

# Measuring the Loschmidt Amplitude for Finite-Energy Properties of the Fermi-Hubbard Model on an Ion-Trap Quantum Computer


Kévin Hémary<sup>1,\*</sup>, Khaldoon Ghanem,<sup>1</sup> Eleanor Crane,<sup>1,2</sup> Sara L. Campbell,<sup>3</sup> Joan M. Dreiling,<sup>3</sup> Caroline Figgatt<sup>3</sup>, Cameron Foltz<sup>3</sup>, John P. Gaebler<sup>3</sup>, Jacob Johansen,<sup>3</sup> Michael Mills<sup>3</sup>, Steven A. Moses,<sup>3</sup> Juan M. Pino,<sup>3</sup> Anthony Ransford,<sup>3</sup> Mary Rowe,<sup>3</sup> Peter Siegfried<sup>3</sup>, Russell P. Stutz,<sup>3</sup> Henrik Dreyer,<sup>1</sup> Alexander Schuckert<sup>1,2</sup> and Ramil Nigmatullin<sup>4</sup>

<sup>1</sup>Quantinuum, Leopoldstrasse 180, Munich 80804, Germany

<sup>2</sup>Joint Quantum Institute and Joint Center for Quantum Information and Computer Science, University of Maryland and National Institute of Standards and Technology (NIST), College Park, Maryland 20742, USA

<sup>3</sup>Quantinuum, 303 S Technology Ct, Broomfield, Colorado 80021, USA

<sup>4</sup>Quantinuum, 13–15 Hills Road, Cambridge CB2 1NL, United Kingdom

 (Received 9 October 2023; revised 29 May 2024; accepted 3 July 2024; published 5 August 2024)

Calculating the equilibrium properties of condensed-matter systems is one of the promising applications of near-term quantum computing. Recently, hybrid quantum-classical time-series algorithms have been proposed to efficiently extract these properties from a measurement of the Loschmidt amplitude  $\langle \psi | e^{-i\hat{H}t} | \psi \rangle$  from initial states  $|\psi\rangle$  and a time evolution under the Hamiltonian  $\hat{H}$  up to short times  $t$ . In this work, we study the operation of this algorithm on a present-day quantum computer. Specifically, we measure the Loschmidt amplitude for the Fermi-Hubbard model on a 16-site ladder geometry (32 orbitals) on the Quantinuum H2-1 trapped-ion device. We assess the effect of noise on the Loschmidt amplitude and implement algorithm-specific error-mitigation techniques. By using a thus-motivated error model, we numerically analyze the influence of noise on the full operation of the quantum-classical algorithm by measuring expectation values of local observables at finite energies. Finally, we estimate the resources needed for scaling up the algorithm.

DOI: [10.1103/PRXQuantum.5.030323](https://doi.org/10.1103/PRXQuantum.5.030323)

## I. INTRODUCTION

Calculating the properties of quantum matter in equilibrium is at the heart of condensed-matter and high-energy physics as well as quantum chemistry. In particular, models containing interacting fermions are key to understanding high-temperature superconductivity [1] and the low-energy properties of quantum chromodynamics [2]. However, despite decades of method development, it remains challenging for classical methods to calculate equilibrium properties of high-dimensional systems with a sign problem such as spin models on frustrated lattices and fermionic models. A paradigmatic example of such systems is the two-dimensional (2D) Fermi-Hubbard (FH) model [3,4]. It has attracted a tremendous amount of interest due to its rich but partially understood phase diagram

[5–7] and its potential application to high-temperature superconductivity [8–10].

In recent years, analogue quantum simulators have established themselves as a complementary means of studying equilibrium properties of fermions [11–13], although both reaching low enough energies and tuning the Hamiltonian beyond a restricted parameter regime remain challenging. Extensive progress has recently been made in the size and control of digital quantum computers, potentially leading to a highly flexible tool for solving high-dimensional fermionic problems. Although many ground-state studies have been performed [14–16], only a few demonstrations of experimentally scalable *finite-energy* or *finite-temperature* quantum algorithms have been carried out so far [17,18].

Recently, *time-series algorithms* have been suggested as an efficient way to obtain equilibrium observables in quantum computers [19,20]. These algorithms require access to only short-time dynamics—i.e., low-depth circuits—on the quantum computer while the equilibrium properties are obtained by classical postprocessing. Despite this relative simplicity, the execution of time-series algorithms on current quantum computers is still challenging due to the

\*Contact author: kevin.hemery@quantinuum.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.

requirement to measure the Loschmidt amplitude

$$G_\psi(t) = \langle \psi | e^{-i\hat{H}t} | \psi \rangle, \quad (1)$$

where  $\hat{H}$  is the Hamiltonian,  $|\psi\rangle$  is an initial state, and  $t$  is time. Indeed, the existing experimental methods for the measurement of Eq. (1) require the measurement of a global observable, making them particularly susceptible to noise and requiring very precise quantum hardware. In this work, we experimentally assess the feasibility on current quantum hardware of the quantum subroutine of the time-series algorithm of Ref. [19]—the computation of the Loschmidt amplitude—for the simulation of the FH model. To this aim, we carry out an experiment in Quantinuum’s 32-qubit digital quantum computer [21]. We analyze the effect of the noise present on the hardware, implement error-mitigation strategies, and extrapolate our results to evaluate the resource requirements needed to scale up the algorithm to larger system sizes. While we study the Loschmidt amplitude in the context of time-series algorithms, note that the kind of interferometry experiment we have performed in this work has important uses in other quantum algorithms [22], notably in quantum phase estimation algorithms [23–27], which has been demonstrated on Quantinuum hardware on small systems with error detection [28].

Our main findings are twofold. First, a quantum computer with an average two-qubit gate fidelity of 0.998, low state-preparation-and-measurement (SPAM) error, as well as all-to-all connectivity—such as the H2 device—allows for the extraction of physical properties of the FH model at finite energies using times-series algorithms. Second, scaling to the classically intractable problems is expensive without further improvement due to the large shot overhead associated with error mitigation as well as the cost of performing Monte Carlo sampling.

To begin with, we summarize in Sec. II the algorithm proposed in Ref. [19] and the protocol we use to measure the Loschmidt amplitude. In Sec. III, we introduce the FH model and explain how we map fermions to qubits in order to perform the dynamics on the digital quantum computer. In Sec. IV, we measure the Loschmidt amplitude of a product initial state for the FH model on the ladder geometry and apply error-mitigation schemes to the data obtained from the quantum device. Then, in Sec. V, we simulate the full operation of the algorithm using matrix-product-state simulations and test the sensitivity of the algorithm to the presence of noise, as it is unlikely that all errors can be mitigated in the near future. Finally, motivated by these results, we evaluate the feasibility of this algorithm and discuss its prospects for quantum advantage in Sec. VI.

## II. ALGORITHM

### A. Review of the time-series algorithm for the microcanonical ensemble

To compute observables of excited states, we implement the quantum subroutine of an algorithm put forward in Ref. [19]. The underlying idea behind this algorithm is that the expectation value of an energy-filtered state with low variance will approach the microcanonical expectation value, even if the width of the filter does not tend toward zero. This can be understood in light of the eigenstate-thermalization hypothesis (ETH) [29–31], which predicts that the expectation values of few body-observables are a smooth function of energy, implying that they do not change abruptly within a small energy window. Although such a low-energy-variance state is difficult to prepare directly on the quantum computer, one can nevertheless use a cosine-filter operator (see also Ref. [20], for a different perspective akin to Wick rotation), which can be decomposed into a sum of time-evolution operators. For convenience, the main steps of the calculation presented in Ref. [19] are reproduced in Appendix A.

The central quantity required for the algorithm is the Loschmidt amplitude given in Eq. (1). We explain Sec. II B how to efficiently measure this quantity. From the Loschmidt amplitude measured at different times, one can approximately calculate the filtered density of states  $D_{\psi,\delta}(E)$  (see Appendix A), defined as

$$D_{\psi,\delta}(E) = \langle \psi | e^{-\frac{(\hat{H}-E)^2}{2\delta^2}} | \psi \rangle. \quad (2)$$

$D_{\psi,\delta}(E)$  can be understood as a weighted sum of the overlaps of  $|\psi\rangle$  with the eigenstates inside a Gaussian energy filter of width  $\delta$  centered around the energy  $E$ . The longer one performs the time evolution, the smaller the width of the filter  $\delta$  becomes. Supposing that one is interested in an observable diagonal within the  $\hat{Z}$ -product-state basis  $\{|\psi_p\rangle, p = 1, \dots, 2^N\}$ , one can compute the microcanonical expectation value in the following way:

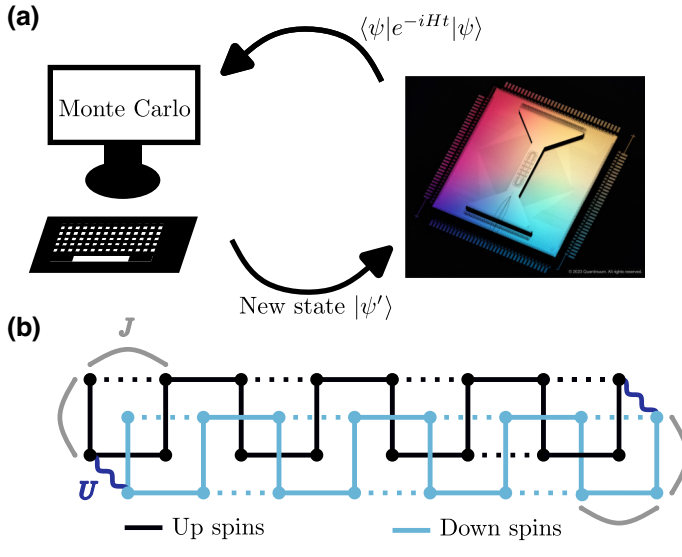
$$\langle \hat{O}_\delta(E) \rangle = \frac{\sum_{|\psi_p\rangle} D_{\psi_p,\delta}(E) O_p}{\sum_{|\psi_p\rangle} D_{\psi_p,\delta}(E)}, \quad (3)$$

where  $O_p = \langle \psi_p | \hat{O} | \psi_p \rangle$  is the corresponding eigenvalue of  $\hat{O}$ . Instead of calculating  $D_{\psi_p,\delta}(E)$  for every  $|\psi_p\rangle$  in order to evaluate the sum in Eq. (3), one can simply use a classical sign-problem-free Monte Carlo algorithm to efficiently sample from the distribution, provided that one can measure the Loschmidt amplitude on a quantum device. In this work, we assess the effect of the noise present on current hardware on the program outlined above and estimate the resources necessary for its application to larger system sizes.

## B. GHZ-like state preparation for the measurement of the Loschmidt amplitude

The Loschmidt amplitude given in Eq. (1) corresponds to the (generally complex) overlap of a time-evolved state with the initial state. This quantity can be calculated on a digital quantum computer using the Hadamard test, where the dynamics  $U(t)$  of a quantum system are controlled on an ancilla qubit that starts in a superposition and interferes with both evolved and nonevolved systems when it is rotated out of the superposition [32]. However, this method is costly in terms of entangling gates, as it requires controlling every gate of  $U(t)$  on an ancilla qubit. Alternatively, all qubits can be prepared in a Greenberger-Horne-Zeilinger– (GHZ) like state corresponding to a superposition between the state of interest and a state that does not evolve under application of the Hamiltonian, i.e., an eigenstate  $|\phi\rangle$ . The Hamiltonian is then applied and the qubits rotated back into the original basis, causing interferometry between the time-evolved and initial states [19,20]—as illustrated in Fig. 1. More precisely, if we define the states  $|\pm(\phi, \psi_0, \xi)\rangle = \frac{1}{\sqrt{2}}(|\phi\rangle \pm e^{i\xi}|\psi_0\rangle)$ , the real part of the Loschmidt amplitude can be extracted from measurements of  $|\langle +(\phi, \psi_0, Et)|\hat{U}(t)|\pm(\phi, \psi_0, Et)\rangle|^2$  in the following way. Consider the quantities

$$\begin{aligned} p_0(t) &:= |\langle +(\phi, \psi, Et)|\hat{U}(t)|+(\phi, \psi, Et)\rangle|^2 \\ &= \frac{1}{4}(1 + |G_\psi(t)|^2 + 2\text{Re}\{G_\psi(t)e^{iEt}\}) \end{aligned} \quad (4a)$$



and

$$\begin{aligned} p_\pi(t) &:= |\langle +(\phi, \psi, Et)|\hat{U}(t)|-(\phi, \psi, Et)\rangle|^2 \\ &= \frac{1}{4}(1 + |G_\psi(t)|^2 - 2\text{Re}\{G_\psi(t)e^{iEt}\}). \end{aligned} \quad (4b)$$

We then find that

$$\text{Re}(G_\psi(t)e^{iEt}) = p_0(t) - p_\pi(t). \quad (5)$$

We have thus reduced the problem of measuring the non-Hermitian observable  $\hat{U}(t)$  to measuring the real quantities  $p_0$  and  $p_\pi$  that correspond to the probability of the circuit shown in Fig. 1(c) to output certain bit strings. We review the details of this technique in Appendix E. Compared to the conditional dynamics technique outlined before, this method reduces the circuit depth by a significant factor, as highlighted in Table I. Note that a GHZ-state preparation can be achieved using a constant-depth circuit with midcircuit measurement or a log-depth circuit without midcircuit measurement. A GHZ-state preparation with 32 qubits has been carried out with 82% fidelity on the device used in this study [21].

We note that in the case in which the initial states are product states, it is also possible to apply a series of single-qubit interferometry experiments in order to circumvent the GHZ-state preparation [19,20]. However, this method introduces a shot overhead proportional to the

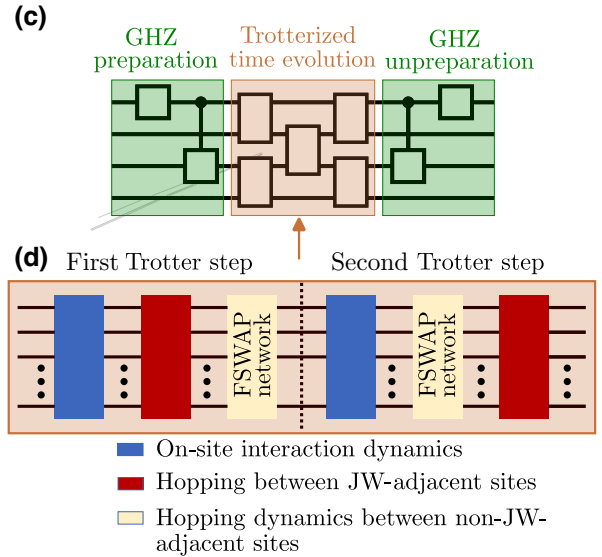


FIG. 1. The hybrid quantum-classical algorithm for the finite-energy properties of the FH model on the H2 quantum computer. (a) An illustration of the quantum-classical loop in the algorithm of Ref. [19], with a modified picture of the Quantinuum H2 surface-trap microchip used in our work. (b) The mapping of the FH model on a  $2 \times 8$  lattice onto 32 qubits. The spin-up (spin-down) fermions are encoded in the black (blue) half of the system. The Jordan-Wigner– (JW) adjacent sites are linked by a solid line, while the non-JW-adjacent sites are linked by a dashed line. (c) A sketch of the circuits used to measure the Loschmidt amplitude using the Greenberger-Horne-Zeilinger– (GHZ) like state preparation. (d) The details of the structure of the time-evolution circuit using two Trotter steps.

