Parallel-in-time quantum simulation via Page and Wootters quantum time

N. L. Diaz,[1,2,3,*] Paolo Braccia,[4,5] Martin Larocca,[4,6] J. M. Matera,[2] R. Rossignoli,[2,7] and M. Cerezo[1]

[1]*Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

[2]*Departamento de Física-IFLP/CONICET, Universidad Nacional de La Plata, C.C. 67, La Plata 1900, Argentina*

[3]*Center for Non-Linear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

[4]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

[5]*Dipartimento di Fisica, Università di Firenze and INFN Sezione di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Italy*

[6]*Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

[7]*Comisión de Investigaciones Científicas (CIC), La Plata 1900, Argentina*

(Received 30 December 2023; accepted 29 August 2025; published 29 September 2025)

In the past few decades, researchers have created a veritable zoo of quantum algorithms by drawing inspiration from classical computing, information theory, and even from physical phenomena. Here, we present quantum algorithms for parallel-in-time simulations that are inspired by the Page and Wootters formalism. In this framework, and thus in our algorithms, the classical time variable of quantum mechanics is promoted to the quantum realm by introducing a Hilbert space of "clock" qubits that are then entangled with the "system" qubits. We show that our algorithms can compute temporal properties over N different times of many-body systems by only using $\log(N)$ clock qubits. As such, we achieve an exponential trade-off between time and spatial complexities. In addition, we rigorously prove that the entanglement created between the system qubits and the clock qubits has operational meaning, as it encodes valuable information about the system's dynamics. We also provide a circuit depth estimation of all the protocols, showing a running time advantage in computation times over traditional sequential-in-time algorithms. In particular, for the case when the dynamics are determined by the Aubry-Andre model, we present a hybrid method for which our algorithms have a depth that only scales as $\mathcal{O}(\log(N)n)$. As a by-product, we can relate the previous schemes to the problem of equilibration of an isolated quantum system, thus indicating that our framework enables a new dimension for studying dynamical properties of many-body systems.

DOI: 10.1103/wpnf-4nnn

I. INTRODUCTION

The field of quantum foundations studies the fundamental principles of quantum theory, such as the nature of quantum states, the interpretation of measurements, the equilibration and thermalization of isolated systems, and the emergence of classicality [1–5]. Another important question that has recently attracted wide attention within this field is that of the role of time in quantum mechanics (QM) [6–34]: It is clear that ever since its inception, time in QM has been treated as an external classical parameter, in asymmetry with other quantum observables. For instance, in the canonical quantization procedure, one promotes the position and momentum variables to operators and the Poisson bracket to a commutator [35]. This quantization is implemented at a fixed time value so the variable t appearing in Schrödinger equation is the same as that appearing in the classical equations of motion. While seemingly innocuous, it is believed that the imbalance between time and space could be a critical issue in developing a quantum theory of gravity [22,36–38]. At the same time, such asymmetry inevitably limits the range of applicability of quantum information and computation tools, as asking questions like "what is the entanglement between the space and time coordinates?" is an entirely moot point within the conventional quantum mechanical framework.

It is tempting to fix the space-time asymmetry by promoting $i\hbar\frac{d}{dt}$ to a quantum operator conjugate to some quantum time observable T, such that $[T, H] = i\hbar$. However, the previous approach has the critical issue that it forces T and H to have exactly the same eigenspectrum, which is generally incompatible (such argument is often attributed to Pauli [39]). Despite this apparent difficulty and other subtleties, there are several proposals to treat time on equal footing with other physical quantities [13,14,40–44]. Here, we will focus on the so-called Page and Wootters (PaW) mechanism [40]. In this framework, the universe is composed by a quantum system of interest plus an ancillary *clock* quantum system, such that the joint state of the universe, the *history state*, is in a stationary state. The previous "Pauli's objection" is circumvented since the operator T acts on the clock system implying $[T, H] = 0$. Remarkably, as long as the system and the clock are correlated in a specific way, the unitary evolution of the system

*Contact author: nldiaz.unlp@gmail.com

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

can be restored by conditioning over the clock states. In this way, measures of system-time entanglement become rigorous quantifiers of the amount of distinguishable evolution undergone by the system during its history [8].

