Lambda} (\frac{1}{2}I_a \otimes \text{tr}_a(X) - X)$ for any X . This is the generator of the fully depolarizing channel.

As before, we consider coherent and dissipative parts separately. For the jump operators, we have

$$\|L_{a,\alpha}^{\beta} - L_{a,\alpha}^0\|_{\infty} \leq \int_{-\infty}^{\infty} \|A^{a,\alpha}(t) - A^{a,\alpha}\|_{\infty} f(t) dt \quad (\text{D49})$$

$$\leq \| [H, A] \|_{\infty} \int_{-\infty}^{\infty} |tf(t)| dt \quad (\text{D50})$$

$$\leq \frac{2}{\sqrt{2\pi}} e^{\frac{1}{8}} \beta J \|A^{a,\alpha}\|_{\infty}. \quad (\text{D51})$$

Here we have used $\|A(t) - A\|_{\infty} \leq |t| \| [H, A] \|_{\infty} \leq 2J|t| \|A\|_{\infty}$ [23]. With $\|A^{a,\alpha}\|_{\infty} \leq 1$, we obtain, with a similar derivation as Eq. (D28),

$$\|\mathcal{D}_{a,\alpha}^{\beta\dagger} - \mathcal{D}_{a,\alpha}^{0\dagger}\|_{\infty \rightarrow \infty} \leq \frac{8\beta J}{\sqrt{2\pi}} e^{\frac{1}{8}}. \quad (\text{D52})$$

For the coherent part, we introduce $c_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(t)g_2(t')dt'dt$. With our choice of the jump operators $A^{a,\alpha}$, we have $(A^{a,\alpha})^2 = \mathbb{I}$, and

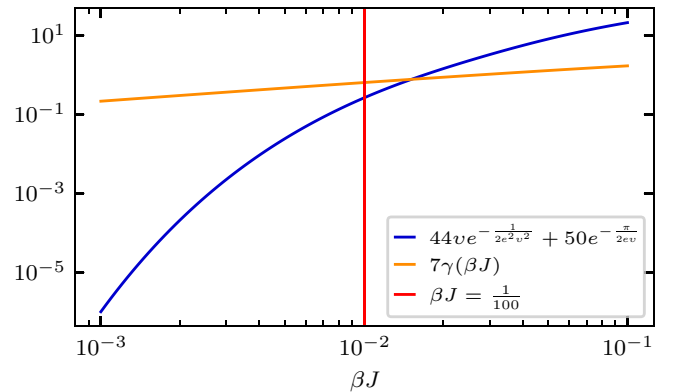


FIG. 12. The contributions in Eq. (D46) for $r = 1$. As long $\beta J \leq \frac{1}{100}$, the term $7\gamma(\beta J)^r$ is dominant.

$$\begin{aligned}
\|G_{a,\alpha}^\beta - c_1 I\|_\infty &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(t)g_2(t')| \|A^{a,\alpha}(-t')A^{a,\alpha}(t') - I\|_\infty dt' dt \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(t)g_2(t')| \|(A^{a,\alpha}(-t') - A^{a,\alpha})A^{a,\alpha}(t') + A^{a,\alpha}(A^{a,\alpha}(t') - A^{a,\alpha})\|_\infty dt' dt \\
&\leq 4J \|A^{a,\alpha}\|_\infty^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(t)g_2(t')| |t'| dt' dt = \frac{e^{\frac{1}{2}}}{\sqrt{\pi}} \beta J \|A^{a,\alpha}\|_\infty^2.
\end{aligned} \tag{D53}$$

This gives

$$\| -i[G_{a,\alpha}^\beta - c_1 I, \cdot] \|_{\infty \rightarrow \infty} \leq \frac{2e^{\frac{1}{2}}}{\sqrt{\pi}} \beta J \|A^{a,\alpha}\|_\infty^2. \tag{D54}$$

Putting Eqs. (D52) and (D54) together, this means we can bound

$$\eta(\beta) = \|\mathcal{L}^{\beta\dagger} - \mathcal{L}^{0\dagger}\|_{\infty \rightarrow \infty} \leq \|\mathcal{D}_{a,\alpha}^{\beta\dagger} - \mathcal{D}_{a,\alpha}^{0\dagger}\|_{\infty \rightarrow \infty} + \| -i[G_{a,\alpha}^\beta - c_1 I, \cdot] \|_{\infty \rightarrow \infty} \leq \left(\frac{2e^{1/2}}{\sqrt{\pi}} + \frac{8e^{1/8}}{\sqrt{2\pi}} \right) \beta J < 6\beta J. \tag{D55}$$

c. Bounding $u(r_0)$

It remains to get a bound for $u(r_0)$ in Eq. (D15),

$$u(r_0) = 5(2r_0 + 1)^{2D} \Delta(r_0) + (5 + 2r_0 + 2(2r_0 + 1)^D) \sum_{l \geq r_0} (2l + 1)^{2D-1} \Delta(l) + 2 \sum_{l' \geq r_0} \sum_{l \geq l'} (2l + 1)^{2D-2} \Delta(l). \tag{D56}$$

We define $s = (\beta J)^{1/D}$. By definition $\gamma(\beta J) < (\beta J)^{1/2} = s^{D/2}$. For $r_0 \geq 1$, and $\beta J \leq \frac{1}{100}$, $\gamma(\beta J) < \frac{1}{2}$ and thus $\Delta(r) < 28\gamma(\beta J)$. We can thus bound

$$\begin{aligned}
\sum_{l \geq r_0} (2l + 1)^{2D-1} \Delta(l) &\leq \sum_{l \geq r_0} (2l + 1)^{2D} \Delta(l) \leq \frac{14}{1 - \gamma(\beta J)} \sum_{l \geq r_0} (2l + 1)^{2D} \gamma(\beta J)^l \leq 28 \sum_{l \geq r_0} ((2l + 1)^4 s^l)^{D/2} \\
&\leq 28 \left(\sum_{l \geq r_0} (2l + 1)^2 s^{l/2} \right)^D.
\end{aligned} \tag{D57}$$

Similarly, we can bound

$$2 \sum_{l' \geq r_0} \sum_{l \geq l'} (2l + 1)^{2D-2} \Delta(l) \leq \sum_{l \geq r_0} (2l + 1)^{2D} \Delta(l) \leq 28 \left(\sum_{l \geq r_0} (2l + 1)^2 s^{l/2} \right)^D. \tag{D58}$$

We have thus

$$\begin{aligned}
u(r_0) &\leq 140((2r_0 + 1)s^{r_0/4})^{2D} + 28(6 + 2r_0 + 2(2r_0 + 1)^D) \left(\sum_{l \geq r_0} (2l + 1)^2 s^{l/2} \right)^D \\
&\leq 140((2r_0 + 1)^2 s^{r_0/2})^D + 140(2r_0 + 1)^D \left(\sum_{l \geq r_0} (2l + 1)^2 s^{l/2} \right)^D.
\end{aligned} \tag{D59}$$

For $s < \frac{1}{2020}$ and $r_0 = 4$, it can be easily shown that $k(r_0, \beta) < 1$ in Eq. (D14), and this concludes the proof of Eq. (D4), with $\beta J < 1/2020^D$.