TABLE I. The two-qubit gate cost for implementing the Trotterized time evolution and measurement of  $\text{Re}[\langle\psi|U_{\text{Trotter}}|\psi\rangle]$  for a  $x \times y$  rectangular FH lattice using the GHZ-state technique and the Hadamard test.

		Lattice		
		$x \times y$	$2 \times 8$	$5 \times 5$
<i>GHZ-state technique</i>				
Number of two-qubit gates	On-site interaction	$xy$	16	25
	Hopping interaction	$yx^2 + 7xy - 4(x + y)$	104	260
	GHZ preparation	$2(xy - 1)$	30	48
	Two Trotter steps	$2yx^2 + 17yx - 8(x + y) - 2$	254	593
Number of qubits		$2xy$	32	50
<i>Hadamard test</i>				
Number of two-qubit gates	On-site interaction	$7xy$	112	175
	Hopping interaction	$5yx^2 + 11xy - 8(y + x)$	256	820
	Two Trotter steps	$29xy + 10yx^2 - 16(y + x)$	624	1815
Number of qubits		$2xy + 1$	33	51

system size and is susceptible to error accumulation. Furthermore, a new interferometry technique employing a short imaginary time evolution has very recently been introduced [22] and could also be used for this algorithm.

### III. MODEL AND QUANTUM CIRCUIT IMPLEMENTATION

The Hamiltonian for the FH model is given by

$$H = H_{\text{hopp}} + H_{\text{int}}, \quad (6)$$

$$H_{\text{hopp}} = -J \sum_{\langle i,j \rangle, \sigma} \left( a_{i\sigma}^\dagger a_{j\sigma} + a_{j\sigma}^\dagger a_{i\sigma} \right), \quad (7)$$

$$H_{\text{int}} = U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (8)$$

where  $a_{i\sigma}$  ( $a_{i\sigma}^\dagger$ ) is a fermionic operator that destroys (creates) a particle at site  $i$  with spin  $\sigma$ ,  $n_{i\sigma} = a_{i\sigma}^\dagger a_{i\sigma}$  is the number operator, and  $\langle i, j \rangle$  denotes adjacent sites on a lattice. The term  $H_{\text{hopp}}$  is the hopping term of the Hamiltonian, which enables fermions to move to neighboring sites. The term  $H_{\text{int}}$  describes the on-site interactions between spin- $\uparrow$  and spin- $\downarrow$  fermions. The  $J$  and  $U$  are the parameters that control the magnitude of the hopping and interaction terms. Throughout this paper, we choose  $J = 0.5$  and  $U = 2$ .

To encode the fermionic operators on a quantum computer, we use the Jordan-Wigner (JW) transform, which maps each fermionic mode to one qubit such that the qubits are interpreted as lying along a one-dimensional (1D) line. The hopping term,  $H_{\text{hopp}}$ , is mapped to

$$J(a_{i\sigma}^\dagger a_{j\sigma} + a_{j\sigma}^\dagger a_{i\sigma}) \rightarrow \frac{J}{2} (X_{i\sigma} X_{j\sigma} + Y_{i\sigma} Y_{j\sigma}) \times Z_{i+1, \sigma} \dots Z_{j-1, \sigma}, \quad (9)$$

where  $X_{j\sigma}$ ,  $Y_{j\sigma}$ , and  $Z_{j\sigma}$  are the Pauli operators acting on the  $j$ th site of  $\sigma \in \{\uparrow, \downarrow\}$  spin sector. The interaction term

$H_{\text{int}}$  is mapped to

$$Un_{i\uparrow}n_{j\downarrow} = Ua_{i\uparrow}^\dagger a_{i\uparrow} a_{j\downarrow}^\dagger a_{j\downarrow} \rightarrow \frac{U}{4} (I_{i\uparrow} - Z_{i\uparrow}) \otimes (I_{i\downarrow} - Z_{i\downarrow}). \quad (10)$$

The sites that are adjacent in the JW ordering will be referred to as JW adjacent. All of the terms of the Hamiltonian between the JW-adjacent sites are two-qubit operators. The terms between non-JW-adjacent sites involve Pauli strings the lengths of which are proportional to the distance between sites in the JW ordering. For for a  $2 \times 8$  rectangular lattice, the interactions are illustrated in Fig. 1(b).

Various methods have been proposed in the literature to perform Hamiltonian simulation on a digital quantum computer, such as Trotter decomposition [33], randomly compiled Hamiltonian simulation [34,35], or classically optimized quantum simulation [36–38]. We use a first-order Trotter decomposition, which approximates  $\hat{U}(t) = e^{-iHt}$  by

$$\hat{U}_{\text{Trotter}}(t) = (e^{-iH_{\text{hopp}}\Delta t} e^{-iH_{\text{int}}\Delta t})^n, \quad (11)$$

where  $\Delta t = t/n$  and  $n$  is the number of steps. Note that for an initial product state, the effect of  $e^{-iH_{\text{int}}\Delta t}$  is simply to add a global phase to the circuit. For generic observables, one would expect the first-order Trotter decomposition to lead to an error  $\mathcal{O}(t^2/n)$ . However, for the Loschmidt amplitude of a Hamiltonian and initial states that are real in the same basis, the first-order decomposition turns out to be surprisingly efficient. As we show in Appendix M, we have

$$\langle\psi|\hat{U}_{\text{Trotter}}(t)|\psi\rangle = \langle\psi|e^{-i\hat{H}t}|\psi\rangle + \mathcal{O}(t^3/n^2), \quad (12)$$

i.e., the first-order decomposition scales just as well as the second-order one, gaining a factor  $t/n$  over the naive scaling.

We now focus on the Trotter-circuit implementation on the H2 quantum computer. The qubits in the H2 charge-coupled-device ion-trap quantum computer are effectively all-to-all connected, as any ion pair can be brought into interaction zones via shuttling [39,40]. The native two-qubit entangling gate on H2 is the ZZPhase gate, which implements an  $e^{-i\theta Z_i Z_j}$  operation between two qubits  $i$  and  $j$  with a tunable phase  $\theta$ . The interaction Hamiltonian  $H_{\text{int}}$  contains the  $Z_i^\uparrow Z_j^\downarrow$  two-body terms. Its time evolution can thus be directly realized using the  $N$  ZZPhase gate, where  $N$  is the number of sites. Note that if the initial state  $\psi$  is a classical bit-string state, then the effect of  $e^{H_{\text{int}}}$  in the first Trotter step can be implemented using single-qubit rotations.

The hopping dynamics between sites that are adjacent in the JW ordering are implemented using operators  $e^{i(XX+YY)\alpha}$ , where  $\alpha = -J\Delta t$ . This can be expressed as a product of XXPhase( $\alpha$ )  $\equiv e^{iXX\alpha}$  and YYPhase( $\alpha$ )  $\equiv e^{iYY\alpha}$ , both of which are equivalent to ZZPhase up to a conjugation by local unitaries. Thus, the cost of implementing hopping dynamics between two JW-adjacent sites is two two-qubit gates. The hopping between the non-JW-adjacent sites is more complex, since it involves operators with long Pauli strings,  $e^{i(XX+YY)Z\dots Z\alpha}$ . An elegant way to compile these operators into two-qubit gates is to use the Fermi-SWAP (FSWAP) networks [41,42]. The FSWAP gates, defined as CZ · SWAP, where “CZ” refers to the controlled-Z gate, swap the states of JW-adjacent fermions while preserving the antisymmetric exchange symmetry of the state vector. A sequence of FSWAP gates can be used to bring the distant fermions into JW-adjacent position. Once the sites are JW adjacent, the hopping dynamics can be implemented as usual using two two-qubit gates. The cost of implementing the FSWAP operation on H2 is only one two-qubit gate, since the SWAP can be implemented by simply relabeling the qubits and the CZ gate can be implemented using ZZPhase( $\pi/4$ ) and local rotations. One round of application of the FSWAP network changes the ordering of the qubits. Thus to restore the original order, the gate sequence in the next Trotter step is reversed as shown in Fig. 1(d). For a square  $L \times L$ , the gate overhead associated with FSWAP gates scales as  $L^3 \propto N^{\frac{3}{2}}$ , where  $N$  is the number of qubits (cf. Appendix F).

For a general  $x \times y$  rectangular lattice, the total number of gates for two Trotter steps is given in Table I. Measuring  $\text{Re}\langle\psi|U(t)|\psi\rangle$  using the GHZ-state technique adds only a linear-in-system-size overhead to the two-qubit gate count of  $n - 2$  gates, where  $n$  is the number of fermions in the state  $|\psi\rangle$ , which is significantly smaller than measuring the Loschmidt amplitude using the Hadamard test. For the  $2 \times 8$  and  $5 \times 5$  lattices, the GHZ technique results in an approximately factor-of-3 reduction in the two-qubit gate count. For details of the gate decomposition into the native gate set, see Appendix F.

#### IV. RESULTS ON THE QUANTUM DEVICE

In order to test the methods discussed above, we benchmark the measurement of the Loschmidt amplitude for the  $2 \times 8$  FH model on the 32-qubit H2 Quantinuum device. As the initial state, we choose the Néel state:  $|\psi_0\rangle = a_{1,\uparrow}^\dagger a_{2,\downarrow}^\dagger a_{3,\uparrow}^\dagger \dots |0\rangle$ , where the ordering of sites follows a “snake,” as illustrated in Fig. 1. The results are shown in Fig. 2. Since the Loschmidt amplitude is a global observable, a single error in any of the gates will cause a corruption of the output. Therefore, we would generically expect to measure a signal that is reduced from its ideal value by a factor

$$q = \prod_i F_i, \quad (13)$$

where  $F_i$  is the fidelity of gate  $i$  and the product runs over all gates in the circuit. For a more in-depth derivation, see Ref. [44] or Appendix H of the present work. In our case, assuming all gates to be of the same quality, we expect [44]

$$\text{Re } G^{\text{noisy}} = q \text{Re } G^{\text{noiseless}}, \quad (14)$$

where  $q$  is given by the two-qubit gate fidelity to the power of the number of two-qubit gates. The two-qubit gate of the H2-1 ion-trap quantum computer has been determined to be 99.8% through extensive cycle benchmarking (see Ref. [21]). Thus, the most straightforward error-mitigation scheme is to simply *rescale* the obtained results according to the above formula. Alternatively, we compare this method to *symmetry-filtered postprocessing*. As our Trotterized time evolution conserves the number of spin-up and spin-down fermions in the system, we can discard the shots where the bit strings do not correspond to the initial number of particles [15]. The data shown in Fig. 2 demonstrate that both techniques yield similar results.

While the rescaling by a factor equal to the inverse of the global fidelity works reasonably well, there are corrections to this simple rescaling. Surprisingly, the raw signal obtained for the first time point for  $p_\pi$  is greater than the clean value, which cannot be captured by the model given by Eq. (14). The reason for this counterintuitive result is that the GHZ state that is involved in time is highly nongeneric: incoherent  $Z$  errors cause a flip between the  $|(\phi, \psi, Et)\rangle$  and  $|-(\phi, \psi, Et)\rangle$  states (cf. Secs. G and H), which can *increase* the values of the measured probabilities  $p_0$  and  $p_\pi$ . Furthermore, we expect memory errors to play a larger role with increasing system size. These errors can be modeled as the coherent evolution  $e^{i\sum_i Z_i \theta_i}$ , where the  $\theta_i$  are angles that depend on the idling time of qubit  $i$ . In particular, a translationally invariant memory error maps  $\text{Re } G \rightarrow \cos(\sum \theta_i) \text{Re } G + \sin(\sum \theta_i) \text{Im } G$  and thus  $p_0$  or  $p_\pi$  can be larger than their noiseless values. Mitigating the effect of coherent errors on the measurement of the

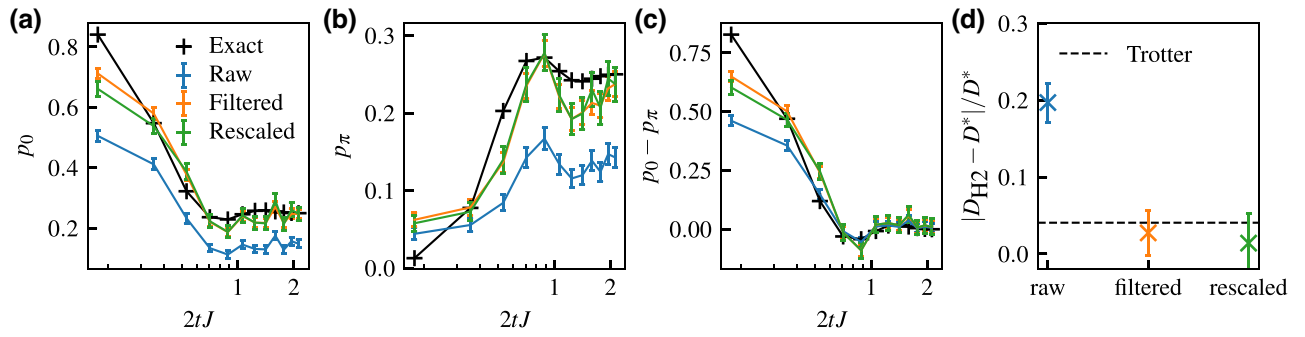


FIG. 2. The experimental data for the Loschmidt amplitude. For each time point, we prepare a GHZ state  $|0 \dots 0\rangle + e^{iEt}|\psi_0\rangle$ , where  $|\psi_0\rangle = |1010 \dots\rangle$  is the Néel state on a  $2 \times 8$  ladder (32 spin orbitals) and  $E = 1.0$ . Subsequently, the system is evolved under the FH model with  $J = 0.5$ ,  $U = 2$  using Trotterized time evolution. Finally, the inverse of the GHZ-state-preparation circuit (without the extra phase  $e^{iEt}$ ) is applied and the probability of the bit strings  $p(0 \dots 0) = p_0$  and  $p(10 \dots 0) = p_\pi$  is measured at the output. The number of Trotter steps is set to one for the first two time points and two for all other time points. (a)–(c) The results for (a)  $p_0 = \frac{1}{4} \|1 + e^{iEt} G_{\psi_0}(t)\|^2$ , (b)  $p_\pi = \frac{1}{4} \|1 - e^{iEt} G_{\psi_0}(t)\|^2$ , and (c)  $p_0 - p_\pi = \text{Re}(e^{iEt} G_{\psi_0}(t))$  compared with exact classically simulated circuits and with data obtained by applying the error-mitigation techniques described in Sec. IV. Note that the raw data for  $p_\pi$  are closer to the exact result than the mitigated data for the first time point. (d) A comparison of the filtered density of states obtained from the exact time evolution  $D^*$  with eight Trotter steps with that obtained from the experimental estimates  $D_{H2}$ . The filtered density of states is obtained by using  $D_{\delta,|\psi\rangle}(E) \approx \sum_{m=0}^R 2c_m \text{Re}\langle \psi | e^{-(H-E)tm} | \psi \rangle$  with  $c_m = (1/2^M) \binom{M}{M/2-m}$ ,  $\delta = 1.0$  leading to  $R = 12$  (see Appendix A). The dotted line represents the relative error of the filtered density of states due to Trotterization. The effect of increasing the final time and the sampling rates is negligible (see Appendix L).