The foundational discussion surrounding the role of time has many intriguing ramifications, even when focusing solely on the PaW mechanism. The interested reader can refer to Appendix A for pertinent discussions. However, the primary objective of this work is computational: In this manuscript, we provide a translation of the PaW mechanism into a useful quantum computational scheme where the quantum aspects of time are captured by clock qubits. This allows us to develop quantum algorithms for studying temporal averages of several dynamical properties of a quantum system. Specifically, given an n-qubit quantum system that is evolving under the action of a time-independent Hamiltonian H, we consider the problem of approximating the infinite-time average of some time-dependent dynamical quantity by a discrete sum over N different times. In a standard setting, we can estimate said discrete average by sequentially running N different quantum circuits (one for each time in the average). However, by leveraging the history state of the Page and Wootters formalism we propose a quantum algorithm for parallel-in-time simulations that uses $\log(N)$ (any logarithm in this manuscript is taken in base 2) ancillary clock qubits and that allows us to evaluate the temporal average with a single quantum circuit. The previous shows that using the history state leads to an exponential trade-off between temporal complexity (running multiple circuits) and spatial complexity (using more qubits).

In addition, we also show that the entanglement between the system and the clock qubits carries operational meaning since it serves as a bound for the infinity time average of the Loschmidt echo and for the temporal variance of expectation values. These results imply that the history state encodes valuable information in its correlations that can be used to study and understand the system's dynamics and equilibration. Given this operational meaning of the entanglement, we present two different schemes to compute the linear entropy of the history state, one based on the state-overlap circuit [45] and another one leveraging classical shadows and randomized measurements [46,47].

We also present a depth study of the circuits showing a clear advantage of using parallel-in-time protocols over the conventional sequential-in-time approaches where time is not mapped to clock qubits. Moreover, we propose a scheme to further reduce the circuit depth needed to prepare the history state via Hamiltonian diagonalization [48,49]. Here, we show that by leveraging tools from variational quantum algorithms [50–53] and quantum machine learning [54–56] to variationally diagonalize H, one can significantly further reduce the required circuit depth. For the special case when H is given by the Aubry-Andre model, we show that all of our algorithms can be implemented with a depth that only scales as $\mathcal{O}(\log(N)n)$, i.e., as the product of the number of clock and system qubits. Finally, we perform simulations that showcase the performance of our algorithms for studying temporal averages of systems evolving under an interacting and noninteracting Aubry-Andre model [57], illustrating the possibility of studying many-body localization.

II. QUANTUM TIME FORMALISM AND ITS DISCRETIZATION

Let us consider an n-qubit quantum system with associated Hilbert space $\mathcal{H}_S$. Then, let H be a time-independent Hamiltonian under which the system evolves. The dynamical evolution of the system is determined by the Schrödinger equation

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle\,, \tag{1}$$

where we have set $\hbar = 1$. It is well known that the solution of Eq. (1) is given by

$$|\psi(t)\rangle = U(t)|\psi_0\rangle\,, \quad \text{with} \quad U(t) = e^{-iHt} \tag{2}$$

and where $|\psi_0\rangle$ is some initial state of the system.

As previously discussed, and as shown in Fig. 1(a), there exists an inherent asymmetry between the space and time variables in quantum mechanics. Namely, the t variable over which we take a derivative is a fully classical parameter that is external to the quantum system. An alternative to fully incorporate time in a quantum framework is to introduce a Hilbert space $\mathcal{H}_T$ spanned by some states $|t\rangle$ [see Fig. 1(b)] such that $T|t\rangle = t|t\rangle$ and $[T, P_T] = i\hbar$, which in the time basis leads to $P_T \equiv -i\hbar\frac{d}{dt}$. Note that P_T is not the Hamiltonian of the system and in fact $[T, H] = 0$ (as they act on distinct Hilbert spaces). Evolution is then recovered from an extended Schrodinger equation, involving both the system and the clock Hilbert spaces, which is given by $\mathcal{J}|\Psi\rangle = 0$, for $\mathcal{J} = P_T \otimes \mathbb{1}_S + \mathbb{1}_T \otimes H$ and $|\Psi\rangle \in \mathcal{H}_T \otimes \mathcal{H}_S$. Here, $\mathbb{1}_T$ and $\mathbb{1}_S$, respectively, denote the identities on $\mathcal{H}_T$ and $\mathcal{H}_S$. In general, the extended Schrodinger equation, together with an initial condition, leads to entanglement between the system and the time Hilbert spaces.