Finally, we prove the statement for truncated Lindbladians [Eq. (D5)]. For truncated Lindbladians $\mathcal{L}^{\beta,r}$ we have that $\Delta(l) = 0$ for $l > r$. Similarly to Eq. (D14), the exponential decay of the oscillator norm is characterized by

$$k'(r_0, \beta) = 4(2r_0 + 1)^{2D} \eta'(\beta) + u'(r_0). \tag{D60}$$

Here $u'(r)$ is a modified version of $u(r)$, with

$$u'(r_0) = 5(2r+1)^{2D} \Delta(r_0) + (5+2r+2(2r_0+1)^D) \sum_{l \geq r_0}^r (2l+1)^{2D-1} \Delta(l) + 2 \sum_{l' \geq r} \sum_{l \geq l'}^r (2l+1)^{2D-2} \Delta(l) \quad (\text{D61})$$

$$\leq 140((2r_0+1)^2 s^{r_0/2})^D + 140(2r_0+1)^D \left(\sum_{l \geq r_0} (2l+1)^2 s^{l/2} \right)^D, \quad (\text{D62})$$

and $\eta'(\beta) = \eta(\beta)$. For $r_0 > r$, $u'(r_0) = 0$ and $k(r_0, \beta)$ in Eq. (D14) is bound by $\eta'(\beta) \leq 6\beta J = 6s^D$. We see thus that for the truncated case as well, for $r \geq 1$, $\kappa' = k'(r, \beta)$ in Eq. (D14) is less than 1 for $s < 1/2020$. This proves Eq. (D5) as well with $\beta J < 1/2020^D$, concluding the proof.

4. Proof of Theorem D1

In order to apply Lemma D1 for proving Theorem D1, it remains to bound $\|\mathcal{L}^\beta - \mathcal{L}^{\beta,r}\|_{1 \rightarrow 1}$. But this follows immediately from the proof for Lemma D2: note that the bounds of $\mathcal{L}_{a,\alpha}^\beta$ and $\mathcal{L}_{a,\alpha}^{\beta,r}$ derived for the $\|\cdot\|_{\infty \rightarrow \infty}$ norm also hold for the $\|\cdot\|_{1 \rightarrow 1}$ -norm. To see this, compare for instance with the bound in Eq. (D27). We have for the transition part $T_{a,\alpha}^r(\rho) = L_{a,\alpha}^{\beta,r} \rho L_{a,\alpha}^{\beta,r\dagger}$,

$$\begin{aligned} \|T_{a,\alpha}^{r+1} - T_{a,\alpha}^r\|_{1 \rightarrow 1} &= \sup_X \frac{\|T_{a,\alpha}^{r+1}(X) - T_{a,\alpha}^r(X)\|_1}{\|X\|_1} \\ &= \sup_X \frac{\|(L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}) X L_{a,\alpha}^{\beta,(r+1)\dagger} + L_{a,\alpha}^{\beta,r} X (L_{a,\alpha}^{\beta,(r+1)\dagger} - L_{a,\alpha}^{\beta,r\dagger})\|_1}{\|X\|_1}. \end{aligned} \quad (\text{D63})$$

However,

$$\begin{aligned} &\|(L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}) X L_{a,\alpha}^{\beta,(r+1)\dagger}\|_1 \\ &\leq \|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_\infty \|X\|_1 \|L_{a,\alpha}^{\beta,(r+1)\dagger}\|_\infty \end{aligned} \quad (\text{D64})$$

and thus

$$\|T_{a,\alpha}^{r+1} - T_{a,\alpha}^r\|_{1 \rightarrow 1} \leq 2 \|L_{a,\alpha}^{\beta,(r+1)} - L_{a,\alpha}^{\beta,r}\|_\infty \|L_{a,\alpha}^{\beta,(r+1)\dagger}\|_\infty. \quad (\text{D65})$$

The same argument holds for the other parts of the Lindbladian $\mathcal{L}_{a,\alpha}^{\beta,r}$. From this observation, it immediately follows, using the same steps as in Appendix D3:

$$\begin{aligned} \sum_\alpha \|\mathcal{L}_{a,\alpha}^\beta - \mathcal{L}_{a,\alpha}^{\beta,r}\|_{1 \rightarrow 1} &\leq \sum_\alpha \sum_{r'=r}^\infty \|\mathcal{E}_{a,\alpha}^{\beta,r'}\|_{1 \rightarrow 1} \\ &\leq \sum_\alpha 7 \frac{\gamma(\beta J)^{r+1}}{1 - \gamma(\beta J)} = \mathcal{O}((\beta J)^{r/2}). \end{aligned} \quad (\text{D66})$$

Adding contributions from n sites, we get

$$\|\mathcal{L}^\beta - \mathcal{L}^{\beta,r}\|_{1 \rightarrow 1} = \mathcal{O}(n(\beta J)^{r/2}). \quad (\text{D67})$$

Also, since we have proved the mixing time scales as $\log n$, it follows from Lemma D1 that the trace distance between

the steady states of the original and truncated evolution is given by

$$\|\rho_{\beta,r} - \rho_\beta\|_1 = \mathcal{O}((\beta J)^{r/2} n \log n). \quad (\text{D68})$$

This expression concludes our proof of Theorem D1.

APPENDIX E: ERROR BOUND ON LOCAL OBSERVABLES

In this Appendix, we provide the proof for the following theorem.

Theorem E1 (Truncation error of local observables; Theorem 3 restated). Under the assumptions of Theorem 1, consider $\beta < \beta^*$ and let $O_X \in \mathbf{S}_X$ be supported on $X \subseteq \Lambda$. Then there exist constants $\gamma, \eta > 0$, depending only on the model parameters and β , such that $\rho_{\beta,r}$ satisfies

$$|\text{tr}(\rho_{\beta,r} O_X) - \text{tr}(\rho_\beta O_X)| \leq \|O_X\| c(|X|) [e^{-\gamma r} + n e^{-\eta d_X}], \quad (\text{E1})$$

where $c(|X|) = \mathcal{O}(\text{poly}(|X|))$, $n \equiv |\Lambda|$ is the number of qubits, and

$$d_X = \min_{a \in X} \ell(a, \partial \Lambda)$$

is the minimal graph distance from any $a \in X$ to the boundary $\partial\Lambda$.

For a fixed region X , the distance d_X , which is distance between X and to the complement Λ^c of the lattice Λ and is thus the distance to the boundaries, increases with system size, which implies that the term $|\Lambda|v_\eta^{-1}(d_X)$ vanishes as the system becomes large. This term can therefore be interpreted as a finite-size correction [35]. The proof exploits two key features: the rapid-mixing property of the Lindbladian $\mathcal{L}^\beta$ and the quasilocal nature of the perturbation introduced by the truncation. Together, they allow us to invoke the stability theory for local, rapidly mixing dissipative systems developed in Ref. [35]. We summarize these foundational results in Appendix E 1 and present the detailed proof of Theorem E1 in Appendix E 2.