Loschmidt amplitude is beyond the scope of the present work but it is clearly important to explore it in the future by, e.g., incorporating dynamical decoupling techniques [45,46].

We note that the rescaling error-mitigation method requires a good device characterization, as its results will be only as precise as the knowledge of the gate fidelity. In contrast, the symmetry-filtering method is device agnostic and does not depend on any calibration parameter. Both methods come at the cost of increasing the uncertainty by a factor of  $q^{-2}$ , which can in turn be offset by increasing the number of shots by the same factor. Interestingly, although based on different principles, the two error-mitigation methods produce similar results.

As we will now see, the estimate yielded by both error-mitigation techniques is sufficiently close to  $G^{\text{noiseless}}$  for the target parameters.

## V. CLASSICAL SIMULATION OF MONTE CARLO SAMPLING

In order to assess the effect of the noise on the final expectation value of the operator of interest in the microcanonical ensemble using classical simulations, we assume that all initial states behave similarly in the presence of noise. We simulate the Markov chains at different energies for the ladder geometry using MPS techniques, with bond dimension  $\chi = 100$ . While it is not expected that the precise behavior of the Loschmidt amplitude at the longest times is perfectly captured with this bond dimension, the filtered densities of states nevertheless converge quickly

with the bond dimension (see Appendix K), in line with the findings of Ref. [20,47]. Therefore, the bond dimension captures the correct features of the target (unnormalized) distribution of the sampling algorithm. To simulate the effect of shot noise of ideal quantum hardware, we add binomial noise to the time series. As can be seen in Fig. 2, some appreciable bias beyond shot noise remains after the various error-mitigation procedures. Therefore, to simulate the effect of systematic hardware errors that cannot be perfectly mitigated, we add random error terms to the time series. They are drawn from a Gaussian distribution centered at zero with standard deviation  $\sigma = 0.05$ . With this simple error model, we artificially introduce more error than observed in Fig. 2. In each case, we use the noisy time series to calculate the filtered density of states. The results are shown in Fig. 3.

A few comments are in order. First, the quantity represented in Fig. 3— $\sum_i \langle n_{i,\uparrow} n_{i,\downarrow} \rangle_\delta / N_{\text{sites}}$ , with  $\delta = 1$ —is the expectation value of the double occupancy per site in the *filter ensemble*. For finite systems, the filter ensemble can be thought of as a moving average of the microcanonical ensemble expectation value over an energy window of width  $\delta$ . By keeping  $\delta$  of the order  $O(1)$  and by increasing the system size, the filter ensemble eventually converges to the microcanonical ensemble for intensive quantities and for generic quantum systems that satisfy the ETH [19,47]. However, in Fig. 3 we already capture the tendency of the FH model to be insulating at high energy and conducting at low energies. Second, the finite-energy algorithm in the presence of noise displays a similar behavior to the finite-temperature scheme investigated in Ref. [48]. Namely, the

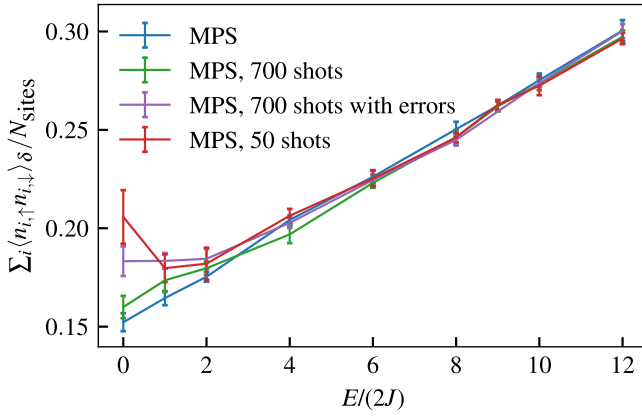


FIG. 3. The performance of the algorithm for different energies as a function of noise. We compare the results obtained for the average double occupancy per site  $\sum_i n_{i,\uparrow} n_{i,\downarrow} / N_{\text{sites}}$  from the Markov chain obtained by simulating the Loschmidt amplitudes using matrix-product-state (MPS) time evolution (blue) to results obtained by adding shot noise for 700 shots (green) and adding shot noise and Gaussian noise (red). For all data points considered, we have used 5000 Monte Carlo samples, except for  $E/(2J) = 0.0, 1.0$  and  $2.0$  in the presence of error, for which we have used 8800 samples. The error bars come from a blocking analysis of the Monte Carlo samples [43]. For all the simulations, the bond dimension used is  $\chi = 100$ .

expectation values of local observables are not sensitive to noise at high energies and/or temperatures. In other words, we expect the densities of states (or Boltzmann weights in the case of the canonical ensemble) to be relatively immune to noise as long as the overlap of the initial state with the targeted energy (temperature) region of the spectrum is large, and therefore we expect the Monte Carlo sampling to perform well in these cases. The algorithm starts to show deviations to the noiseless values close to  $E = 0.0$ . Note that  $E = 0.0$  is the lowest energy that can be targeted in a scalable manner with  $\hat{Z}$  product states. As explained in Ref. [19,47], it is not possible to explore lower energies using the  $\hat{Z}$ -product-state basis, since the low overlap of these initial states with the corresponding eigenstates will decay with the system size, yielding a vanishing density of states when approaching the thermodynamic limit. In contrast to the finite-temperature algorithm, we find that we need very few shots (as low as 50 for the energy considered) to converge toward the correct expectation value. This indicates that the finite-energy scheme is much more resilient to noise than the finite-temperature algorithm. This is explained by the fact that the Boltzmann weights  $W_\psi$  are not simply equal to the density of states as in the present work but are convolved by a factor  $e^{-\beta\omega}$ :  $W_\psi = \int dE e^{-\beta E} D_\psi(E)$ . Therefore, the low-energy sector is multiplied by an exponential factor and any error at low energy caused by noise will be amplified accordingly. Our results thus support the hypothesis that finite-energy

properties are more amenable to quantum techniques in the near term than finite-temperature ones, at least in the time-series framework proposed in Ref. [19]. However, we note that other initial states—with higher overlap with the low-energy sector—could be chosen, as demonstrated in Ref. [47]. It would be interesting to compare the performance of the two schemes when sampling from these initial states.

## VI. PROSPECTS OF QUANTUM ADVANTAGE

It has been demonstrated in Ref. [47] that for the finite-energy algorithm studied in our work, choosing the maximal time constant as a function of the system size is sufficient to reach the microcanonical expectation value in the thermodynamic limit, for generic quantum systems satisfying the ETH. Naively, one might conclude that quantum circuits of constant depth are thus sufficient. However, as the system size increases, so does the Trotter error. It has been proven in Ref. [49] that for a local Hamiltonian acting on  $N$  qubits, the number of Trotter steps required to reach time  $t$  with a fixed precision scales as  $N^{1/p} t^{1+1/p}$  in the worst case, where  $p$  is the order of the Trotter product formula. The number of entangling gates needed per second-order Trotter step using the JW encoding for a rectangular geometry of size  $x \times y$  is given by

$$n_{\text{JW}} = 2yx^2 + 3xy + 14x - 2y - 15, \quad (15)$$

while the number of entangling gates needed for a recently proposed local encoding [50] is given by

$$n_{\text{compact}} = 26xy - 24(x + y). \quad (16)$$

While the overall scaling of the algorithm is extremely favorable in a fault-tolerant setup, the strength of the signal decreases exponentially with the number of gates for a noisy intermediate-scale quantum (NISQ) device, as  $G_\psi(E)_{\text{measured}} \propto f^{N_{\text{gates}}} := q$ , where  $f$  is the gate fidelity. Therefore, the number of shots will increase exponentially with the system size, in order to reach the precision that would be obtained on a noiseless ideal device. Based on these considerations, we present the shot overhead as a function of the size of the system for the square lattice geometry in Fig. 4. We choose the units such that  $t = 1$ , yielding  $n_{\text{Trotter}} = \left\lceil \frac{2}{\sqrt{32}} \sqrt{N_{\text{qubits}}} \right\rceil$ , such that for 32 qubits we use two Trotter steps as in the present work and we use  $p = 2$  (second-order Trotter decomposition). Our resource estimate is likely pessimistic, as it would in principle be possible to take the final time as small as  $t \propto 1/\sqrt{N}$  [47]. However, choosing the maximum time constant ensures a faster convergence to the microcanonical value as a function of the system size.

Note that this resource estimate will likely be further improved by both hardware and software improvements. In particular, improved gate fidelity and optimization of

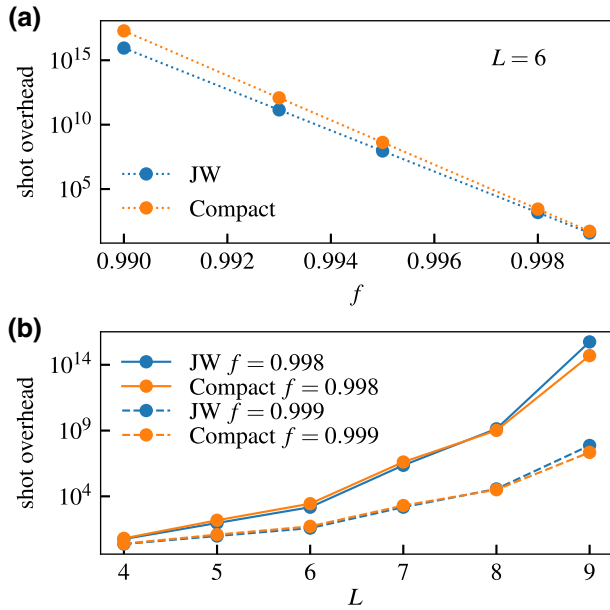


FIG. 4. A comparison of the estimated shot overhead between Fermion encodings: (a) shot overhead as function of fidelity  $f$  for system size  $L = 6$ ; (b) shot overhead as function of system size for different fidelities. We consider the cost of time evolving a square-lattice FH model of size  $L \times L$  with  $J = 0.5$  and  $U = 2$  up to  $t = 1$ , with a Trotter error comparable to that demonstrated in the experimental results of this work. The shot overhead  $1/q^2$  is the number of shots required to compensate the exponential damping of the signal by the global fidelity, given in terms of the fidelity to the power of the number of two-qubit gates,  $q$  (cf. Eq. (13)). We assume that single-qubit gates can be carried out with unit fidelity. We compare the JW encoding and the compact encoding for the two-qubit gate fidelities  $f = 0.998$  and  $f = 0.999$ . A second-order Trotterization is assumed for all time points, which, in order to reach the same precision as in the present experiment, requires two Trotter steps at  $L = 4$ , three steps at  $L = 5, 6$ , and four steps at  $L = 7, 8$ .

the time-evolution circuits [36–38,51] have the potential to reduce the resources requirement by several orders of magnitude.

The whole algorithm requires at least  $10^3$  Monte Carlo iterations for each energy density. For each iteration, we need about  $2\sqrt{N}$  time steps, measured at least  $10^2$  times each. Assuming 99.8% gate fidelity and an  $6 \times 6$  lattice, the shot overhead from error mitigation is about  $10^2$ . Therefore, each MC iteration requires about  $10^5$  shots. Assuming a shot time of around a second, the full algorithm would require a run time of the order of 1700 days. While this estimate is likely pessimistic, as it should be understood as an upper bound of the total run time given our current assumption, we conclude that the scheme presented in this work will require at least an order of magnitude of improvement to be practical on NISQ devices. The improvements will most likely come from

the hardware, as well as from more efficient Monte Carlo sampling and time-evolution algorithms.

Note that the resulting scaling of the algorithm presented in Fig. 4 is roughly exponential in the system size. This is expected, as most of the NISQ error-mitigation techniques, indispensable for obtaining sensible results, scale exponentially [52], as the increasing number of gates would result in an exponential increase of the noise.

If one is interested in the center of the spectrum only [47], one could instead evaluate the trace equation [Eq. (A5)] directly, e.g., by using the preparation of approximate unitary 2-designs [53]. This would circumvent the need for Monte Carlo sampling and bring down the costs but would not allow for the exploration of low energies properties of the spectrum, due to the exponentially vanishing filtered density of states.

While classical simulations of this algorithm can be performed efficiently in one dimension [47] (we exploit this fact in Sec. V), MPS simulations would require computational resources growing exponentially with one of the dimensions of the system. Nevertheless, we expect that the particular quantum routine demonstrated in this work could also be performed for larger system sizes by other state-of-the-art classical techniques. First, 2D tensor network classes could be used in principle, including projected entangled pair states (PEPSs). There, it is expected that computing the Loschmidt amplitude would be challenging, due to the complexity of the PEPS contraction [54–57]. Furthermore, neural-network simulations have proven competitive for performing the dynamics of 2D systems for short times [58–60] and it would be interesting to investigate whether they are able to capture the Loschmidt amplitude with the precision required by the algorithm used in this work. Our investigation of the effect of noise in Sec. V is encouraging, as it suggests that even approximate time series could yield satisfying observable expectation values.