The previous scheme can also be regarded as the mathematical basis of the PaW mechanism. Under this framework, the *universe state* $|\Psi\rangle$ is stationary (as $\mathcal{J}|\Psi\rangle = 0$), while the unitary evolution of the subsystem S emerges by conditioning on the rest. In our previous notation, this means that given a universe state

$$|\Psi\rangle = \int dt\ |t\rangle\,|\psi(t)\rangle, \tag{3}$$

we can recover the state of the system as $|\psi(t)\rangle = \langle t|\Psi\rangle$ [assuming $\langle t|t'\rangle = \delta(t - t')$]. More notably, one can readily see that $\mathcal{J}|\Psi\rangle = (i\partial_t - H)|\psi(t)\rangle = 0$ precisely recovers the standard Schrödinger equation with the index t being demoted from a quantum state label to a time parameter.

In order to make the states $|\Psi\rangle$ accessible to conventional (discrete) qubit-based quantum computers, one needs a proper discrete-time framework. Fortunately, it is easy to guess the form of a discrete-time *history state*. Namely, we start by introducing a finite-dimensional Hilbert space $\mathcal{H}_T$, which we denote as the *time* or *clock* Hilbert space with basis $|t\rangle$ satisfying $\langle t'|t\rangle = \delta_{tt'}$ for $t = 0, \ldots, N-1$. A discrete history state is then defined as the state

$$|\Psi\rangle = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}|t\rangle|\psi(\varepsilon t)\rangle, \tag{4}$$

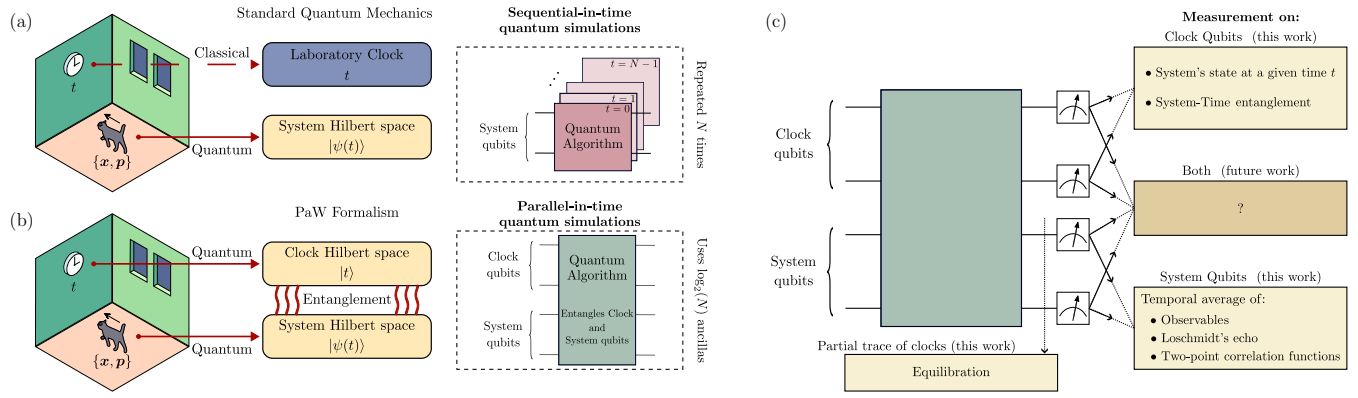


FIG. 1. Quantum algorithms based on the standard quantum mechanics or the PaW formalism. (a) In Hamiltonian classical mechanics, dynamical variables are functions of the phase-space coordinates position $\boldsymbol{x}$ and momentum $\boldsymbol{p}$. In standard quantum mechanics, one promotes $\boldsymbol{x}$ and $\boldsymbol{p}$ to quantum operators, but the time variable t is treated as a classical parameter that is external to the quantum system being studied. Quantum algorithms for studying dynamical properties based on this framework are implemented for some fixed time t. If we want to compute an average of N times, we need to repeat the run N different sequential-in-time experiments. (b) In the PaW formalism, time is treated as a quantum variable, with its own associated Hilbert space. In this work, we present quantum algorithms for parallel-in-time simulations that trade circuit repetitions for ancilla clock qubits. (c) After a proper entangling protocol, one has access not only to properties of the system at a given time, but also to their complete history. This information can be retrieved by performing measurements at the end of the circuit, which now can involve the clock qubits, the system qubits, or both. Measurements on the system that are conditioned to a certain time value give properties of the system at a given time. More interestingly, if one only measures on the system and completely ignores the clock's values, temporal averages are obtained. This is a consequence of the entanglement between the system and the clock, which induces a useful quantum channel when the clock is treated as an environment. Because of the quantum nature of the simulated clock and system, many other measurements can be proposed, meaning that the different protocols we discuss in this manuscript do not exhaust all the possibilities opened by this computational framework.

with $|\psi(\varepsilon t)\rangle = U(\varepsilon t)|\psi_0\rangle \in \mathcal{H}_S$. Here, we have $\varepsilon = T/N$ the time spacing for a given time window T, while t denotes a discrete dimensionless index (so that εt is a physical time interval).