1. Stability of local and rapidly mixing Lindbladians

We will make use of the following stability result, adapted from Ref. [35]:

Theorem E2 (Adapted from Theorem 7 in Ref. [35]). Let $\mathcal{L}$ be a uniform family of local Lindbladians with finite range interactions, a unique fixed point, and satisfying rapid mixing. Consider a perturbation of the form

$$\mathcal{E}^\Lambda = \sum_{a \in \Lambda} \sum_{r' \geq 0} \mathcal{E}_{a,r'},$$

where

- (1) Each $\mathcal{E}_{a,r'}$ acts on $B_a(r')$ and $\|\mathcal{E}_{a,r'}\|_{1 \rightarrow 1} \leq \epsilon e(r')$, where $\epsilon > 0$ is a small parameter characterizing the perturbation strength and $e(r')$ is a rapidly decaying function with $e(r') < 1$.
- (2) $\mathcal{L}_{a,r'} + \mathcal{E}_{a,r'}$ is a Lindbladian itself for all a and $r' \geq 0$.

For any finite lattice Λ , define $\mathcal{L}^\Lambda = \sum_{B_a(r) \subset \Lambda} \mathcal{L}_{a,r}$ the truncated version of $\mathcal{L}$ on Λ

$$T_t = \exp(t\mathcal{L}^\Lambda) \quad (\text{E2})$$

and

$$S_t = \exp(t(\mathcal{L}^\Lambda + \mathcal{E}^\Lambda)). \quad (\text{E3})$$

Then, for any observable O_X supported on $X \subset \Lambda$, the following bound holds:

$$\|T_t^\dagger(O_X) - S_t^\dagger(O_X)\| \leq c(|X|) \|O_X\| (\epsilon + |\Lambda|v_\eta^{-1}(d_X)), \quad (\text{E4})$$

where $d_X = \inf_{a \in X, b \in \Lambda^c} \ell(a, b)$ is the distance between the set X and the complement of Λ with respect to $\mathbb{Z}^d$,

$v_\eta^{-1}(x) = e^{-\eta x}$ and η is a positive constant independent of both Λ and t . Moreover, $v_\eta^{-1}(d_X) \leq (1 + d_X)^{-D-1}$, and $c(|X|)$ is polynomially bounded in $|X|$ and independent of Λ and t .

Before proceeding, we make a few comments concerning the result above. First, note that the result above is for a *uniform family* of Lindbladians. We refer to Definition 3 in Ref. [35] for a precise definition, but informally one can think of a uniform family of Lindbladians as a function that generates a finite-size version $\mathcal{L}^\Lambda$ defined on Λ , starting from a Lindbladian $\mathcal{L}$ defined on $\mathbb{Z}^d$. Such uniform families provide a rigorous framework for dealing with lattice size scaling. We will apply Theorem E2 to the uniform family $\mathcal{L}^{\beta,r}$, for which the finite-size version for any specific finite lattice Λ is defined as $\mathcal{L}^{\beta,r,\Lambda}$. Additionally, we note that Theorem 7 in Ref. [35] is more general than the adaption stated above, in that it allows for non-Lindbladian perturbations, periodic boundary conditions, perturbations with boundary terms, and local Lindbladians with either exponentially decaying or quasilocal interactions (in which case the function v_η will differ). The proof in this case is completely analogous, but we focus here on the case of open boundary conditions to streamline the discussion.

However, none of these generalizations are necessary for our purposes, and so we have omitted details relevant to those cases from our presentation.

Additionally, as pointed out in Ref. [35]:

- (1) For a fixed region A , the distance d_X , which is distance between X and to the complement Λ^c of the lattice Λ and is thus the distance to the boundaries, increases with system size, which implies that the term $|\Lambda|v_\eta^{-1}(d_X)$ vanishes as the system becomes large. This term can therefore be interpreted as a finite-size correction.
- (2) The theorem allows O_X to act on disconnected subsets of the lattice Λ , meaning that it also applies to correlation functions.

2. Proof of Theorem 3

We prove Theorem 3 by applying Theorem E2. Specifically, as mentioned above, we consider as a starting point the uniform family of Lindbladians $\mathcal{L}^{\beta,r}$ for which $\mathcal{L}^{\beta,r,\Lambda}$ is defined via the RHS of Eq. (19). We also define $\mathcal{L}^{\beta,\Lambda}$ as the RHS of Eq. (3). Now, let $\mathcal{E}_{a,\alpha}^{\beta,r}$ be as defined in Appendix D 3 a below Eq. (D16). We then note that if we define

$$\mathcal{E}_a^{\beta,r} = \sum_{\alpha=1}^3 \mathcal{E}_{a,\alpha}^{\beta,r} \quad (\text{E5})$$

and

$$\mathcal{E}^{\beta,r,\Lambda} = \sum_{a \in \Lambda} \sum_{r'=r}^{\infty} \mathcal{E}_a^{\beta,r'} \quad (\text{E6})$$

then via Eqs. (3), (19) and (D17) we have

$$\mathcal{L}^{\beta,\Lambda} = \mathcal{L}^{\beta,r,\Lambda} + \mathcal{E}^{\beta,r,\Lambda}. \quad (\text{E7})$$

As $\mathcal{L}^{\beta,\Lambda}$ is itself a valid Lindbladian, the second condition in Theorem E2 is fulfilled. To apply Theorem E2 for our purposes, we still need:

- (1) To prove that $\mathcal{L}^{\beta,r}$ is rapid mixing and admits a unique fixed point.
- (2) To prove $\|\mathcal{E}_a^{\beta,r'}\|_{1 \rightarrow 1} \leq \epsilon e(r')$ for rapidly decaying $e(r')$, for all $r' \geq r$.

We note that rapid mixing of $\mathcal{L}^{\beta,r}$ is implied by the logarithmic mixing time of $\mathcal{L}^{\beta,r}$, as proven in Theorem 2 [35], and that the existence of a unique fixed point was proven in Theorem 1. For the final remaining condition, note that via the arguments of Appendix D 4 we have already proven that

$$\|\mathcal{E}_{a,\alpha}^{\beta,r'}\|_{1 \rightarrow 1} \leq 7\gamma(\beta J)^{r'} \quad (\text{E8})$$

as long as $\beta J \leq 1/100$. Using this, we straightforwardly have for all $r' \geq r$ that

$$\|\mathcal{E}_a^{\beta,r'}\|_{1 \rightarrow 1} \leq \sum_{\alpha=1}^3 \|\mathcal{E}_{a,\alpha}^{\beta,r'}\|_{1 \rightarrow 1} \quad (\text{E9})$$