Nonetheless, it is desirable to go beyond the program that we have applied in the present work. First, in order to reach smaller energies and/or temperatures, one needs to prepare an initial state with significant overlap with the low-energy sector, which would significantly increase both the classical and quantum resources needed [47]. Furthermore, while we have studied only static properties, a similar approach could give access to the finite-energy expectation values of dynamical observables, at the cost of deeper circuits. Due to the exponential resources needed to perform time evolution classically with most commonly used methods [61], it is likely that such a program would be out of reach for classical computers.

## VII. DISCUSSION AND OUTLOOK

In this work, we have demonstrated that the current capabilities of the Quantinuum H2 trapped-ion quantum



device allow for the execution of the quantum subroutine of one of the simplest time-series algorithms on a condensed-matter system. Although the noise of the machine still affects the results, the high fidelity of the gates as well as the low memory and SPAM error allow us to obtain satisfying data after error mitigation, in the sense that the physical features of the system should be well captured at the end of the hybrid quantum classical algorithm (see Fig. 3). We have compared two different and independent error-mitigation techniques, one based on symmetry and the other based on the probability of success of our circuit, and found that both give comparable results. Furthermore, using classical simulations, we have provided numerical evidence that the remaining errors, not correctly taken into account by our error-mitigation schemes, would have a low impact on the final prediction of the finite-energy properties of the system for the system size that is reachable with the current experimental setup (32 qubits). In other words, the particular Monte Carlo sampling explored here seems to be relatively resilient to noise. Furthermore, assuming that one is able to scale up the device while conserving the same level of fidelity, we have evaluated the resources necessary in order to run the algorithm for larger system sizes. To summarize the analysis: we have evaluated the number of shots necessary to recover the time series to the same accuracy as in the current experiment, given the dependency of the error-mitigation techniques on the number of gates and the noise level, and taking into account the known scaling of the algorithm. Assuming that this shot budget is allocated, the weight of the Monte Carlo sampling procedure would be obtained with the same precision and therefore the Monte Carlo chain should yield the same accuracy for the expectation values of local observables, which we have found adequate in the present case.

Overall, we have demonstrated that while time-series algorithms require the measurement of a global observable, which is in principle maximally sensitive to noise, the precision of an existing device today is sufficient to run these hybrid quantum-classical schemes. We have found that the effect of noise is not entirely explained by a global damping of the signal by the global fidelity. On the other hand, our resource estimates indicate that sampling over tens of thousands of initial states makes this algorithm prohibitively expensive to run on ion-trap devices before a significant drop in cost per sample caused, e.g., by the advent of a manufacturing age in which many quantum computers can execute coherent evolutions of intermediate depth in parallel.

It is as yet an open question to evaluate how precisely classical methods are able to capture the Loschmidt amplitude at moderately short times for 2D systems, although some theoretical studies have already been performed [62], which leaves open the possibility that the algorithm studied in this work could be carried out classically. However,

more involved versions of this algorithm would be necessary to explore the low-energy properties as well as the linear-response behavior of strongly correlated systems. These would likely be very challenging to execute classically, indicating the possibility of near-term useful quantum advantage with time-series algorithms.

The numerical data that support the findings of this study are available at Ref. [63]. The code used for numerical simulations is available from the corresponding author upon reasonable request.

## ACKNOWLEDGMENTS

K.H. and K.G. are supported by the German Federal Ministry of Education and Research (BMBF) through the “Efficient Quantum Algorithms for the Hubbard Model” (EQUAHUMO) project (Grant No. 13N16069) within the funding program “Quantum Technologies—from Basic Research to Market.” A.S. and E.C. acknowledge support from the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator.

## APPENDIX A: REVIEW OF THE ALGORITHM

The central quantity used in Ref. [19] is the filter operator:

$$\hat{P}_\delta(E) = \exp\left(-\frac{(\hat{H} - E)^2}{2\delta^2}\right), \quad (\text{A1})$$

where  $E$  is the target energy and  $\delta$  is the width of the filter. It can be shown that as long as  $\|H - E\|_\infty < N\pi/2$ , where  $\|\cdot\|_\infty$  denotes the operator norm,

$$\hat{P}_\delta(E) \approx \cos\left(\frac{\hat{H} - E}{N}\right) \Big|_{\lfloor \frac{N^2}{\delta^2} \rfloor_2}, \quad (\text{A2})$$

where  $N$  is the system size and  $\lfloor \cdot \rfloor_2$  denotes the nearest even integer. By writing the cosine as a sum of two complex exponentials, using the binomial formula, and truncating the resulting series, one finds [19]

$$\hat{P}_\delta(E) \approx \sum_{-R}^R c_m e^{-i(H-E)t_m}, \quad (\text{A3})$$

where  $c_m = (1/2^M) \binom{M}{M/2-m}$ ,  $R = \lfloor x/\delta \rfloor$ , in which  $x$  is a scalar controlling the truncation of the series and  $t_m = 2m/N$ . Furthermore, it has been shown that one can further reduce the number of measurements by choosing  $R = \lfloor x\alpha/\delta \rfloor$ ,  $t_m = 2m/\alpha$  and  $\alpha \propto \sqrt{N}$  if the initial states are product states. In order to relate the microcanonical

expectation value and the cosine-filter operator, Ref. [19] considers

$$\langle \hat{A} \rangle_\delta(E) = \frac{\text{tr} \hat{A} \hat{P}_\delta(E)}{\text{tr} \hat{P}_\delta(E)} \quad (\text{A4})$$

$$= \frac{\sum_i D_{\delta,|i\rangle}(E) A_{\delta,|i\rangle}(E)}{\sum_i D_{\delta,|i\rangle}(E)}, \quad (\text{A5})$$

where the sum over the index  $i$  denotes the sum over the product states in the  $\hat{Z}$  basis and

$$A_{\delta,|\psi\rangle} = \frac{\langle \psi | \hat{A} \hat{P}_\delta(E) + \hat{P}_\delta(E) \hat{A} | \psi \rangle}{2 \langle \psi | \hat{P}_\delta(E) | \psi \rangle}, \quad (\text{A6})$$

$$D_{\delta,|\psi\rangle}(E) = \langle \psi | P_\delta(E) | \psi \rangle. \quad (\text{A7})$$

$$\approx \sum_0^R 2c_m \text{Re}(\langle \psi | e^{-i(H-E)t_m} | \psi \rangle). \quad (\text{A8})$$

The sum in Eq. (A5) is sampled using classical Monte Carlo, with  $D_{\delta,|i\rangle}$  being the (unnormalized) target distribution of the sampling algorithm. Note that in this work, we choose the observable  $\hat{A}$  to be diagonal in the  $\hat{Z}$ -product-state basis; therefore,  $A_{\delta,|i\rangle}$  reduces to the eigenvalue of  $\hat{A}$  corresponding to the product state  $|i\rangle$ .

## APPENDIX B: SCALING OF THE ALGORITHM

The scaling of the algorithm with the system size has been investigated in Ref. [47]. For convenience, we summarize the authors' main arguments and results in this appendix.

The key finding concerns the scaling of the width of the filter as a function of the system size. In order to understand the requirements analytically, it is important to note that for systems that satisfy the ETH [30], the value of  $\sum_{E \in [E_0 - \epsilon, E_0 + \epsilon]} \langle E | \hat{A} | E \rangle$ , where  $|E\rangle$  denotes the eigenstate of the Hamiltonian of interest, converges to the micro-canonical ensemble in the thermodynamic limit if  $\hat{A}$  is extensive and if  $\epsilon/n \rightarrow 0$ . This is because for a system satisfying the ETH, the values of  $\langle E | \hat{A} | E \rangle$  are smooth functions of  $E$ . Therefore, in order to understand the convergence of the filter ensemble of Eq. (A5) to the micro-canonical ensemble, one needs to determine the width of the density of states in the filter ensemble as a function of the width  $\delta$ . This can be done by performing a convolution of the density of states of the Hamiltonian—which is a Gaussian with variance  $\sigma_0$  for most interacting model of interest [64,65]—and of the envelope of the filter. The result is a density of states proportional to a Gaussian with variance  $\Delta = \delta / \sqrt{1 + \delta^2 / N \sigma_0^2}$ . According to the above discussion, a choice of  $\delta$  as big as  $\delta \propto \sqrt{N}$  (or smaller) is sufficient to ensure that  $\Delta / N \rightarrow 0$ .

The choice of  $\alpha$  governing the number of measurements is related to the fact that one uses a cosine filter given in Eq. (A2) to approximate an Gaussian filter given in Eq. (A1). Because of the periodicity of the cosine, one needs to make sure that  $\langle \psi | (H - E) / \alpha | \psi \rangle < (\pi/2)$  for all states  $|\psi\rangle$  on which the filter operator is applied. If one samples over product states, it is enough to choose  $\alpha \propto \sqrt{N}$ , due to the limited support of the energy distribution of product states.

In the present work,  $\alpha = 2\sqrt{N}$ ,  $\delta = 1$  and  $x = 1$ .

## APPENDIX C: BEHAVIOR OF LOSCHMIDT AMPLITUDE AS A FUNCTION OF SYSTEM SIZE

In order to better understand the scaling of the algorithm, it is interesting to study the decay of the Loschmidt amplitude  $G_\psi = \langle \psi | e^{-iHt} | \psi \rangle$  as a function of the system size  $N$ . Note that the considerations presented in Appendix B have been derived from the knowledge of the density of states of the interacting system and since the Loschmidt amplitude can be understood as the Fourier transform of the density of states, the two pictures are equivalent. However, it might be more intuitive for the reader to look at the Loschmidt amplitude directly, since it is the quantity that is measured on the quantum computer in the algorithm used throughout this work.

We start by deriving an analytical expression for the Loschmidt echo for a particular class of initial states. We consider local Hamiltonians of the form

$$\hat{H} = \sum_i h_i^0 + h_{i,i+1}^{\text{int}}, \quad (\text{C1})$$

where  $h_i^0$  acts on the subsystem  $i$  and  $h_{i,i+1}^{\text{int}}$  is the interaction term between subsystem  $i$  and  $i + 1$ . It has been proven in Ref. [64] that for the product states  $|a\rangle = \prod_i |a_i\rangle$  such that  $h_i^0 |a_i\rangle = \epsilon_i |a_i\rangle$ , one has

$$\lim_{N \rightarrow \infty} \langle a | \delta(H - E) | a \rangle = \frac{\exp\left(-\frac{(E - E_a)^2}{2\sigma_a^2}\right)}{\sqrt{2\pi}\sigma_a}, \quad (\text{C2})$$

where  $E_a = \langle a | \hat{H} | a \rangle$  and  $\sigma_a^2 = \langle a | (H - E)^2 | a \rangle$ . Furthermore, one has

$$\langle a | \delta(H - E) | a \rangle = \frac{1}{2\pi} \int \langle a | e^{-iHt} | a \rangle e^{iEt} dt. \quad (\text{C3})$$

Taking the inverse Fourier transform gives

$$G_a(t) = \int \frac{\exp\left(-\frac{(E - E_a)^2}{2\sigma_a^2}\right)}{\sqrt{2\pi}\sigma_a} e^{-iEt} dE \quad (\text{C4})$$

$$= e^{-t^2 \sigma_a^2 / 2} e^{-iE_a t}. \quad (\text{C5})$$

Because the states  $|a\rangle$  have finite correlation length and we have supposed  $\hat{H}$  to be local, we have  $\sigma_a = \sigma_0 \sqrt{N}$ , where

$\sigma_0$  depends on the details of the Hamiltonian. Therefore, the Loschmidt amplitude is a modulated Gaussian with variance proportional to  $1/\sqrt{N}$ :

$$|G_a(t)| = e^{-t^2 \sigma_0^2 N/2}. \quad (\text{C6})$$

If we define  $t_\epsilon$  as the time such that  $\forall t > t_\epsilon, |G_a(t)| < \epsilon$ , which is the time at which one can truncate the series given in Eq. (A3), one obtains

$$t_\epsilon(N) \propto 1/\sqrt{N}. \quad (\text{C7})$$

Note that the scaling of the algorithm makes intuitive sense, as the number of measurements increases as  $\sqrt{N}$ ; therefore, one obtains a constant number of nonzero measurements as one increases the system size. In order to check that these results are representative of the behavior of the Loschmidt amplitude at relatively small system sizes, we perform a numerical study of the following 1D tilted-field Ising model:

$$\hat{H} = \sum_i \hat{Z}_i \hat{Z}_{i+1} + h_x \hat{X}_i + h_z \hat{Z}_i, \quad (\text{C8})$$

with parameters to ensure that the system is ergodic:  $h_x = 0.8090$  and  $h_z = 0.9045$  [66], at various chain lengths  $L$ . We simulate the time evolution of this model at short time using the time-evolving block decimation (TEBD) technique, using bond dimension  $\chi = 400$ , starting from  $\hat{Z}$  product states (for details, see the caption of Fig. 5). We measure the Loschmidt amplitude as a function of time. In Fig. 5(a), we plot one typical Loschmidt amplitude as a function of the system size. We fit the results with a Gaussian with variance proportional to  $1/\sqrt{L}$  and we find relatively good qualitative behavior for already small system sizes. Note that these numerics go beyond the analytical calculation above, since the  $\hat{Z}$  product states are not eigenstates of the noninteracting part of the Hamiltonian. In Fig. 5(b), we plot the statistics of the decay exponent  $\kappa$ , defined as  $t_\epsilon(L) = CL^\kappa$ . We find that in most cases,  $\kappa$  is close to  $-\frac{1}{2}$  as expected. We have obtained similar results with various choices of small  $\epsilon$ .

Note that Eqs. (C5) and (C6) are valid only in the case in which the initial product states are the eigenstates  $\{|a\rangle\}$  of  $\sum_i h_i^0$ . This means that for large enough systems, the algorithm is tractable analytically *only* for operators that are diagonal in the  $\{|a\rangle\}$  basis. While this opens the door to rigorous benchmarking of the algorithm in the future, when quantum computers will reach the realm of hundreds of qubits, this is not applicable to the vast majority of observables of interest. In particular, the dynamical observables of the form  $e^{i\hat{H}t} \hat{A} e^{-i\hat{H}t}$ , which are extremely relevant to experimental settings, are not amenable to analytical treatments.