In analogy with the continuum case, one can recover the state of the system at a given time by conditioning as $|\psi(\varepsilon t)\rangle\langle\psi(\varepsilon t)| = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|\Pi_t]/\langle\Psi|\Pi_t|\Psi\rangle$ for $\Pi_t = |t\rangle\langle t| \otimes \mathbb{1}_S$. In this way, the unitarily evolved state is recovered for the time values allowed by $\mathcal{H}_T$. Notice that this operation is different from a direct partial trace over the clock states, which generally yields a mixed state. It turns out that the partial trace induces a quantum channel that also encodes useful information about the system's dynamics and its (eventual) equilibration. In fact, one can think about the history states as a purification of that particular quantum channel. This is related to the system-time entanglement, as we discuss in Sec. IV and Appendix B.

Notice also that if the system itself is composite, namely, $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$, a partial trace over system B induces a mixed history state of the form $\mathrm{Tr}_B|\Psi\rangle\langle\Psi| = \frac{1}{N}\sum_{t,t'}|t\rangle\langle t'| \otimes \mathrm{Tr}_B\{|\psi(\epsilon t)\rangle\langle\psi(\epsilon t')|\}$ (see also Ref. [8]). Conversely, the history of a mixed state ρ can always be obtained from Eq. (4) by a purification of the initial state leading to $\mathrm{Tr}_B|\Psi\rangle\langle\Psi| = \frac{1}{N}\sum_{t,t'}|t\rangle\langle t'| \otimes U(\epsilon t)\rho U^\dagger(\epsilon t')$ that is just the convex combination of histories. The standard evolved reduced density matrix is again obtained by conditioning, namely, $\rho_A(t) = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|\Pi_t]/\langle\Psi|\Pi_t|\Psi\rangle$. For these reasons, we mainly focus on pure initial states, as the history states of mixed states are just obtained by standard convex combinations.

Here, we note that for the case of N being a power of two, the discrete history state can be prepared with the quantum phase estimation-like circuit of Fig. 2. For N being a power of two, one requires $\log(N)$ ancillary or clock qubits (we henceforth assume the logarithms to be base 2). As such, the clock Hilbert space $\mathcal{H}_T$ is of dimension $\dim(\mathcal{H}_T) = 2^{\log(N)} = N$. This result has been reported recently in Refs. [8,9], where the discrete history state of Eq. (4) has also been extensively studied.

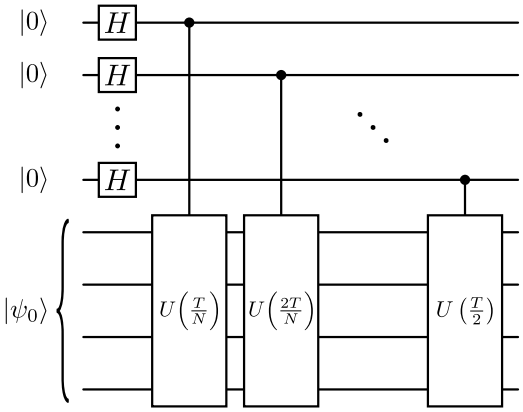


FIG. 2. Circuit for preparing history states. As shown above, the initial state to the clock qubits is $|0\rangle^{\otimes\log(N)}$ while that of the system is $|\psi_0\rangle$. The action of the Hadamard gates is to map the initial state to $|+\rangle^{\otimes\log(N)} \otimes |\psi_0\rangle$. Here, we find it convenient to write $|+\rangle^{\otimes\log(N)} = \frac{1}{\sqrt{N}}\otimes_{j=1}^{\log N}(|0_j\rangle + |1_j\rangle) = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}|t\rangle$ where we have expressed t in its binary form $t = \sum_{j=1}^{\log N} t_j 2^{j-1}$. Next, the $\log N$ controlled gates $U(2^{j-1}\frac{T}{N}) = U(\frac{T}{N})^{2^{j-1}}$ for $j = 1, \ldots, \log N$ perform the operations $U(\varepsilon t)|\psi_0\rangle = |\psi(\varepsilon t)\rangle$ for $U(\varepsilon t) = U(\frac{T}{N})^{\sum_{j=1}^{\log N} t_j 2^{j-1}}$.