$$\leq 21\gamma(\beta J)^{r'} \quad (\text{E10})$$

$$= [21\gamma(\beta J)^r] \gamma(\beta J)^{(r'-r)}. \quad (\text{E11})$$

As such, we can apply Theorem E2 with $\epsilon = 21\gamma(\beta J)^r$ and $e(r') = \gamma(\beta J)^{r'-r}$, which is rapidly decaying for all $r' \geq r$. More specifically, Defining $T_t = \exp\{(t\mathcal{L}^{\beta,r,\Lambda})\}$ and $S_t = \exp\{(t\mathcal{L}^{\beta,\Lambda})\}$ Theorem E2 yields

$$\begin{aligned} & \|T_t^\dagger(O_A) - S_t^\dagger(O_A)\| \\ & \leq c(A) \|O_A\| (21\gamma(\beta J)^r + |\Lambda|v_\eta^{-1}(d_X)), \end{aligned} \quad (\text{E12})$$

which by taking $t \rightarrow \infty$ gives

$$\begin{aligned} & |\text{tr}(\rho_{\beta,r} O_A) - \text{tr}(\rho_\beta O_A)| \\ & \leq c(A) \|O_A\| (21\gamma(\beta J)^r + |\Lambda|v_\eta^{-1}(d_X)), \end{aligned} \quad (\text{E13})$$

which implies the claim of Theorem 3.

APPENDIX F: PROOF OF LEMMA 1

In this Appendix, we present a proof of Lemma 1. We begin with a brief restatement of the relevant definitions and notations from the main text.

We are interested in providing a compilation procedure for each elementary channel $\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})$, which represents the randomized evolution described in Eq. (30). To do this, we define the Hermitian operator

$$O_r^{a,\alpha} := \begin{pmatrix} \sqrt{\tau} G_r^{a,\alpha} & L_r^{a,\alpha\dagger} \\ L_r^{a,\alpha} & \sqrt{\tau} G_r^{a,\alpha} \end{pmatrix}, \quad (\text{F1})$$

where $L_r^{a,\alpha}$ and $G_r^{a,\alpha}$ are the jump operators and associated coherent terms, respectively, as was previously defined in Eq. (21).

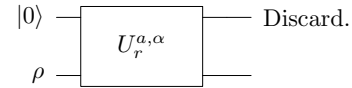
We consider the quantum channel

$$\mathcal{C}_{r,\tau}^{a,\alpha}(\rho) := \text{Tr}_{\text{anc}}(U_r^{a,\alpha}(\rho \otimes |0\rangle\langle 0|_{\text{anc}})U_r^{a,\alpha\dagger}), \quad (\text{F2})$$

with $\text{Tr}_{\text{anc}}(\cdot)$ denoting the partial trace over the ancilla qubit and the unitary $U_r^{a,\alpha}$ describing the evolution over a time interval $\sqrt{\tau}$ as

$$U_r^{a,\alpha} = \exp(-iO_r^{a,\alpha}\sqrt{\tau}). \quad (\text{F3})$$

The circuit for the implementation of the channel described by Eq. (F2) is given by



The error of this implementation is given by Lemma 1, which we restate here.

Lemma F1 (Lemma 1 restated). The channel in Eq. (F2) satisfies

$$\|\mathcal{C}_{r,\tau}^{a,\alpha} - \exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})\|_\diamond = O(\tau^2). \quad (\text{F4})$$

To prove this lemma, we first state the following auxiliary lemma, a standard result in functional analysis (e.g., see Chap. XIII, § 6 of Ref. [38]):

Lemma F2. Let X be a Banach space and let $\mathcal{B}(X)$ denote the space of bounded linear operators on X , equipped with an operator norm $\|\cdot\|_{\mathcal{B}(X)}$. Suppose $F: \mathbb{R} \rightarrow \mathcal{B}(X)$ is $(K+1)$ -times continuously differentiable. Then, for any $x \in \mathbb{R}$, the Taylor expansion of $F(x)$ about

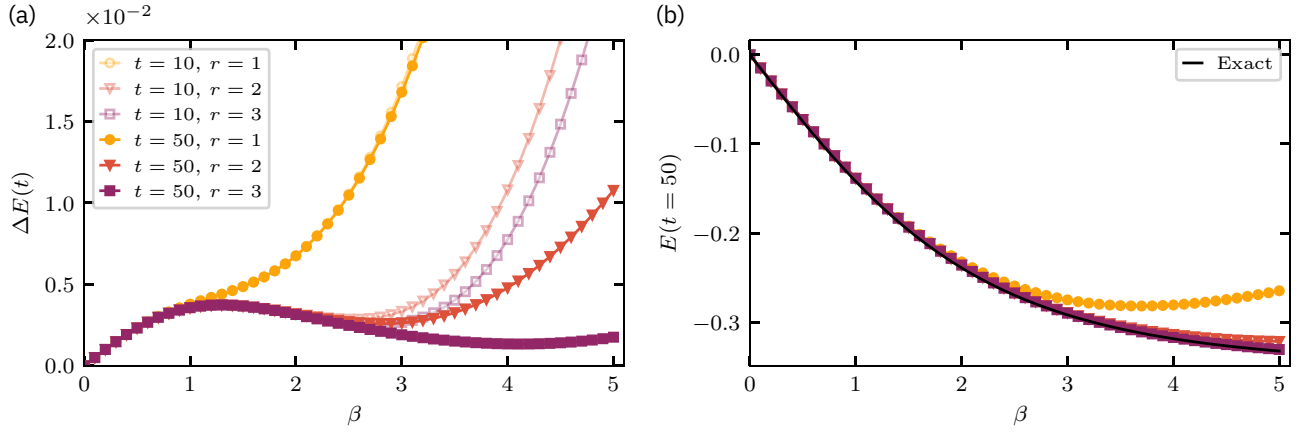


FIG. 13. Energy convergence across temperatures in the transverse-field Ising model. Energy density convergence for the transverse-field Ising Hamiltonian in Eq. (G1) on $n = 12$ sites as a function of inverse temperature β , using a Trotter step size $\tau = 0.1$. (a) Relative energy error $\Delta E(t)$ after $t = 10$ and $t = 50$: errors decrease monotonically with increasing β , while the growing gap at large β signals mixing times $t_{\text{mix}} > 50$; variations across truncation radii r reflect residual truncation inaccuracies. (b) Absolute energy compared to exact diagonalization: a truncation radius of $r = 3$ achieves near-ground-state fidelity up to high β , with slight deviations at larger β due to finite-radius effects.

$x = 0$ satisfies

$$F(x) = \sum_{m=0}^K \frac{x^m}{m!} F^{(m)}(0) + R_{K,F}(x),$$

where $F^{(m)}$ denotes m th derivative and the remainder $R_{K,F}(x)$ satisfies the operator norm bound,

$$\|R_{K,F}(x)\|_{B(X)} \leq \frac{|x|^{K+1}}{(K+1)!} \sup_{s \in [0, |x|]} \|F^{(K+1)}(s)\|_{B(X)}.$$

With this lemma, we prove Lemma F1.