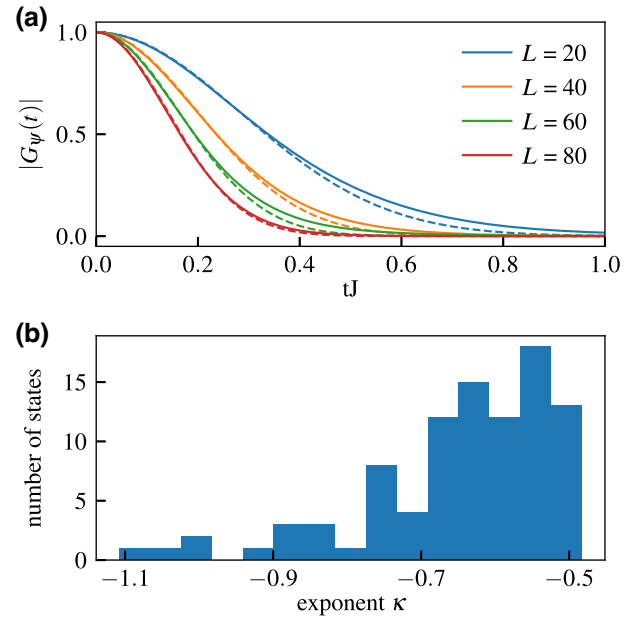


FIG. 5. The numerical study of the decay of the absolute value of the Loschmidt amplitude  $|G_\psi(t)|$ . (a) The typical decay of the Loschmidt amplitude as a function of the system size at fixed density, starting from a  $\hat{Z}$  product state, simulated with MPS:  $|G_\psi(t)|$  is plotted in full lines, while the dashed lines represent the fit with a Gaussian  $e^{-t^2 N \sigma_0^2 / 2}$ , where  $\sigma_0$  is obtained by fitting the largest system size. The product states  $|\psi\rangle$  are given by  $|\psi\rangle = |\downarrow\downarrow\downarrow\downarrow\uparrow\uparrow\uparrow\downarrow\downarrow\downarrow\uparrow\uparrow\rangle^{\otimes L/10}$  (b) The statistics of the decay of the absolute value of the Loschmidt echo. One starts with a pattern that is a product state of the form  $|i_1 \dots i_{10}\rangle^{\otimes (L/10)}$ , where  $i_j$  is either  $\uparrow$  or  $\downarrow$ , to ensure that we are at fixed energy density. From the curve  $|G_{\psi,L}(t)|$  at different system sizes, we obtain the time  $t_{0.1}(L)$  such that  $|G_{\psi,L}(t_{0.1}(L))| = 0.1$ . We then fit  $t_{0.1,L}$  with the function  $Ct^\kappa$ .

#### APPENDIX D: DETAILS OF THE MONTE CARLO SAMPLING

In order to extract expectation values of observables from the time-series algorithm outlined in Appendix A, we sample Eq. (A5) using the Metropolis-Hasting algorithm. As our goal is to probe the sector with a specific number of spin-up and spin down fermions, we use the following update scheme. We propose a new state  $|\psi'\rangle$  from the previous one  $|\psi\rangle$  by applying a random hopping of one fermion from one site to one of its unoccupied nearest-neighbor sites with the same spin. We then accept this new state with the acceptance ratio

$$A = \min\left(1, \frac{D_{\delta,|\psi'\rangle}(E) P_{\psi \rightarrow \psi'}}{D_{\delta,|\psi\rangle}(E) P_{\psi' \rightarrow \psi}}\right), \quad (\text{D1})$$

where  $P_{\psi' \rightarrow \psi}$  is the probability of hopping from  $\psi'$  to  $\psi$  and is related to the number of unoccupied neighboring sites. At the end of the sampling procedure, we obtain a list of product states. Since the observable that we study

in this work, the double occupancy, is diagonal in the product-state basis, the expectation value is estimated as the average double occupancy of the sampled product states.

### APPENDIX E: GHZ-LIKE STATE PREPARATION

Suppose that there exists a state  $|\phi\rangle$  such that  $\hat{U}|\phi\rangle = |\phi\rangle$ . Let us define  $|\pm(\phi, \psi, t)\rangle = \frac{1}{\sqrt{2}}(|\phi\rangle \pm e^{iEt}|\psi\rangle)$ . We further introduce

$$\begin{aligned} p_0(t) &= ||\langle +(\phi, \psi, t)|\hat{U}(t)| + (\phi, \psi, t)\rangle||^2 \\ &= \frac{1}{4} (1 + |G(t)|^2 + 2\text{Re}\{G(t)e^{iEt}\}) \end{aligned} \quad (\text{E1})$$

and

$$\begin{aligned} p_\pi(t) &= ||\langle +(\phi, \psi, t)|\hat{U}(t)| - (\phi, \psi, t)\rangle||^2 \\ &= \frac{1}{4} (1 + |G(t)|^2 - 2\text{Re}\{G(t)e^{iEt}\}), \end{aligned} \quad (\text{E2})$$

where  $G(t) = \langle \psi|e^{-iHt}|\psi\rangle$  as in the main text. In order to run the microcanonical algorithm, one only needs the real part of the Loschmidt amplitude modulated by a time-dependent phase, as made explicit in Eq. (A8). It is straightforward to see that

$$\text{Re}(\langle \psi|\hat{U}(t)|\psi\rangle e^{iEt}) = p_0(t) - p_\pi(t). \quad (\text{E3})$$

When  $|\psi\rangle$  is a product state, as it is the case in this paper, this procedure is very similar to a GHZ-state preparation. Note that both  $p_0$  and  $p_\pi$  can be obtained from measuring all qubits of only one circuit if  $|\psi\rangle$  is a product state. This can be understood by inspecting Fig. 6, which shows the circuit for the three-qubit GHZ-like state with the initial product state  $|\psi\rangle = |101\rangle$ .

Indeed,  $p_0 = \frac{1}{4} ||\langle (101) + \langle 000|\hat{U}(t)(|000\rangle + |101)\rangle||^2$  correspond to the probability of obtaining the bit string “000.”  $p_\pi = \frac{1}{4} ||\langle (000) - \langle 101|\hat{U}(t)(|000\rangle + |101)\rangle||^2$  correspond to the probability of measuring the bit string “000” after introducing a single-qubit  $X$  gate on the right of the leftmost Hadamard gate. Equivalently,  $p_\pi$  corresponds to the probability of measuring the bit string “100.” For this reason, we denote “000” (“100”) the 0 string ( $\pi$  string).

We note that as an alternative to applying the inverse of the GHZ-state preparation at the end of the circuit, one may directly measure  $\Pi_0 \otimes 1/M \sum_{k=1}^M (-1)^k (\cos(k\pi/M)X + \sin(k\pi/M)Y)^{\otimes M}$ , where  $M$  is the number of  $|1\rangle$ s in the initial state and  $\Pi_0$  is the projector on  $|0\rangle$  on the other qubits [67].

### APPENDIX F: IMPLEMENTING FH-MODEL TROTTER CIRCUITS

In this appendix, we give more details on the implementation of Trotterized dynamics on the H2 32-qubit quantum

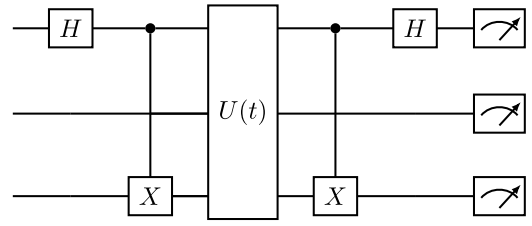


FIG. 6. A circuit diagram illustrating the GHZ-like state technique on three qubits, with an initial state  $|\psi_0\rangle = |101\rangle$ .

device. We show in Fig. 7 the important gates used in the Trotterized-dynamics circuit.

The native one-qubit gates on the H2 are rotations  $U_{1q}(\theta, \phi) \equiv e^{-\frac{1}{2}(\cos(\phi)\hat{X} + \sin(\phi)\hat{Y})}$  and  $R_z(\theta) \equiv e^{-\frac{1}{2}i\theta\hat{Z}}$  for  $\theta, \phi \in [0, 2\pi]$  and the native two-qubit gate is the ZZPhase-gate-implementing operation  $e^{-\frac{1}{2}i\alpha(\hat{Z}\otimes\hat{Z})}$  for  $\alpha \in [0, 2\pi]$ .

In the JW encoding, some interaction terms become strings of Pauli operators, the length of which is proportional to the size of the system. In the 2D FH model, these are the hopping terms between sites that are not adjacent in the JW ordering. Operators of the form  $e^{i\alpha(XX+YY)Z\dots Z}$  can be implemented by means of a staircase circuit using  $2(n-1)$  two-qubit gates [41,42]. The gate overhead associated with the implementation of long Pauli strings can be reduced by using FSWAP networks [68,69]. The FSWAP operator swaps the states of the neighboring qubits in the JW ordering while preserving the fermionic antisymmetric exchange statistics. The SWAP network is a sequence of FSWAP gates that brings the non-JW-adjacent sites into JW-adjacent positions, so that the hopping term between them can be implemented locally. This can be viewed as a succession of rotations into a basis in which the nonlocal Pauli string become two-body terms.

For a rectangular grid, an efficient way to implement an FSWAP network is described in Ref. [69]. The procedure consists of repeatedly applying the operator  $VW$ , where  $V$  swaps odd-numbered columns with those to their right and  $W$  swaps even-numbered columns with those to their right. After each application of  $VW$ , a new set of qubits that were previously not JW adjacent are made JW adjacent and the hopping term can be implemented locally via gate  $e^{-i(X\otimes X)\alpha} e^{-i(Y\otimes Y)\alpha}$ . After implementing all of the vertical hopping interactions, the FSWAP operations would normally be applied in reverse to return the qubits to their original positions. However, for Trotterized dynamics this is not necessary, since the order can be restored in the next Trotter step, by implementing the hopping-interaction gates in reverse order.

If the Trotter circuit involves an odd number of steps, then the final ordering needs to be restored by adding the FSWAP gates in reverse at the end of the circuit. However, for classical input states  $|\psi\rangle$  (tensor products of  $|0\rangle$  and

$|1\rangle$ ), the effect of the FSWAP network can be efficiently computed classically and thus observables of the form  $|\langle\psi|U(t)|\psi\rangle|^2$ , can still be obtained without applying the reversed FSWAP network on the last Trotter step.

On the H2 device, the simplest way to implement FSWAP is to use a single CZ gate and a software swap, i.e., virtually relabeling the qubits. The FSWAP operator can be expressed as a product of CZ and SWAP gates,  $\text{FSWAP} = \text{CZ} \cdot \text{SWAP}$ . Since H2 has all-to-all connectivity, the relabeling of the qubits does not add any overheads in implementation of subsequent gates. Thus, each FSWAP operation costs one two-qubit gate on H2.

The FSWAP network gate sequence for the ladder geometry and for a three-column geometry are illustrated in Figs. 8(a) and 8(b), respectively. The experiments on the H2 devices have been carried out on a  $2 \times 8$  ladder geometry. In general, for an  $x \times y$  rectangular lattice, the number of  $VW$  repetitions in the FSWAP network is  $\lfloor (x-1)/2 \rfloor$ . Each column swap operator involves  $y$  swaps and there are  $\lfloor x/2 \rfloor$  column swaps in operators  $V$  and  $W$ . Thus, the total number of FSWAP operations in one Trotter step is  $yx(x-1)$ . For a square 2D system, where  $x = y = L$ , the number of two-qubit gates in the FSWAP network is proportional to  $\propto L^3$  or  $\propto N^{3/2}$ , where  $N$  is the number of qubits. The superlinear scaling with the system size can be avoided by using the local fermion to qubits encoding [50,70–72] instead of JW encoding but for the small systems considered in this work, the JW encoding is more resource efficient.

## APPENDIX G: ERROR MITIGATION

Error mitigation is crucial for obtaining meaningful results on NISQ devices. We have considered two different error-mitigation strategies. The first strategy utilizes the number-conservation symmetry of the FH Trotterized evolution. It simply involves discarding the shots that violate the number symmetry and thereby reducing the error in the observed quantities. The second strategy involves rescaling the measured quantities to compensate for the effect of noise. Here, we detail the theoretical details of this heuristic.

The prepared GHZ-like states that undergo the Trotterized evolution are given by

$$|\psi_0\rangle = V_0|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\psi\rangle), \quad (\text{G1})$$

$$|\psi_\pi\rangle = V_\pi|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |\psi\rangle), \quad (\text{G2})$$

where  $|0\rangle \equiv |0\rangle^{\otimes n}$  is the vacuum state and  $|\psi\rangle$  is a product state. The operators  $V_0$  and  $V_j$  prepare the GHZ-like state from a product state. As explained in Appendix E,  $V_0$  can be constructed using a Hadamard gate on a selected qubit,

$j$ , and a series of controlled-NOT (CNOT) gates with a control on the  $j$  and targets on qubits where  $|\psi\rangle$  is in state  $|1\rangle$ . It is easy to show that  $V_\pi = X_j V_0$ .

The GHZ-like measurement technique obtains the real part of the Loschmidt amplitude through a difference of expectation values:

$$\text{Re}[G(t)] = |\langle 0|V_0^\dagger UV_0|0\rangle|^2 - |\langle 0|V_\pi^\dagger UV_0|0\rangle|^2 \quad (\text{G3})$$

$$\equiv |\langle 0|W_0|0\rangle|^2 - |\langle 0|W_\pi|0\rangle|^2 \quad (\text{G4})$$

$$= p_0 - p_\pi, \quad (\text{G5})$$

where  $W_0 \equiv V_0^\dagger UV_0$  and  $W_\pi \equiv V_\pi^\dagger UV_0 = X_j W_0$ .

Now let us consider the effect of gates affected by incoherent noise on  $\langle\psi_0|U|\psi_0\rangle$  and  $\langle\psi_0|U|\psi_\pi\rangle$ . First, note that a  $Z$  flip on a single qubit can flip the GHZ state  $|\psi_0\rangle \leftrightarrow |\psi_\pi\rangle$ :

$$Z_k \frac{1}{\sqrt{2}}(|0\rangle + |\psi\rangle) = \frac{1}{\sqrt{2}}(|0\rangle - |\psi\rangle), \quad (\text{G6})$$

if  $|\psi\rangle$  is in state  $|1\rangle$  in position  $k$ . Similarly,

$$Z_k \frac{1}{\sqrt{2}}(|0\rangle - |\psi\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + |\psi\rangle). \quad (\text{G7})$$

If the  $Z$  error occurs in the position of where  $|\psi\rangle$  is in state  $|0\rangle$ , it has no effect on the GHZ states. In a fully occupied FH lattice model,  $|\psi\rangle$  has an equal number of qubits in  $|0\rangle$  and  $|1\rangle$  states, and hence the probability of a single error  $Z$  causing the flip between GHZ states is equal to the probability of the state being unaffected. Other types of Pauli noise, i.e.,  $X$  and  $Y$  flip, will randomize the states  $|\psi_0\rangle$  and  $|\psi_\pi\rangle$ .