The advantages of encoding history states in a quantum computer become clear once one starts considering measurements on the end of the circuit that are different from simple conditioning: While conditioned measurements allow one to recover properties of the system at a given time, genuinely quantum possibilities become accessible through the clock qubits. A small summary of such possibilities is provided in Fig. 1(c).

III. FROM QUBIT CLOCKS TO PARALLEL-IN-TIME SIMULATIONS

Here, we discuss how the mathematical formalism of qubit clocks presented in the previous section can be leveraged to create quantum algorithms aimed at studying averages of dynamical-in-time properties of quantum systems. In particular, in this section we focus on developing parallel-in-time-type algorithms that estimate time averages of physical quantities.

A. Setting

Given a time-independent Hamiltonian H acting on n qubits, and its associated time evolution operator $U(t) = e^{-iHt}$, we consider the problem of estimating general quantities of the form

$$\overline{F}(O_1, O_2, \omega) = \lim_{T\to\infty} \int_0^T \frac{dt}{T} e^{-i\omega t} \langle O_1(t) O_2 \rangle_\rho \tag{5}$$

$$= \lim_{T\to\infty} \int_0^T \frac{dt}{T} e^{-i\omega t} \mathrm{Tr}[\rho O_1(t) O_2], \tag{6}$$

where $O_1(t) = U^\dagger(t) O_1 U(t)$. Here, ρ is an n-qubit state acting on the d-dimensional Hilbert space $\mathcal{H}_S$ (with $d = 2^n$), O_1 and O_2 are two operators, and $\omega \in \mathbb{R}$.

To illustrate the relevance of the quantity $\overline{F}(\rho, O_1, O_2, \omega)$ in Eq. (5), let us consider several special cases. First, let $\omega = 0$ and $O_2 = \mathbb{1}$, which leads to

$$\overline{F}(O_1) := \overline{F}(O_1, \mathbb{1}, 0)$$

$$= \lim_{T\to\infty} \int_0^T \frac{dt}{T} \langle O_1(t) \rangle_\rho. \tag{7}$$

We can see that $\overline{F}(\rho, O_1)$ simply corresponds to an infinite temporal average of the observable O_1. These quantities are crucial to understanding the dynamical properties of closed quantum systems and in particular their equilibration [4,58,59]. They are also relevant to the study of quantum quench processes in field theories [60] and signatures of nonequilibrium quantum phase transition through infinite-time averages of Loschmidt echoes [61–65]. Next, when $\omega = 0$, we have

$$\overline{F}(O_1, O_2) := \overline{F}(O_1, O_2, 0)$$

$$= \lim_{T\to\infty} \int_0^T \frac{dt}{T} \langle O_1(t) O_2 \rangle_\rho. \tag{8}$$

Here, we can recognize $\langle O_1(t) O_2 \rangle_\rho$ as a two-point correlation function (also known as a dynamical Green's function). Two-point correlation functions are used to describe the behavior of a system under perturbations and are a widely used tool in quantum many-body systems and condensed matter physics [66–69]. The infinite-time average of $\langle O_1(t) O_2 \rangle_\rho$ has been recently considered in Ref. [70] to study thermodynamic properties of closed quantum systems, such as the emergence of dissipation at late times.

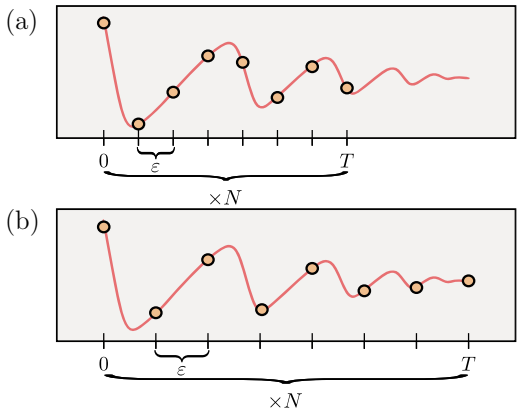


FIG. 3. Trade-off between accuracy and resolution. Consider the approximation of the infinite-time average of Eq. (5) given by the discrete sum in Eq. (9). For some fix number of time steps N, there exists a trade-off between the window size ε and the final time T. Namely, larger T implies larger window size ε, and hence less accuracy. On the other hand, smaller final time T implies more resolution in the temporal average at the cost of less accuracy.