Proof of Lemma F1. A proof excluding a coherent term was given in Ref. [28]. The present version extends that result to incorporate the modification introduced by an additional coherent term.

Let us use the following notation:

$$T_{m,F} := \frac{1}{m!} F^{(m)}(0). \quad (\text{F5})$$

Then, consider the Taylor expansion of both operators in the lemma statement as

$$\begin{aligned} \mathcal{C}_{r,\tau}^{a,\alpha} &= \mathcal{I} + \tau T_{2,\mathcal{C}_{r,\tau}^{a,\alpha}} + R_{3,\mathcal{C}_{r,\tau}^{a,\alpha}}(\sqrt{\tau}), \\ \exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r}) &= \mathcal{I} + \tau T_{1,\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})} + R_{1,\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})}(\tau), \end{aligned} \quad (\text{F6})$$

where $\mathcal{I}$ is the identity superoperator. Note that terms odd in $\sqrt{\tau}$ in the first expansion are canceled due to tracing out

the ancilla qubit. Furthermore, the lowest-order terms in this expansion coincide,

$$T_{1,\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})} = T_{2,\mathcal{C}_{r,\tau}^{a,\alpha}} = \mathcal{L}_{a,\alpha}^{\beta,r}. \quad (\text{F7})$$

Thus, using the triangle inequality, we obtain

$$\begin{aligned} \|\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r}) - \mathcal{C}_{r,\tau}^{a,\alpha}\|_{\diamond} &= \|R_{1,\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})}(\tau) - R_{3,\mathcal{C}_{r,\tau}^{a,\alpha}}(\sqrt{\tau})\|_{\diamond} \\ &\leq \|R_{1,\exp(\tau \mathcal{L}_{a,\alpha}^{\beta,r})}(\tau)\|_{\diamond} + \|R_{3,\mathcal{C}_{r,\tau}^{a,\alpha}}(\sqrt{\tau})\|_{\diamond} \leq C\tau^2, \end{aligned} \quad (\text{F8})$$

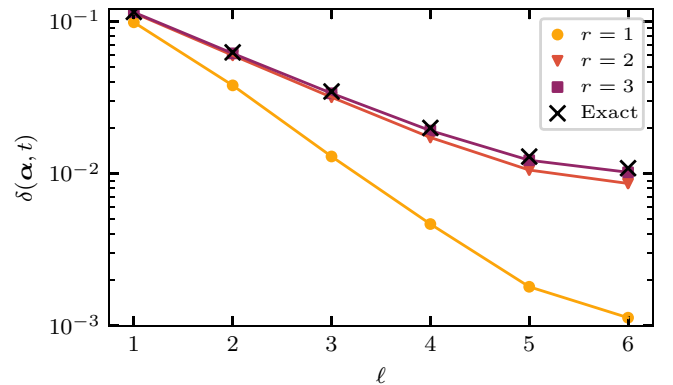


FIG. 14. Spatial profile of the two-spin correlators. The correlator $\delta(\alpha, t)$, for the transverse-field Ising model defined in Eq. (G1). The Trotter step size is $\Delta t = 0.1$, $\beta = 3.0$. Results of the correlator are shown after $t = 10$ and truncation radius $r = 1$ (yellow), $r = 2$ (orange), $r = 3$ (red), and the exact results (black crosses). For a truncation radius $r \geq 2$, the proposed protocol can reproduce the correlator at large distances ℓ .

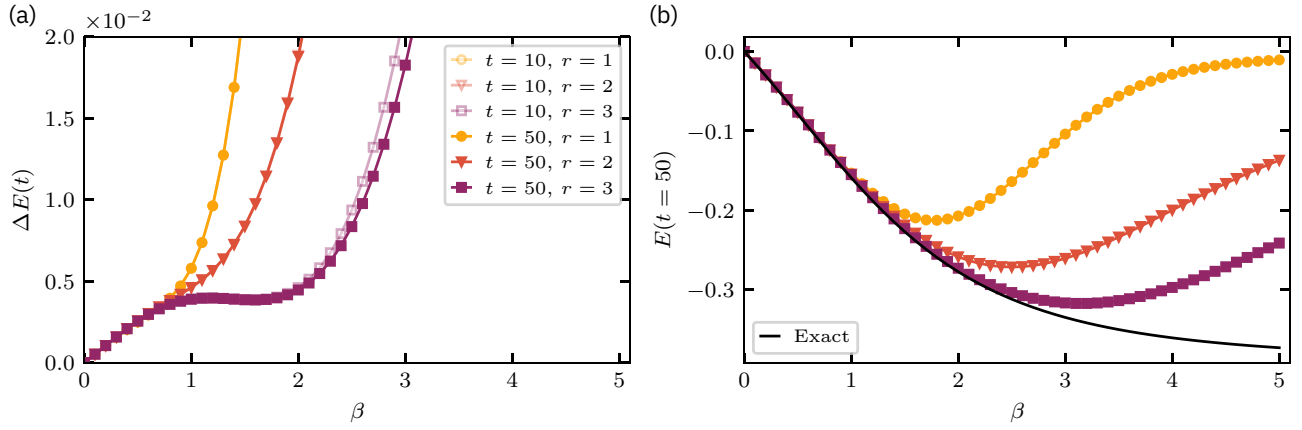


FIG. 15. Energy convergence across temperatures in the XXZ model. Energy density convergence for the XXZ Hamiltonian in Eq. (G2) on $n = 12$ sites as a function of inverse temperature β , using a Trotter step size $\tau = 0.1$. (a) Relative energy error $\Delta E(t)$ after $t = 10$ and $t = 50$ steps: the error decreases monotonically with increasing β , while the widening gap between the two curves highlights increasingly slow mixing at larger β . Even a truncation radius $r = 3$ fails to achieve rapid convergence at long times. (b) Absolute energy compared with exact diagonalization: deviations at high β (low temperatures) reflect residual truncation errors and extended mixing times.

where we used Lemma F2 on both terms separately and C is a constant. This completes the proof of Lemma F1. ■

APPENDIX G: ADDITIONAL MODELS AND NUMERICS DETAILS

1. Additional models

We consider two additional spin-chain models. The objective is to show that our approach remains effective for integrable models and models with symmetries, indicating that it does not rely on full thermalization.

a. 1D transverse-field Ising model

In the following, we present results for a simplified version of the model in Eq. (41), namely the transverse-field Ising model with periodic boundary conditions,

$$H = \sum_{\langle i,j \rangle} S_i^z S_j^z + g \sum_{i=1}^n S_i^x. \quad (\text{G1})$$

Throughout, we set $g = 0.6$, placing the system within the ordered phase of the zero-temperature quantum phase transition at $g = 1.0$. Unlike the mixed-field Ising model, the transverse-field Ising model is integrable and can be mapped to a one-dimensional system of spinless free fermions via a Jordan–Wigner transformation [56]. However, the thermalizing Lindbladian dynamics we consider enables dynamics among the different sectors and thus breaks the integrability.