More formally, let  $\tilde{W}_0$  be a noisy channel corresponding to noisy circuit  $W_0$ . The noise is assumed to be Markovian and incoherent. The effect of the noisy channel on the input state  $\rho$  is

$$\tilde{W}_0[\rho] \approx (q + \gamma)W_0\rho W_0^\dagger + \gamma W_\pi\rho W_\pi^\dagger + (1 - q - 2\gamma)\frac{I}{2^{2N}}, \quad (\text{G8})$$

where  $q$  is the probability of an error that randomizes the state.

The noisy probabilities  $p_0^*$  and  $p_\pi^*$  are obtained by applying the  $\tilde{W}_0$  on  $\rho_0 = |0\rangle\langle 0|$  and projecting on the state  $\Pi_0 = \rho_0 = |0\rangle\langle 0|$  or  $X_j \Pi_0 X_j$ :

$$\begin{aligned} p_0^* &= \text{Tr}[\Pi_0 \tilde{W}_0[\rho_0]] \\ &= (q + \gamma)\text{Tr}[\Pi_0 W_0 \rho_0 W_0^\dagger] + \gamma \text{Tr}[\Pi_0 W_\pi \rho_0 W_\pi^\dagger] \\ &\quad + (1 - q - 2\gamma)\text{Tr}\left[\frac{\Pi_0}{2^{2N}}\right] \\ &= (q + \gamma)p_0 + \gamma p_\pi + \mathcal{K}, \end{aligned} \quad (\text{G9})$$

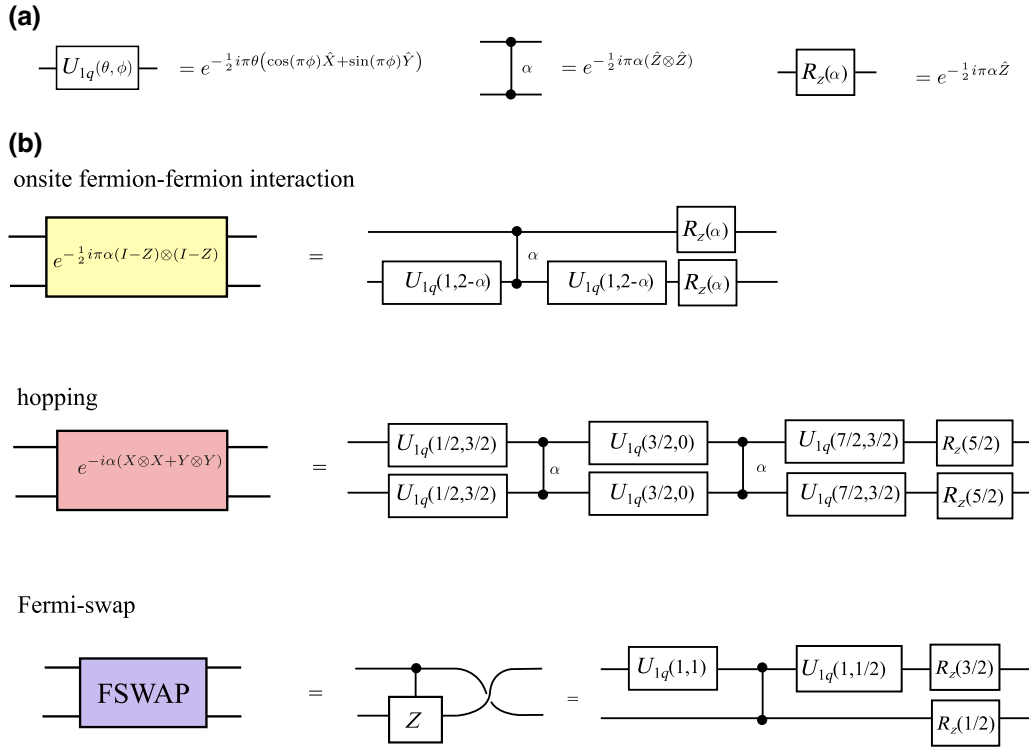


FIG. 7. (a) The H2 native gates. (b) The operations used in the FH Trotterized dynamics circuits—on-site interaction dynamics, hopping, and FSWAP—and their decomposition into one- and two-qubit gates native to the H2 device.

where  $\mathcal{K} = (1 - q - 2\gamma)/2^{2N}$  and

$$\begin{aligned}
 p_\pi^* &= \text{Tr}[X_j \Pi_0 X_j \tilde{W}_0[\rho_0]] = (q + \gamma) \text{Tr}[X_j \rho_0 X_j W_0 \rho_0 W_0^\dagger] \\
 &+ \gamma \text{Tr}[X_j \Pi_0 X_j W_\pi \rho_0 W_\pi^\dagger] + (1 - q - 2\gamma) \text{Tr} \\
 &\times \left[ \frac{X_j \Pi_0 X_j}{2^{2N}} \right] \\
 &= (q + \gamma)p_\pi + \gamma p_0 + \mathcal{K}.
 \end{aligned} \tag{G10}$$

From, Eqs. (G9) and (G10), it follows that

$$p_0^* - p_\pi^* = q(p_0 - p_\pi). \tag{G11}$$

Thus, from the measured noisy value of  $\text{Re}[G(t)]^* = p_0^* - p_\pi^*$ , one can obtain the true value using rescaling  $(p_0^* - p_\pi^*)/q$ . The factor  $q$  can be estimated from the microscopic model of the system. For example, in the presence of two-qubit depolarizing noise, where gate errors occur with probability  $p$ , the factor  $q$  is given by  $q = (1 - p)^n$ , where  $n$  is the number of two-qubit gates in the circuit. This is the approach used to mitigate errors in the experimental results presented in Fig. 2. The scaling parameters  $q$  and  $\gamma$  can be more accurately obtained by performing a zero-noisy extrapolation (ZNE) experiment.

ZNE involves systematically varying the amount of noise in the circuit and deducing the scaling parameters

from the change in the observed expectation values. One possible ZNE scheme for measuring  $q$  is to use “folding,” i.e., to apply  $\tilde{W}_0^\dagger \circ \tilde{W}_0$ , where  $\tilde{W}_0^\dagger$  denotes a noisy reversed circuit  $W^\dagger$ . The combined map  $\tilde{W}_0^\dagger \circ \tilde{W}_0$  is equal to identity in the absence of noise. Its effect on the input state  $\rho_0 = |0\rangle\langle 0|$  is

$$\begin{aligned}
 \tilde{W}_0^\dagger \circ \tilde{W}_0[\rho_0] &= (q + \gamma)^2 W_0^\dagger W_0 \rho_0 W_0^\dagger W_0 \\
 &+ (q + \gamma)\gamma W_0^\dagger W_\pi \rho_0 W_\pi^\dagger W_0 \\
 &+ \gamma(q + \gamma) W_\pi^\dagger W_0 \rho_0 W_0^\dagger W_\pi \\
 &+ \gamma^2 W_\pi^\dagger W_\pi \rho_0 W_\pi^\dagger W_\pi \\
 &+ (1 - q - 2\gamma) \frac{I}{2^{2(N-1)}}.
 \end{aligned} \tag{G12}$$

Using  $W_0 = V_0^\dagger U V_0$  and  $W_\pi = V_0^\dagger U V_\pi$ , we have

$$W_0^\dagger W_\pi = V_0^\dagger U^\dagger V_0 V_0^\dagger U V_\pi = V_0^\dagger V_\pi \tag{G13}$$

and

$$W_\pi^\dagger W_0 = V_\pi^\dagger U^\dagger V_0 V_0^\dagger U V_0 = V_\pi^\dagger V_0. \tag{G14}$$

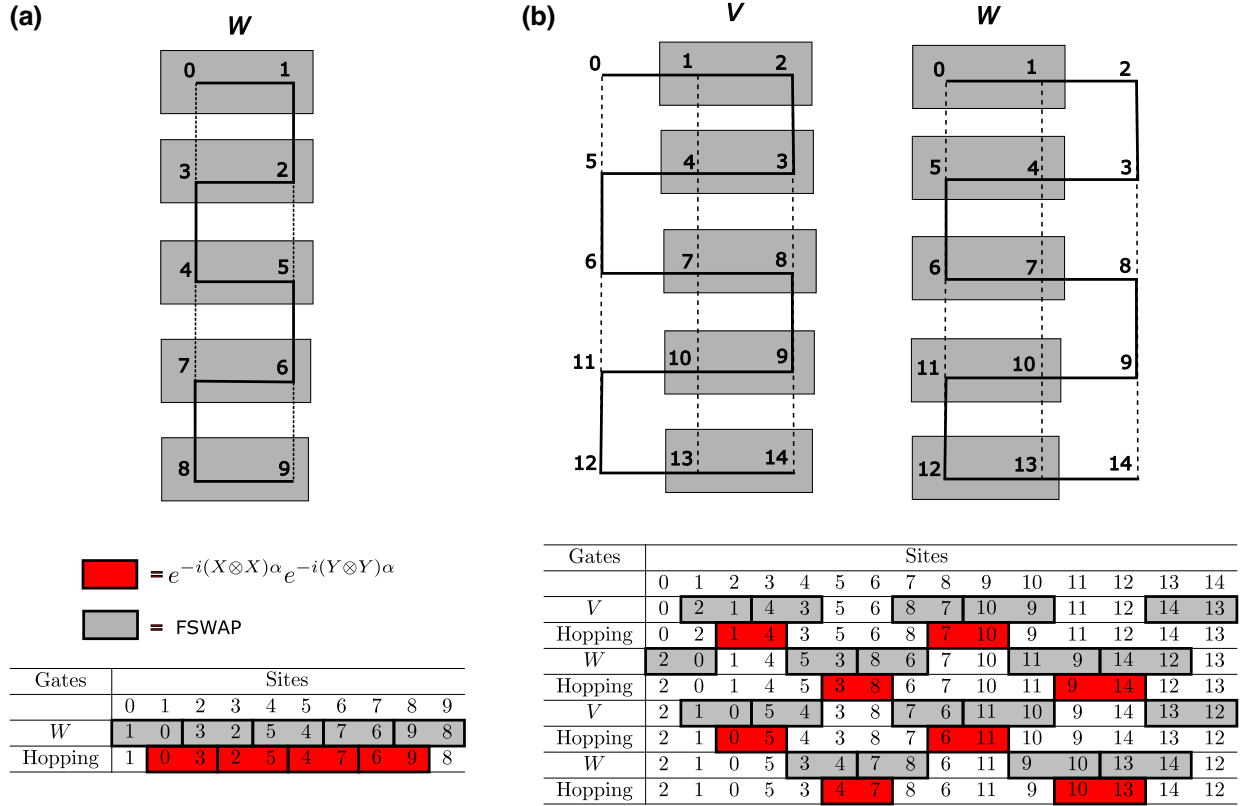


FIG. 8. The FSWAP network for an  $x \times y$  lattice with (a)  $x = 2$  and  $y = 5$  and (b)  $x = 3$  and  $y = 5$ . The network consists of applying the column swap operations  $V$  and  $W$ . The tables show the full sequence of operations without returning qubits to their original order.

Using Eqs. (G13) and (G14), Eq. (G12) simplifies to

$$\begin{aligned}
 \tilde{W}_0^\dagger \circ \tilde{W}_0[\rho_0] &= (q + \gamma)^2 \rho_0 + (q + \gamma) \gamma V_0^\dagger V_\pi \rho_0 V_\pi^\dagger V_0 \\
 &\quad + \gamma (q + \gamma) V_\pi^\dagger V_0 \rho_0 V_0^\dagger V_\pi \\
 &\quad + \gamma^2 \rho_0 + (1 - q - 2\gamma) \frac{I}{2^{2(N-1)}} \\
 &= (q + \gamma)^2 \rho_0 + 2\gamma (q + \gamma) X_j \rho_0 X_j \\
 &\quad + \gamma^2 \rho_0 + 2(1 - q - 2\gamma) \frac{I}{2^{2N}}. \quad (\text{G15})
 \end{aligned}$$

After applying  $\tilde{W}_0^\dagger \circ \tilde{W}_0$ , all qubits are measured in the computational basis. The probability of obtaining the outcome 0 on all qubits, i.e., projecting on the state  $\rho_0$ , is

$$\begin{aligned}
 p_0^{\text{ZNE}} &= \text{Tr}[\rho_0 \tilde{W}_0^\dagger [\tilde{W}_0[\rho_0]]] \\
 &= (q + \gamma)^2 \text{Tr}[\rho_0^2] + 2(q + \gamma) \gamma \text{Tr}[\rho_0 X_j \rho_0 X_j] \\
 &\quad + \gamma^2 \text{Tr}[\rho_0^2] + 2\mathcal{K} \\
 &= (q + \gamma)^2 + \gamma^2 + 2\mathcal{K}. \quad (\text{G16})
 \end{aligned}$$

Similarly, projecting on state  $X_j \rho_0 X_j$  gives

$$\begin{aligned}
 p_\pi^{\text{ZNE}} &= \text{Tr}[X_j \rho_0 X_j \tilde{W}_0^\dagger [\tilde{W}_0[\rho_0]]] \\
 &= (q + \gamma)^2 \text{Tr}[X_j \rho_0 X_j \rho_0] + 2\gamma (q + \gamma) \text{Tr} \\
 &\quad \times [X_j \rho_0 X_j X_j \rho_0 X_j] + \gamma^2 \text{Tr}[X_j \rho_0 X_j \rho_0] + 2\mathcal{K} \\
 &= 2\gamma (q + \gamma) + 2\mathcal{K}. \quad (\text{G17})
 \end{aligned}$$

Solving Eqs. (G16) and (G17) for  $q$  and  $\gamma$  gives

$$q = \sqrt{p_0^{\text{ZNE}} - p_\pi^{\text{ZNE}}}, \quad (\text{G18})$$

$$\gamma = -q + \sqrt{q^2 + (p_\pi^{\text{ZNE}} - 2\mathcal{K})}. \quad (\text{G19})$$

In practice, the ZNE approach involves performing “folding” experiments to evaluate  $q$  and  $\gamma$  using Eqs. (G18) and (G19), which are then used to obtain noiseless  $p_0$  and  $p_\pi$  from  $p_0^*$  and  $p_\pi^*$  using Eqs. (G9) and (G10).