Finally, we note that the general function $\overline{F}(O_1, O_2, \omega)$ corresponds to a Fourier transform of the two-point correlation function, which is commonly referred to as the dynamical structure factor in the condensed matter community [71,72]. Crucially, the dynamical structure factors are used to study dynamical properties of a given system and have the properties of being experimentally accessible [73,74], and usually being hard to compute via classical simulations [72].

While the importance of Eq. (5) is clear, the computation of $\overline{F}(O_1, O_2, \omega)$ might not be straightforward. On the one hand, the classical simulation of some quantum mechanical dynamical process is generally expected to be exponentially expensive in classical computers. Such scaling can be mitigated by using a quantum computer. Here, there are several schemes capable of computing fixed-time quantities of the form $\langle O_1(t) O_2 \rangle_\rho$ [71,75–77]. Still, the issue remains that one needs to perform the time average. In practice, this can be achieved via the discrete-time approximation

$$\widetilde{F}(O_1, O_2, \omega) = \frac{1}{N} \sum_{t=0}^{N-1} e^{-i\omega\varepsilon t} \langle O_1(\varepsilon t) O_2 \rangle_\rho, \tag{9}$$

where we have $\varepsilon = T/N$ (for simplicity, we will henceforth assume that N is a power of 2). That is, for a given (finite) time window T, we are computing the average over N points separated by a spacing ε. As shown in Fig. 3, the spacing ε determines the level of accuracy in the approximation, as a smaller ε leads to a more precise discretization of the integral and a better approximation of the true infinite-time average. On the other hand, the final time T determines the resolution of the approximation, as a larger T allows for a longer time interval to be averaged over, capturing more information about the system's behavior over time. One can see that both

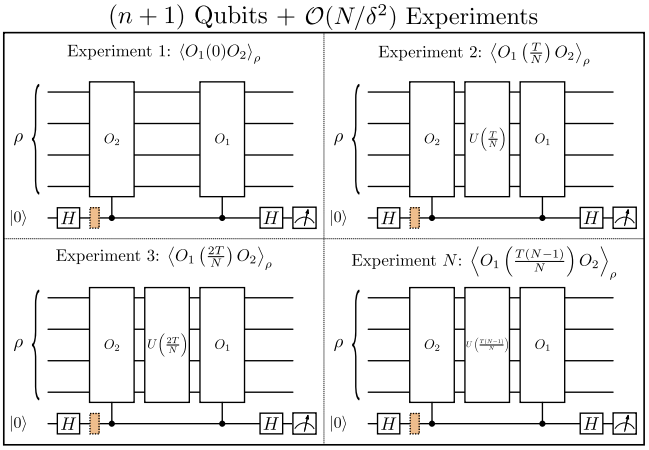


FIG. 4. Algorithm for sequential-in-time estimation of Eq. (9). The algorithm show can be used to individually compute each term in the summation. That is, the circuits can be used to estimate quantities of the form $\langle O_1(t)O_2\rangle_\rho$. Then, one can combine those expectation values classically (as well as add the appropriate phases $e^{-i\omega\varepsilon t}$) to estimate the quantity $\widetilde{F}(O_1, O_2, \omega)$ up to precision δ. The colored dashed gate is replaced with an identity (an $S^\dagger$ gate) to compute the real (imaginary) part of $\langle O_1(t)O_2\rangle_\rho$. This approach requires a quantum device with $(n+1)$ qubits and $\mathcal{O}(N/\delta^2)$ different experiments.

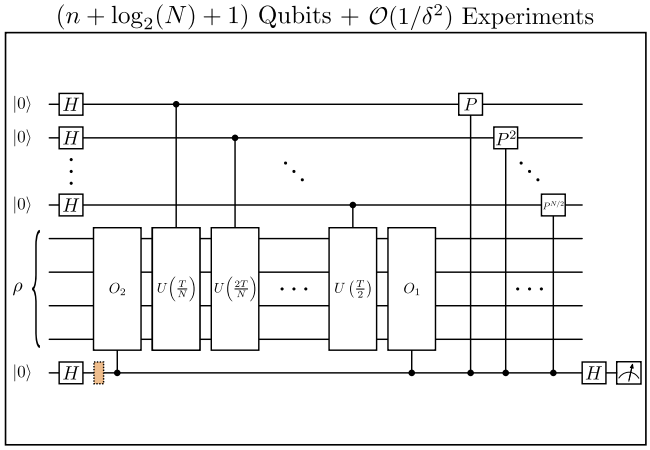


FIG. 5. Algorithm for parallel-in-time estimation of Eq. (9). The algorithm shown can be used to directly estimate the quantity $\widetilde{F}(O_1, O_2, \omega)$ up to precision δ. In the figure, P denotes a $\omega\varepsilon$ phase gate. It is clear that this algorithm contains as a subroutine the circuit for preparing the history state of Fig. 2. This approach requires a quantum device with $(n+\log(N)+1)$ qubits and $\mathcal{O}(1/\delta^2)$ experiments. The colored dashed gate is replaced with an identity (an $S^\dagger$ gate) to compute the real (imaginary) part of $\widetilde{F}(O_1, O_2, \omega)$.

the resolution and the accuracy can be improved by a larger number of discrete time steps N.