Figure 13 shows the convergence of the energy toward the thermal value. As with the mixed-field Ising model, we observe rapid convergence to the Gibbs-state energy density using a small truncation radius, $r \leq 3$. Similarly,

Fig. 14 demonstrates that correlation functions are accurately reproduced with the same truncation. These results confirm that the algorithm remains effective even in integrable settings.

b. XXZ model

Finally, we present results for the antiferromagnetic XXZ model

$$H = \sum_{\langle i,j \rangle} (S_i^x S_{i+1}^x + S_i^y S_j^y) + \Delta \sum_{\langle i,j \rangle} S_i^z S_j^z, \quad (\text{G2})$$

with $\Delta = 0.6$, which puts the model far from the isotropic point $\Delta = 1$ and in the middle of the gapless phase that

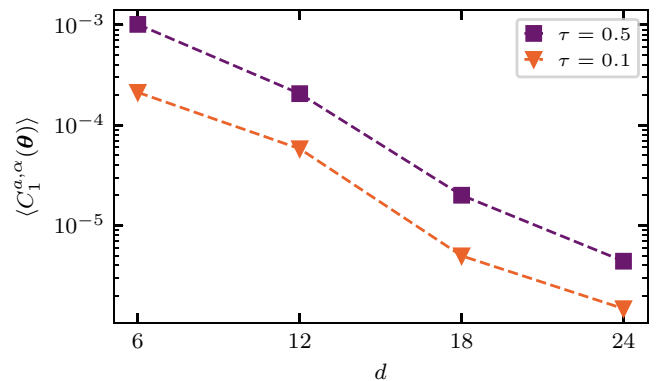


FIG. 16. Variational compilation accuracy as a function of template depth. Final loss values obtained by optimizing circuit templates of depth d for different Trotter step sizes τ , for fixed truncation radius $r = 1$. Each curve corresponds to a fixed τ , showing a systematic, monotonic decrease in compilation error as the template depth increases.

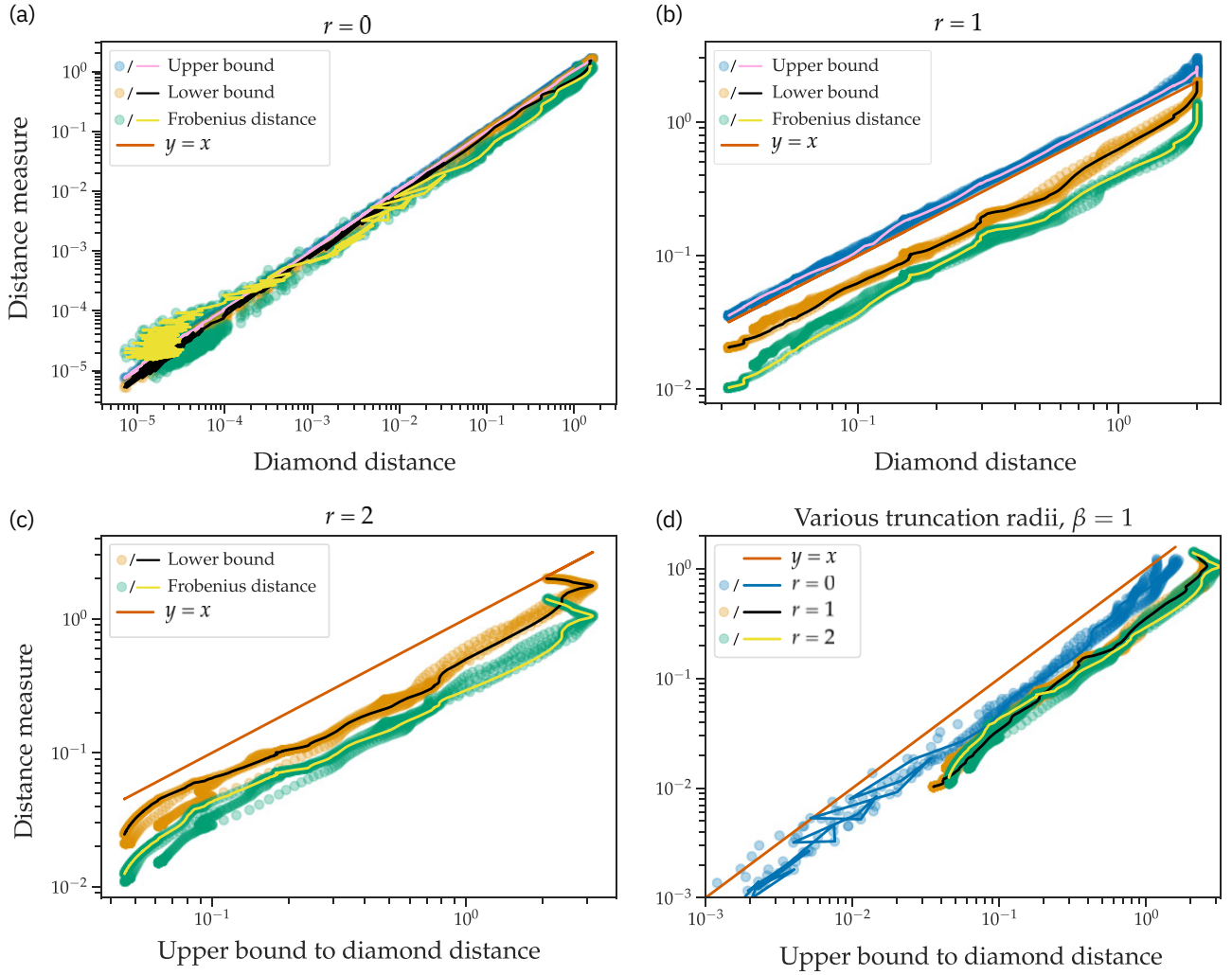


FIG. 17. Correlation between diamond-distance proxies. Distance measures between the target quantum channel and the trained circuit during variational compilation, shown for five independent trajectories illustrating the evolution of the distance under optimization from different random circuit initializations. Dots represent raw data points, while solid lines depict a representative example trajectory. Results are shown for different gadget constructions: (a) a one-qubit gadget ($r = 0$), (b) a T -gadget ($r = 1$), and (c) a long- T -gadget ($r = 2$). (d) Same data as (a)–(c), focusing on the modified Frobenius distance and the upper bound to the diamond distance for various radii r of the channel support.

exhibits quasi-long-range correlations at zero temperature. The results are shown in Fig. 15. While the results indicate that our algorithm still works even in this case for high temperatures, the results for large inverse temperatures $\beta > 1$ show more significant deviations from the exact thermal state than in comparison to the previous models.

2. Renormalization of the Lindbladian

In Sec. VC and the corresponding Fig. 8, we compare the effect of different envelope functions $q(v)$ in Eq. (6). For better comparison, we fix the jump operators for different envelope functions to have the same norm as the Gaussian case in Eq. (11).