We have tested this procedure numerically by performing noisy circuit simulations on a QISKIT Aer backend [73], where all two-qubit gates experience a uniformly depolarizing noise with probability of gate error  $\epsilon$ . The simulations have been carried out for a  $2 \times 2$  FH model with two spin-up and two spin-down fermions with  $|\psi\rangle = |1010\rangle \otimes$

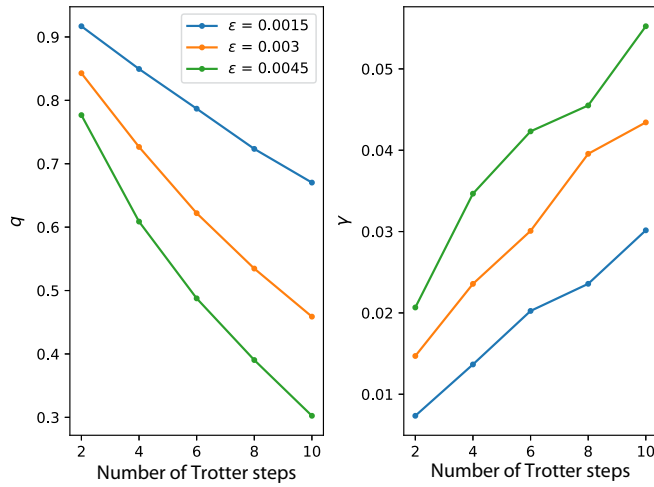


FIG. 9. Measuring the  $q$  and  $\gamma$  factors from numerical simulations of ZNE experiments. The two-qubit gate noise is modeled as a depolarizing channel with the probability of gate error set to  $\epsilon = 0.0015, 0.003, 0.0045$ . The simulations have been carried out for the  $2 \times 2$  FH model at half-filling, with  $U = 2.0, J = 0.5$ . The final time has been set to  $T = 1.6$  and the number of Trotter steps has been varied, thereby changing the depth of the circuit. Each data point has been obtained from 5000 shots.

|0101). The depth of the circuit is varied by changing the number of Trotter steps, while keeping the final time  $T$  constant. To apply a rescaling procedure, the factors  $q$  and  $\gamma$  have been obtained using the folding procedure. The resulting  $q$  and  $\gamma$  factors are shown in Fig. 9. As expected, the  $q$  factor decreases with the number of Trotter steps (and hence noisy gates), as well as the probability of gate error  $\epsilon$ . The factor  $\gamma$  is much smaller than  $q$ , since it arises only from the phase errors in the GHZ-state-preparation circuit. In Fig. 10, we present that probabilities  $p_0^*$  and  $p_\pi^*$  obtained in noisy simulations and the rescaled values using the measured  $q$  and  $\gamma$  factors. The rescaled values appear to be close to the real values for a range of circuit depths, which supports the validity of our model.

The rescaling procedure also works for biased Pauli noise. In Fig. 10, we show the rescaling error mitigation for simulations on H1E, which models realistic noise in the H1 device. On the H1E, the Pauli errors affecting the two-qubit gates are not equally likely—in particular,  $Z$  errors are more likely than  $X$  and  $Y$  errors. In addition, the H1E includes the coherent memory errors. Since the coherent errors are not included in the error-mitigation model, we expect that the memory errors will degrade the performance of the rescaling procedure. Indeed, from Fig. 10, we can see that the rescaling performs better when the coherent memory error is switched off. As the number of qubits and the depth of the circuit grows, the memory errors will become much more significant and will have to be accounted for in the error-mitigation procedure. Dynamical decoupling methods can be used to reduce the number

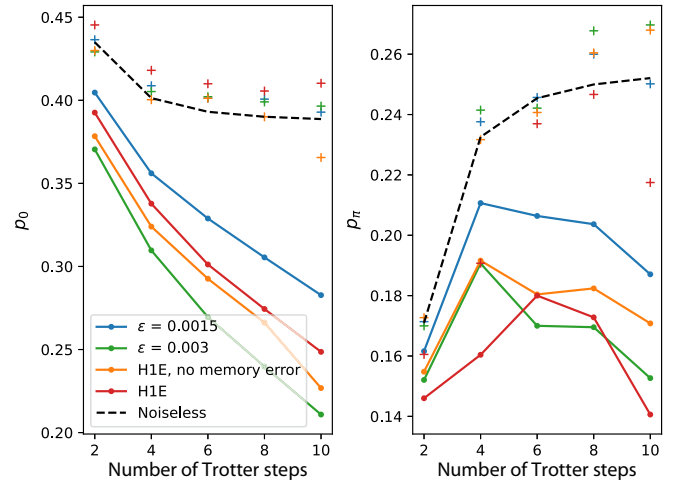


FIG. 10. Numerically testing rescaling error mitigation for  $p_0$  and  $p_\pi$  on the Aer simulator and the H1E emulator. The Aer simulator has been used with the depolarizing noise model with probability of two-qubit error  $\epsilon = 0.0015$  and  $0.003$ . The H1E emulator has been used with the default device-realistic error model and a model with the memory error switched off. The parameters  $q$  and  $\gamma$  have been extracted using simulated ZNE experiments. The simulations have been carried out for the  $2 \times 2$  FH model at half-filling with  $U = 2.0, J = 0.5$  and the final time  $T = 1.6$ . Each data point has been obtained from 5000 shots.

of accumulated memory errors. In addition, the rescaling model can be modified to include extra parameters associated with the build-up of memory error, as discussed in Appendix H.

## APPENDIX H: MEMORY ERROR

Let us define

$$\rho_{0/\pi}(t) = U(t)V_{0/\pi}\rho_0V_{0/\pi}U(t)^\dagger. \quad (\text{H1})$$

The effect of memory error alone can be thought of as

$$\rho_{0/\pi}(t)^* = \int d\theta p(\theta) |\Omega_\theta^{0/\pi}(t)\rangle \langle \Omega_\theta^{0/\pi}(t)|, \quad (\text{H2})$$

where  $\Omega_\theta^{0/\pi}(t) = \frac{1}{\sqrt{2}} (|\phi\rangle \pm e^{i\theta} e^{iEt} |\psi\rangle)$ . This leads to

$$p_{0/\pi} = \frac{q}{4} (1 + |G(t)|^2 \pm (2\langle \cos(\theta) \rangle \text{Re}G(t) - 2\langle \sin(\theta) \rangle \text{Im}G(t))). \quad (\text{H3})$$

If we call  $\mathcal{W}_0[\cdot]^*$  the noisy channel, which takes into account only memory error, and  $\tilde{\mathcal{W}}_0[\rho]$  the channel that takes into account both depolarizing noise and memory



error, we have

$$\begin{aligned} \tilde{W}_0[\rho] &\approx (q + \gamma)W_0^* \rho W_0^{\dagger} \\ &+ \gamma W_{\pi}^* \rho W_{\pi}^{\dagger} + (1 - q - 2\gamma) \frac{I}{2^{2N}}, \end{aligned} \quad (\text{H4})$$

according to Appendix G. Therefore we find that

$$\begin{aligned} p_{0/\pi} &= (1 - q - 2\gamma) \frac{I}{2^N} + \frac{q + 2\gamma}{4} (1 + |G(t)|^2) \\ &\pm \frac{q}{4} (2\langle \cos(\theta) \rangle \text{Re}G(t) - 2\langle \sin(\theta) \rangle \text{Im}G(t)). \end{aligned} \quad (\text{H5})$$

Note that in Appendix G, we have outlined how the effect of depolarizing noise could be mitigated through a simple zero-noise extrapolation procedure. In principle, similar schemes could be developed in the presence of memory error but since there are three parameters ( $\theta$ ,  $q$ , and  $\gamma$ ), the procedure would likely be more complicated. This difficulty could potentially be avoided using dynamical-decoupling techniques.

#### APPENDIX I: TROTTER ERROR ANALYSIS

The results of the time series as well as the error on the corresponding filtered density of states are presented in Fig. 11. All simulations have been performed using the WII method [74], with  $dt = 0.025$  and  $\chi = 300$ .

#### APPENDIX J: PURIFICATION RESULTS

In order to compare the minimum energy reachable with the product state with other studies performed in the canonical ensemble, we have performed a simulation of the system using the purification algorithm [75,76] using the TeNPy library [77] and measured the expectation value of the Hamiltonian as a function of the temperature. The

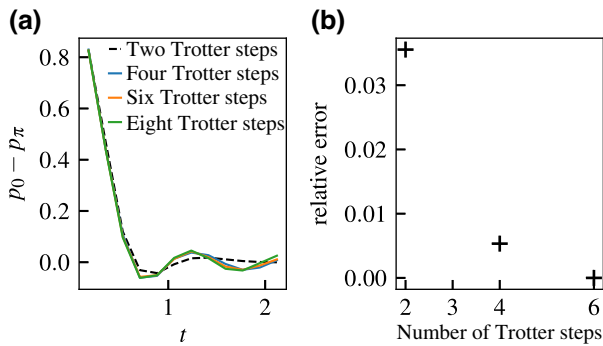


FIG. 11. (a) The real part of the Loschmidt amplitude  $\text{Re}(G(t)e^{iEt})$  with  $E = 1.0$  as a function of time for different numbers of Trotter steps. (b) The relative error on the final filtered density of states  $D$  as a function of the number of Trotter steps, with the reference value being with eight Trotter steps.

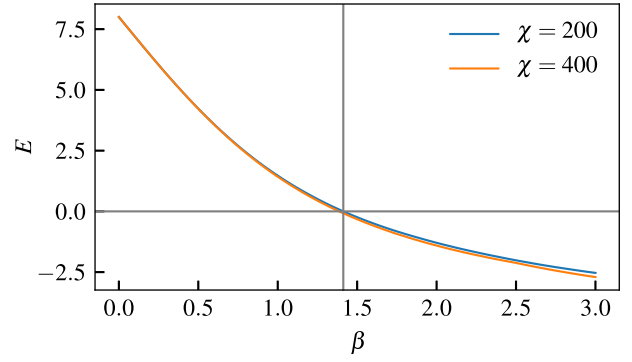


FIG. 12. The average value of the energy as a function of the inverse temperature obtained through purification.

results are presented in Fig. 12. The minimum energy reachable using  $Z$  product states is  $E = 0$ , which according to our simulation, corresponds to approximately  $\beta = 1.4$ . Note that lower temperatures can be reached by preparing initial states that have higher overlaps with the low-energy sector of the Hamiltonian.

#### APPENDIX K: CONVERGENCE OF THE MARKOV CHAIN WITH BOND DIMENSION

When checking the bond-dimension convergence of the MPS simulation, we have found that low bond dimensions are sufficient to capture the filtered density of states with high accuracy. We show the relative error (compared to  $\chi = 200$ ) on the filtered density of states for 60 different states in Fig. 13, at  $E = 1.0$ . For  $\chi = 100$ , which we use in the main text, the average relative error is below 1% with very few outliers around 2%.

#### APPENDIX L: INFLUENCE OF CUT-OFF TIME AND MEASUREMENT FREQUENCY ON FILTERED DENSITY OF STATES

In order to quantify the error on the truncation of the series of Eq. (A3), we have calculated the value of the filtered density of states with  $x = 0.5$ , for the Néel state and  $E = 1.0$ , with eight Trotter steps and  $\alpha = 2\sqrt{L}$ . We have found  $D = 0.2468$  with  $x = 1.0$  and  $D = 0.2447$  for  $x = 0.5$ , indicating a good convergence in time. Similarly, we have tested the dependence on measurement by doubling  $\alpha$ . With  $\alpha = 4\sqrt{L}$ , we have found an almost identical result,  $D = 0.2469$ , indicating a good convergence with the number of measurements.

#### APPENDIX M: SECOND-ORDER SCALING FOR FIRST-ORDER TROTTERIZATION OF LOSCHMIDT AMPLITUDES

It is known that the error of first-order Trotterization generally scales as  $\mathcal{O}(t^2/n)$ , while the error of second-order Trotterization scales as  $\mathcal{O}(t^3/n^2)$ , where  $t$  is the total

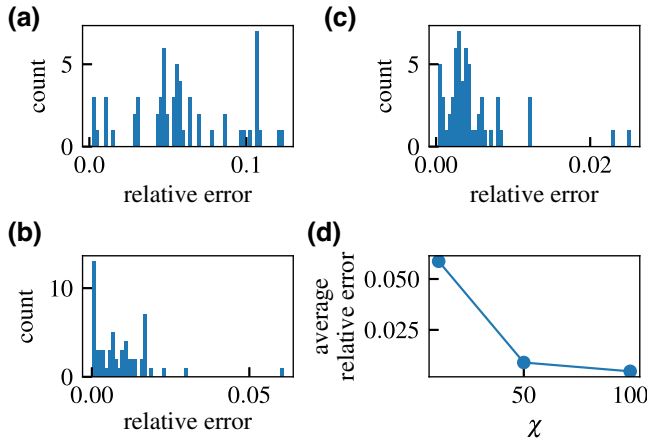


FIG. 13. The convergence of the filtered density of states for a set of states as a function of the bond dimension. The states are the first 60 samples of a Markov chain, obtained with  $\chi = 200$ . (a)–(c) Histograms of the relative error on the filtered density of states for (a)  $\chi = 10$ , (b)  $\chi = 50$ , and (c)  $\chi = 100$ . (d) The convergence of the average of the relative errors on the filtered density of states as a function of the bond dimension. The reference value in calculating the relative error is the one corresponding to  $\chi = 200$ .

time and  $n$  is the number of Trotter steps. In the following, we prove that for Hamiltonians the noncommuting terms of which are all *real* in some basis, then for all *real* wave functions in the same basis, first-order Trotterization of its Loschmidt amplitudes scales, rather, as  $\mathcal{O}(t^3/n^2)$ . Note that the notion of realness is basis dependent and that this proposition is true as long as there exists some basis in which all quantities are real.