B. Sequential- and parallel-in-time protocols

Let us now consider the task of estimating $\widetilde{F}(O_1, O_2, \omega)$ when O_1 and O_2 are Pauli operators by either sequential- or parallel-in-time simulations. Here, by sequential, we mean that each term in the sum in Eq. (9) is estimated on a quantum device by running some finite number of "experiments." For instance, consider the circuit in Fig. 4, as explicitly shown in the Appendix C, it can be used to estimate an expectation value of the form $\langle O_1(\varepsilon t)O_2\rangle_\rho$. Thus, we have that the following proposition holds.

Proposition 1. The circuit in Fig. 4, which requires $(n+1)$ qubits, can be used to estimate the quantity $\widetilde{F}(O_1, O_2, \omega)$ of Eq. (9) up to δ accuracy with $\mathcal{O}(N/\delta^2)$ experiments.

The proof of Proposition 1, as well as that of all other main results, is presented in Appendix C.

Clearly, the fact that we need to sequentially estimate $\langle O_1(\varepsilon t)O_2\rangle_\rho$ for each $t = 0, \ldots, N-1$ leads to a complexity in the number of experiments (i.e., number of calls to the quantum computer) that scales as $\mathcal{O}(N)$. As we now show, this complexity can be reduced by using a scheme based on the discrete history state formalism, which allows us to directly estimate the whole sum of Eq. (9). That is, the following result holds.

Theorem 1. The circuit in Fig. 5, which requires $(n+\log(N)+1)$ qubits, can be used to estimate the quantity $\widetilde{F}(O_1, O_2, \omega)$ of Eq. (9) up to δ accuracy with $\mathcal{O}(1/\delta^2)$ experiments. Comparing Proposition 1 and Theorem 1 reveals that by leveraging the discrete history state formalism we can trade the $\mathcal{O}(N)$ experiment complexity for $\mathcal{O}(\log(N))$-ancillary qubits. That is, the parallel-in-time algorithm of Fig. 5 allows an exponential temporal-to-qubit resource trade-off. Here, we remark that one can see from Fig. 5 that the key step behind the algorithm to compute $\widetilde{F}(O_1, O_2, \omega)$ is the discrete history state. In fact, the circuit in Fig. 2 used to create the history state is a subroutine in Fig. 5. Thus, by leveraging $\log(N)$ ancillas, one can simultaneously implement all N time evolution operators $U(\varepsilon t)$ for $t = 0, \ldots, N-1$ and concomitantly compute all the terms in the summation leading to $\widetilde{F}(O_1, O_2, \omega)$.

Note that while Proposition 1 and Theorem 1 are derived and proved for the case of O_1 and O_2 being unitary operators, one can readily generalize the previous results for the case when they are instead expressed as a linear combination of Pauli operators. In particular, if

$$O_i = \sum_{\mu=1}^{M_1} c_\mu^{(i)} U_\mu, \tag{10}$$

for U_μ being a Pauli operator, then the experiment complexities in Proposition 1 and Theorem 1, respectively, change as $\mathcal{O}(NM_1M_2/\delta^2)$ and $\mathcal{O}(M_1M_2/\delta^2)$. Here, we again recover an exponential temporal-to-qubit resource trade-off by using the parallel-in-time algorithm.