Let $L_{a,\alpha}^{\beta,r}|_{q(v)}$ and $G_{a,\alpha}^{\beta,r}|_{q(v)}$ denote the jump operators and coherent terms, respectively, introduced in Eq. (21) evaluated using the envelope function $q(v)$. Then, we compute the mean of the norms for all operators,

$$\varphi_{\beta,r}[q(v)] := \frac{1}{3} \sum_{\alpha=1}^3 \left\| L_{a,\alpha}^{\beta,r}|_{q(v)} \right\|_2. \quad (\text{G3})$$

Note that the norm does not depend on the qubit index a due to translational invariance. The normalized jump operators are then obtained by

$$L_{a,\alpha}^{\beta,r}|_{q(v),\text{norm.}} = \frac{\varphi_{\beta,r} \left[\exp \left(-\frac{(\beta v)^2}{8} \right) \right]}{\varphi_{\beta,r}[q(v)]} L_{a,\alpha}^{\beta,r}|_{q(v)}. \quad (\text{G4})$$

In order to maintain detailed balance, the coherent terms should be normalized as

$$G_{a,\alpha}^{\beta,r}|_{q(v),\text{norm.}} = \left(\frac{\varphi_{\beta,r}[\exp(-\frac{(\beta v)^2}{8})]}{\varphi_{\beta,r}[q(v)]} \right)^2 G_{a,\alpha}^{\beta,r}|_{q(v)}. \quad (\text{G5})$$

The performance of the renormalized jump operators is shown in Fig. 8.

3. Optimization details

In this Appendix subsection, we provide details of the optimization procedure of the compilation using the Adam optimizer [44].

As the hyperparameters, we choose a learning rate $l_r = 10^{-3}$, first-moment decay rate $\beta_1 = 0.99$, second-moment decay rate $\beta_2 = 0.99$, and regularization $\epsilon = 10^{-3}$. For a given template, we initialize 50 different random configurations, iterate 8000 steps and choose the best configuration after the optimization process. Figure 16 shows the optimal outcomes, averaged over the unitaries $U_r^{a,1}$, $U_r^{a,2}$, and $U_r^{a,3}$.

4. Choice of loss function in variational compilation

In this work we consider the diamond distance between channels and two proxies for the diamond distance, namely its upper and lower bounds. This is because while computing the diamond distance of two channels from their specified Choi-Jamiołkowski matrix is formally efficient via semidefinite programs [57], its upper and lower bounds [58] are even more efficiently computable and particularly amenable for use in a framework employing automatic differentiation, such as JAX. Specifically, suppose a channel $\mathcal{C} : D_{X_1} \rightarrow D_{X_2}$ mapping input states with dimension d_1 to states with dimension d_2 has a Choi-Jamiołkowski matrix given by

$$J(\mathcal{C}) = \frac{1}{d_1} \sum_{i,j=1}^{d_1} |i\rangle\langle j| \otimes \mathcal{C}(|i\rangle\langle j|). \quad (\text{G6})$$

The upper and lower bounds on the diamond distance $\|\mathcal{C}_1 - \mathcal{C}_2\|_\diamond$ between two channels $\mathcal{C}_1$ and $\mathcal{C}_2$ can be written in terms of the difference in their Choi-Jamiołkowski matrix [59] $J(\mathcal{C}_1) - J(\mathcal{C}_2) =: J(\Delta\mathcal{C})$ [58]. Specifically, the lower bound is

$$\|\mathcal{C}_1 - \mathcal{C}_2\|_\diamond \geq \|J(\Delta\mathcal{C})\|_1, \quad (\text{G7})$$

while the upper bound is

$$\|\mathcal{C}_1 - \mathcal{C}_2\|_\diamond \leq 2^{d_1} \|\text{Tr}_2 J(\Delta\mathcal{C})\|_\infty, \quad (\text{G8})$$

where Tr_2 denotes tracing out the second register in Eq. (G6), namely the register to which the channel is applied, and $|A|$ denotes $\sqrt{A^\dagger A}$.

In Fig. 17, we compare these distance measures with each other for gadgets involving $2r + 1$ system qubits, with the central qubit coupled to one additional ancilla qubit. These plots show the progress of variational compilation with random initialization for five different initializations each. The “Frobenius distance” refers to the square root of the quantity in Eq. (40). We observe that the various distance proxies are well-correlated with each other and with the diamond distance across a large range of values. In Fig. 17(c), we do not compute the diamond distance and instead use its upper bound as a proxy, having built evidence in Figs. 17(a) and 17(b) that it correlates very well with the diamond distance.

In Fig. 17(d), we plot the (modified) Frobenius distance and the upper bound to the diamond distance from these same runs. We illustrate data from different truncation radii on the same plot to emphasize that there is no dimension-dependent loss in going between the Frobenius distance and the upper bound to the diamond distance.

-
- [1] M. Vojta, Quantum phase transitions, *Rep. Prog. Phys.* **66**, 2069 (2003).
 - [2] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
 - [3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University, Cambridge, 2010).
 - [4] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme *et al.*, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
 - [5] Á. M. Alhambra, Quantum many-body systems in thermal equilibrium, *PRX Quantum* **4**, 040201 (2023).
 - [6] C. Gogolin and J. Eisert, Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems, *Rep. Prog. Phys.* **79**, 056001 (2016).
 - [7] F. G. S. L. Brandão, A. Kalev, T. Li, C. Y.-Y. Lin, K. M. Svore, and X. Wu, in *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)* (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2019), Vol. 132, p. 27:1.
 - [8] M. Kieferová and N. Wiebe, Tomography and generative training with quantum Boltzmann machines, *Phys. Rev. A* **96**, 062327 (2017).
 - [9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087 (1953).
 - [10] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times* (American Mathematical Society, Providence, Rhode Island, 2017), 2nd ed.
 - [11] D. Poulin and P. Wocjan, Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer, *Phys. Rev. Lett.* **103**, 220502 (2009).