Let the Hamiltonian be  $H = A + B$ , where  $A$  and  $B$  are noncommuting operators with commutator  $C := [A, B]$ . Using the Baker-Campbell-Hausdorff formula, we can write the first-order Trotterization as

$$U(t, n) := [e^{-iA\frac{t}{n}} e^{-iB\frac{t}{n}}]^n = e^{-it(A+B) - \frac{t^2}{n} \frac{C}{2} + \mathcal{O}\left(\frac{t^3}{n^2}\right)}. \quad (\text{M1})$$

The Taylor series of a function  $f(\delta) = e^{X+\delta Y}$  reads

$$f(\delta) = e^X + \delta \int_0^1 d\alpha e^{\alpha X} Y e^{(1-\alpha)X} + \mathcal{O}(\delta^2). \quad (\text{M2})$$

Using this formula to expand the right-hand side of Eq. (M1) as a function of small time steps  $\delta := \frac{t}{n}$ , with  $X = -itH$  and  $Y = -tC/2 + \mathcal{O}(t^2/n)$ , we obtain

$$U(t, n) = e^{-itH} - \frac{t^2}{n} \int_0^1 d\alpha e^{-itH\alpha} C e^{-itH(1-\alpha)} + \mathcal{O}(t^3/n^2). \quad (\text{M3})$$

The leading error term in the Loschmidt amplitude  $\langle \psi | e^{-itH} | \psi \rangle$  would then be  $(t^2/n)\epsilon$ , with

$$\epsilon := \int_0^1 d\alpha \langle \psi | e^{-itH\alpha} C e^{-itH(1-\alpha)} | \psi \rangle \quad (\text{M4})$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} d\alpha \langle \psi | e^{-itH(\alpha+\frac{1}{2})} C e^{-itH(\frac{1}{2}-\alpha)} | \psi \rangle \quad (\text{M5})$$

$$= \int_0^{\frac{1}{2}} d\alpha \langle \psi | e^{-itH(\alpha+\frac{1}{2})} C e^{-itH(\frac{1}{2}-\alpha)} + e^{-itH(\frac{1}{2}-\alpha)} C e^{-itH(\frac{1}{2}+\alpha)} | \psi \rangle \quad (\text{M6})$$

$$= \int_0^{\frac{1}{2}} d\alpha \langle \psi(t_1) | C | \psi(t_2) \rangle + \langle \psi(-t_2) | C | \psi(-t_1) \rangle \quad (\text{M7})$$

$$= \int_0^{\frac{1}{2}} d\alpha \langle \psi(t_1) | C | \psi(t_2) \rangle - \langle \psi(-t_1) | C | \psi(-t_2) \rangle^*, \quad (\text{M8})$$

where  $t_1 := -t(\alpha + \frac{1}{2})$  and  $t_2 := t(\frac{1}{2} - \alpha)$ , and we have used the fact that  $C$  is anti-Hermitian (because it is the commutator of two Hermitian operators). Let us assume that the operators  $A$  and  $B$  are real in some basis  $|a\rangle$ . Then, for any wave function that is real in that basis, we have the time-reversal symmetry  $\langle a | \psi(t) \rangle = \langle a | \psi(-t) \rangle^*$ . Moreover, the matrix elements  $\langle a' | C | a \rangle$  of the commutator are also real in the same basis. Therefore,

$$\begin{aligned} & \langle \psi(-t_1) | C | \psi(-t_2) \rangle^* \\ &= \left[ \sum_{a,a'} \langle \psi(-t_1) | a' \rangle \langle a' | C | a \rangle \langle a | \psi(-t_2) \rangle \right]^* \\ &= \langle \psi(t_1) | C | \psi(t_2) \rangle \end{aligned} \quad (\text{M9})$$

and the leading error term  $\epsilon$  vanishes. The next error term in Loschmidt amplitude is then proportional to  $\mathcal{O}(t^3/n^2)$ .

Extending this proof to Hamiltonians with more than two terms is straightforward by replacing the commutator  $C$  with the sum of all commutators.

- 
- [1] B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen, From quantum matter to high-temperature superconductivity in copper oxides, *Nature* **518**, 179 (2015).
  - [2] C. W. Bauer, *et al.*, Quantum simulation for high-energy physics, *PRX Quantum* **4**, 027001 (2023).
  - [3] J. Hubbard, Electron correlations in narrow energy bands, *Proc. R. Soc. Lond. A* **238** (1963).
  - [4] J. Hubbard, Electron correlations in narrow energy bands III. An improved solution, *Proc. R. Soc. Lond. A* **276**, 238 (1964).

- [5] A. Wietek, R. Rossi, F. Šimkovic, M. Klett, P. Hansmann, M. Ferrero, E. M. Stoudenmire, T. Schäfer, and A. Georges, Mott insulating states with competing orders in the triangular lattice Hubbard model, *Phys. Rev. X* **11**, 041013 (2021).
- [6] T. Schäfer, *et al.*, Tracking the footprints of spin fluctuations: A multimethod, multimessenger study of the two-dimensional Hubbard model, *Phys. Rev. X* **11**, 011058 (2021).
- [7] J. P. F. LeBlanc, *et al.*, (Simons Collaboration on the Many-Electron Problem), Solutions of the two-dimensional Hubbard model: Benchmarks and results from a wide range of numerical algorithms, *Phys. Rev. X* **5**, 041041 (2015).
- [8] T. Timusk and B. Statt, The pseudogap in high-temperature superconductors: An experimental survey, *Rep. Prog. Phys.* **62**, 61 (1999).
- [9] M. R. Norman, D. Pines, and C. Kallin, The pseudogap: Friend or foe of high  $T_c$ ?, *Adv. Phys.* **54**, 715 (2005).
- [10] P. A. Lee, N. Nagaosa, and X.-G. Wen, Doping a Mott insulator: Physics of high-temperature superconductivity, *Rev. Mod. Phys.* **78**, 17 (2006).
- [11] A. Bohrdt, L. Homeier, C. Reinmoser, E. Demler, and F. Grusdt, Exploration of doped quantum magnets with ultracold atoms, *Ann. Phys.* **435**, 168651 (2021).. Special issue on Philip W. Anderson
- [12] A. Mazurenko, C. S. Chiu, G. Ji, M. F. Parsons, M. Kanász-Nagy, R. Schmidt, F. Grusdt, E. Demler, D. Greif, and M. Greiner, A cold-atom Fermi-Hubbard antiferromagnet, *Nature* **545**, 462 (2017).
- [13] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samajdar, H. Pichler, W. W. Ho, S. Choi, S. Sachdev, M. Greiner, V. Vuletić, and M. D. Lukin, Quantum phases of matter on a 256-atom programmable quantum simulator, *Nature* **595**, 227 (2021).
- [14] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, *Phys. Rep.* **986**, 1 (2022)..
- [15] S. Stanisic, J. L. Bosse, F. M. Gabbetta, R. A. Santos, W. Mruczkiewicz, T. E. O'Brien, E. Ostby, and A. Montanaro, Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer, *Nat. Commun.* **13**, 5743 (2022).
- [16] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, *Quantum* **6**, 759 (2022).
- [17] F. Turro, Quantum imaginary time propagation algorithm for preparing thermal states, [arXiv:2306.16580](https://arxiv.org/abs/2306.16580).
- [18] A. Summer, C. Chiaracane, M. T. Mitchison, and J. Goold, Calculating the many-body density of states on a digital quantum computer, *Phys. Rev. Res.* **6**, 013106 (2024).
- [19] S. Lu, M. C. Bañuls, and J. I. Cirac, Algorithms for quantum simulation at finite energies, *PRX Quantum* **2**, 020321 (2021).
- [20] A. Schuckert, A. Bohrdt, E. Crane, and M. Knap, Probing finite-temperature observables in quantum simulators of spin systems with short-time dynamics, *Phys. Rev. B* **107**, L140410 (2023).
- [21] S. A. Moses, *et al.*, A race track trapped-ion quantum processor, *Phys. Rev. X* **13**, 041052 (2024).
- [22] Y. Yang, A. Christianen, M. C. Bañuls, D. S. Wild, and J. I. Cirac, Phase-sensitive quantum measurement without controlled operations, *Phys. Rev. Lett.* **132**, 220601 (2024).
- [23] T. E. O'Brien, B. Tarasinski, and B. M. Terhal, Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments, *New J. Phys.* **21**, 023022 (2019).
- [24] C. Yi, C. Zhou, and J. Takahashi, Quantum phase estimation by compressed sensing, [arXiv:2306.07008](https://arxiv.org/abs/2306.07008).
- [25] L. Lin and Y. Tong, Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers, *PRX Quantum* **3**, 010318 (2022).
- [26] Z. Ding and L. Lin, Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation, *PRX Quantum* **4**, 020331 (2023).
- [27] R. D. Somma, Quantum eigenvalue estimation via time series analysis, *New J. Phys.* **21**, 123025 (2019).
- [28] K. Yamamoto, S. Duffield, Y. Kikuchi, and D. M. Ramo, Demonstrating Bayesian quantum phase estimation with quantum error detection, *Phys. Rev. Res.* **6**, 013221 (2024).
- [29] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
- [30] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
- [31] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, *Nature* **452**, 854 (2008).
- [32] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Direct estimations of linear and nonlinear functionals of a quantum state, *Phys. Rev. Lett.* **88**, 217901 (2002).
- [33] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, *Phys. Lett. A* **146**, 319 (1990).
- [34] A. M. Childs, A. Ostrander, and Y. Su, Faster quantum simulation by randomization, *Quantum* **3**, 182 (2019).
- [35] E. Campbell, Random compiler for fast Hamiltonian simulation, *Phys. Rev. Lett.* **123**, 070503 (2019).
- [36] C. Mc Keever and M. Lubasch, Classically optimized Hamiltonian simulation, *Phys. Rev. Res.* **5**, 023146 (2023).
- [37] 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).
- [38] R. Mansuroglu, T. Eckstein, L. Nützel, S. A. Wilkinson, and M. J. Hartmann, Variational Hamiltonian simulation for translational invariant systems via classical pre-processing, *Quantum Sci. Technol.* **8**, 025006 (2023).
- [39] D. Kielpinski, C. Monroe, and D. J. Wineland, Architecture for a large-scale ion-trap quantum computer, *Nature* **417**, 709 (2002).
- [40] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, and B. Neyenhuis, Demonstration of the trapped-ion quantum CCD computer architecture, *Nature* **592**, 209 (2021).
- [41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

- [42] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Simulation of electronic structure Hamiltonians using quantum computers, *Mol. Phys.* **109**, 735 (2011).
- [43] H. Flyvbjerg and H. G. Petersen, Error estimates on averages of correlated data, *J. Chem. Phys.* **91**, 461 (1989).
- [44] Y. Yang, A. Christianen, S. Coll-Vinent, V. Smelyanskiy, M. C. Bañuls, T. E. O'Brien, D. S. Wild, and J. I. Cirac, Simulating prethermalization using near-term quantum computers, *PRX Quantum* **4**, 030320 (2023).
- [45] L. Viola, E. Knill, and S. Lloyd, Dynamical decoupling of open quantum systems, *Phys. Rev. Lett.* **82**, 2417 (1999).
- [46] K. N. Smith, G. Subramanian Ravi, P. Murali, J. M. Baker, N. Earnest, A. Javadi-Abhari, and F. T. Chong, Error mitigation in quantum computers through instruction scheduling, [arXiv:2105.01760](https://arxiv.org/abs/2105.01760).
- [47] Y. Yang, J. I. Cirac, and M. C. Bañuls, Classical algorithms for many-body quantum systems at finite energies, *Phys. Rev. B* **106**, 024307 (2022).
- [48] K. Ghanem, A. Schuckert, and H. Dreyer, Robust extraction of thermal observables from state sampling and real-time dynamics on quantum computers, *Quantum* **7**, 1163 (2023).
- [49] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, *Phys. Rev. X* **11**, 011020 (2021).
- [50] C. Derby, J. Klassen, J. Bausch, and T. Cubitt, Compact fermion to qubit mappings, *Phys. Rev. B* **104**, 035118 (2021).
- [51] N. Astrakhantsev, S.-H. Lin, F. Pollmann, and A. Smith, Time evolution of uniform sequential circuits, *Phys. Rev. Res.* **5**, 033187 (2023).
- [52] S. N. Filippov, S. Miscalco, and G. García-Pérez, Scalability of quantum error mitigation techniques: From utility to advantage, [arXiv:2403.13542](https://arxiv.org/abs/2403.13542).
- [53] K. Seki and S. Yunoki, Energy-filtered random-phase states as microcanonical thermal pure quantum states, *Phys. Rev. B* **106**, 155111 (2022).
- [54] N. Schuch, M. M. Wolf, F. Verstraete, and I. J. Cirac, Computational complexity of projected entangled pair states, *Phys. Rev. Lett.* **98**, 140506 (2007).
- [55] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Computational complexity of projected entangled pair states, *Phys. Rev. Lett.* **98**, 140506 (2007).
- [56] S. Gonzalez-Garcia, S. Sang, T. H. Hsieh, S. Boixo, G. Vidal, A. C. Potter, and R. Vasseur, Random insights into the complexity of two-dimensional tensor network calculations, *Phys. Rev. B* **109**, 235102 (2024).
- [57] R. Vasseur, A. C. Potter, Y.-Z. You, and A. W. W. Ludwig, Entanglement transitions from holographic random tensor networks, *Phys. Rev. B* **100**, 134203 (2019).
- [58] M. Schmitt and M. Heyl, Quantum many-body dynamics in two dimensions with artificial neural networks, *Phys. Rev. Lett.* **125**, 100503 (2020).
- [59] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [60] I. L. Gutiérrez and C. B. Mendl, Real time evolution with neural-network quantum states, *Quantum* **6**, 627 (2022).
- [61] S.-H. Lin and F. Pollmann, Scaling of neural-network quantum states for time evolution, *Phys. Status Solidi (b)* **259**, 2100172 (2022).
- [62] D. S. Wild and A. M. Alhambra, Classical simulation of short-time quantum dynamics, *PRX Quantum* **4**, 020340 (2023).
- [63] <https://doi.org/10.5281/zenodo.8330634>
- [64] M. Hartmann, G. Mahler, and O. Hess, Spectral densities and partition functions of modular quantum systems as derived from a central limit theorem, *J. Stat. Phys.* **119**, 1139 (2005).
- [65] All traceless local models will have this property.
- [66] H. Kim and D. A. Huse, Ballistic spreading of entanglement in a diffusive nonintegrable system, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [67] O. Gühne, C.-Y. Lu, W.-B. Gao, and J.-W. Pan, Toolbox for entanglement detection and fidelity estimation, *Phys. Rev. A* **76**, 030305 (2007).
- [68] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum simulation of electronic structure with linear depth and connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).
- [69] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Strategies for solving the Fermi-Hubbard model on near-term quantum computers, *Phys. Rev. B* **102**, 235122 (2020).
- [70] F. Verstraete and J. I. Cirac, Mapping local Hamiltonians of fermions to local Hamiltonians of spins, *J. Stat. Mech.: Theory Exp.* **2005**, P09012 (2005).
- [71] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, *Ann. Phys.* **298**, 210 (2002).
- [72] K. Setia, S. Bravyi, A. Mezzacapo, and J. D. Whitfield, Superfast encodings for fermionic quantum simulation, *Phys. Rev. Res.* **1**, 033033 (2019).
- [73] Qiskit contributors, QISKIT: An open-source framework for quantum computing (2023).
- [74] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Time-evolving a matrix product state with long-ranged interactions, *Phys. Rev. B* **91**, 165112 (2015).
- [75] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [76] T. Barthel, U. Schollwöck, and S. R. White, Spectral functions in one-dimensional quantum systems at finite temperature using the density matrix renormalization group, *Phys. Rev. B* **79**, 245101 (2009).
- [77] J. Hauschild and F. Pollmann, Efficient numerical simulations with tensor networks: Tensor network PYTHON (TeNPy), *SciPost Phys. Lect. Notes* **5** (2018).