- [12] O. Shtanko and R. Movassagh, Preparing thermal states on noiseless and noisy programmable quantum processors, [arXiv:2112.14688](#).
- [13] M. Motta, C. Sun, A. T. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. Brandao, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, *Nat. Phys.* **16**, 205 (2020).
- [14] M. Consiglio, in *Numerical Computations: Theory and Algorithms: 4th International Conference, NUMTA 2023* (Springer-Verlag, Berlin, Heidelberg, 2023), p. 56.
- [15] A. Deshpande, M. Hinsche, S. Najafi, K. Sharma, R. Sweke, and C. Zoufal, Dynamic parameterized quantum circuits: Expressive and barren-plateau free, [arXiv:2411.05760](#).
- [16] Y. Ilin and I. Arad, Dissipative variational quantum algorithms for Gibbs state preparation, *IEEE Trans. Quantum Eng.* **6**, 1 (2025).
- [17] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, Quantum Metropolis sampling, *Nature* **471**, 87 (2011).
- [18] M.-H. Yung and A. Aspuru-Guzik, A quantum–quantum Metropolis algorithm, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 754 (2012).
- [19] E. B. Davies, Markovian master equations, *Commun. Math. Phys.* **39**, 91 (1974).
- [20] P. Rall, C. Wang, and P. Wocjan, Thermal state preparation via rounding promises, *Quantum* **7**, 1132 (2023).
- [21] C.-F. Chen, M. J. Kastoryano, F. G. S. L. Brandão, and A. Gilyén, Quantum thermal state preparation, [arXiv:2303.18224](#).
- [22] C.-F. Chen, M. J. Kastoryano, and A. Gilyén, An efficient and exact noncommutative quantum Gibbs sampler, [arXiv:2311.09207](#).
- [23] C. Rouzé, D. Stilck França, and Á. M. Alhambra, Optimal quantum algorithm for Gibbs state preparation, *Phys. Rev. Lett.* **136**, 060601 (2026).
- [24] Š. Šmíd, R. Meister, M. Berta, and R. Bondesan, Polynomial time quantum Gibbs sampling for Fermi-Hubbard model at any temperature, *Nat. Commun.* **16**, 10736 (2025).
- [25] Y. Tong and Y. Zhan, Fast mixing of weakly interacting fermionic systems at any temperature, *PRX Quantum* **6**, 030301 (2025).
- [26] Z. Ding, B. Li, and L. Lin, Efficient quantum Gibbs samplers with Kubo–Martin–Schwinger detailed balance condition, *Commun. Math. Phys.* **406**, 67 (2025).
- [27] A. Gilyén, C.-F. Chen, J. F. Doriguello, and M. J. Kastoryano, Quantum generalizations of Glauber and Metropolis dynamics, [arXiv:2405.20322](#).
- [28] H. Chen, B. Li, J. Lu, and L. Ying, A randomized method for simulating Lindblad equations and thermal state preparation, *Quantum* **9**, 1917 (2025).
- [29] E. Brunner, L. Coopmans, G. Matos, M. Rosenkranz, F. Sauvage, and Y. Kikuchi, Lindblad engineering for quantum Gibbs state preparation under the eigenstate thermalization hypothesis, *Quantum* **9**, 1843 (2025).
- [30] D. Hahn, S. A. Parameswaran, and B. Placke, Provably efficient quantum thermal state preparation via local driving, [arXiv:2505.22816](#).
- [31] C. Rouzé, D. Stilck França, and Á. M. Alhambra, Efficient thermalization and universal quantum computing with quantum Gibbs samplers, [arXiv:2403.12691](#).
- [32] E. B. Davies, Markovian master equations. II, *Math. Ann.* **219**, 147 (1976).
- [33] While terminology may vary across the literature, we use the term *Lindblad operators* or *Lindbladians* to refer to the full operators $\mathcal{L}^\beta$ and $\mathcal{L}_{a,\alpha}^\beta$, including both their dissipative and coherent components. This contrasts with definitions that reserve the term only for the dissipative part.
- [34] Note that Ding *et al.* [26] introduced an additional regularization for large frequencies ν to ensure a well-defined time representation in the thermodynamic limit. We avoid the necessity of regularization by choosing either rapidly decaying filter functions or restricting to small subsystems, as it becomes clear below.
- [35] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia, Stability of local quantum dissipative systems, *Commun. Math. Phys.* **337**, 1275 (2015).
- [36] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, Dissipative quantum Church-Turing theorem, *Phys. Rev. Lett.* **107**, 120501 (2011).
- [37] J. Han, W. Cai, L. Hu, X. Mu, Y. Ma, Y. Xu, W. Wang, H. Wang, Y. P. Song, C.-L. Zou, and L. Sun, Experimental simulation of open quantum system dynamics via Trotterization, *Phys. Rev. Lett.* **127**, 020504 (2021).
- [38] S. Lang, *Real and Functional Analysis*, 3rd ed., Graduate Texts in Mathematics, Vol. 142 (Springer-Verlag, New York, 1993).
- [39] Z. Ding, X. Li, and L. Lin, Simulating open quantum systems using Hamiltonian simulations, *PRX Quantum* **5**, 020332 (2024).
- [40] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [41] M. S. J. Tepaske, D. Hahn, and D. J. Luitz, Optimal compression of quantum many-body time evolution operators into brickwall circuits, *SciPost Phys.* **14**, 073 (2023).
- [42] H. Kim and D. A. Huse, Ballistic spreading of entanglement in a diffusive nonintegrable system, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [43] S. Hesselmann and S. Wessel, Thermal Ising transitions in the vicinity of two-dimensional quantum critical points, *Phys. Rev. B* **93**, 155157 (2016).
- [44] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](#).
- [45] V. Shende, S. Bullock, and I. Markov, Synthesis of quantum-logic circuits, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **25**, 1000 (2006).
- [46] J. Lloyd, A. A. Michailidis, X. Mi, V. Smelyanskiy, and D. A. Abanin, Quasiparticle cooling algorithms for quantum many-body state preparation, *PRX Quantum* **6**, 010361 (2025).
- [47] Y. Zhan, Z. Ding, J. Huhn, J. Gray, J. Preskill, G. K. Chan, and L. Lin, Rapid quantum ground state preparation via dissipative dynamics, *Phys. Rev. X* **16**, 011004 (2026).
- [48] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, *Commun. Math. Phys.* **28**, 251 (1972).

- [49] M. B. Hastings, Lieb-Schultz-Mattis in higher dimensions, *Phys. Rev. B* **69**, 104431 (2004).
- [50] M. B. Hastings and T. Koma, Spectral gap and exponential decay of correlations, *Commun. Math. Phys.* **265**, 781 (2006).
- [51] B. Nachtergaele and R. Sims, Lieb-Robinson bounds and the exponential clustering theorem, *Commun. Math. Phys.* **265**, 119 (2006).
- [52] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, Quantum algorithm for simulating real time evolution of lattice Hamiltonians, *SIAM J. Comput.* **52**, FOCS18-250 (2023).
- [53] K. Temme and M. J. Kastoryano, How fast do stabilizer Hamiltonians thermalize? [arXiv:1505.07811](https://arxiv.org/abs/1505.07811).
- [54] Note that we use a different normalization convention for the Choi-Jamiolkowski matrix from the authors in Ref. [58], giving us different normalization factors in the upper and lower bounds.
- [55] When $r > \max_{b \in \Lambda} \ell(a, b)$ we define the error term to be zero.
- [56] S. Suzuki, J. ichi Inoue, and B. K. Chakrabarti, *Quantum Ising Phases and Transitions in Transverse Ising Models* (Springer, Berlin, 2012), Vol. 862.
- [57] J. Watrous, Semidefinite programs for completely bounded norms, [arXiv:0901.4709](https://arxiv.org/abs/0901.4709).
- [58] I. Nechita, Z. Puchała, Ł. Paweł, and K. Życzkowski, Almost all quantum channels are equidistant, *J. Math. Phys.* **59**, 052201 (2018).
- [59] The first term decreases slower with increasing r than the other two terms, therefore it suffices to check $r = 1$.