Next, let us consider $\rho = O_1 = |\psi_0\rangle\langle\psi_0|$, $O_2 = \mathbb{1}$ and $\omega = 0$. In this special case,

$$\overline{F}(|\psi_0\rangle\langle\psi_0|, \mathbb{1}, 0) = \lim_{T\to\infty} \int_0^T \frac{dt}{T} |\langle\psi_0|U(t)|\psi_0\rangle|^2. \tag{11}$$

The quantity on the right-hand side is the infinite-time Loschmidt echo average [61,62,64,65], which we denote as $\bar{\mathcal{L}}(\psi_0)$. We see that

$$\bar{\mathcal{L}}(\psi_0) \equiv \overline{F}(|\psi_0\rangle\langle\psi_0|, \mathbb{1}, 0). \tag{12}$$

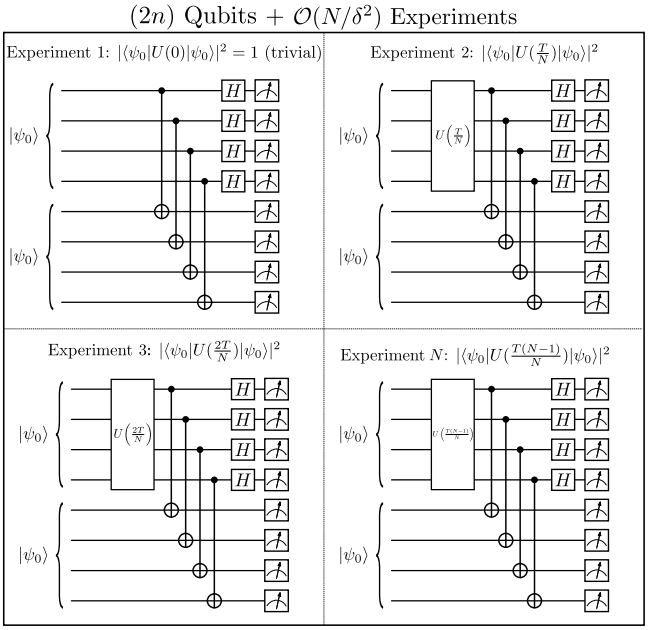


FIG. 6. Algorithm for sequential-in-time estimation of the Loschmidt echo of Eq. (13). We show an algorithm that computes, up to precision δ, the overlap between $|\psi_0\rangle$ and $U(\varepsilon t)\,|\psi_0\rangle$ for $t = 0, \ldots, (N-1)$. The algorithm is based on Bell-basis measurements as described in Ref. [45]. Once these overlaps are estimated, we can average them classically to estimate discrete-time temporal average of the Loschmidt echo $\widetilde{\mathcal{L}}(\psi_0)$. This approach requires a quantum device with $(2n)$ qubits and $\mathcal{O}(N/\delta^2)$ different experiments.

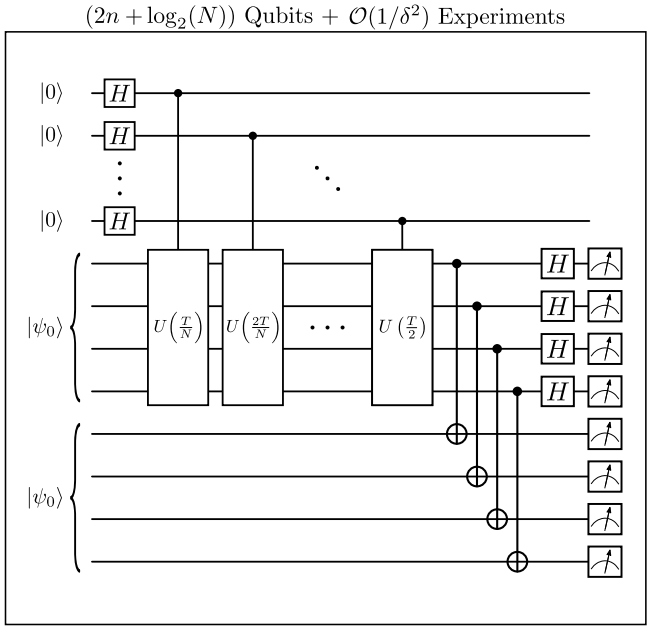


FIG. 7. Algorithm for parallel-in-time estimation of the Loschmidt echo of Eq. (13). We show an algorithm that computes, up to precision δ, the overlap between the discrete history state $|\Psi\rangle\langle\Psi|$ and $\mathbb{1}\otimes|\psi_0\rangle\langle\psi_0|$. As shown in Eq. (14), the overlap between these two states is equal to $\widetilde{\mathcal{L}}(\psi_0)$. The algorithm is based on Bell-basis measurements as described in Ref. [45]. This approach requires a quantum device with $(2n + \log(N))$ qubits and $\mathcal{O}(1/\delta^2)$ different experiments.

Similarly, for its discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$, we can write

$$\widetilde{\mathcal{L}}(\psi_0) \equiv \widetilde{F}(|\psi_0\rangle\langle\psi_0|, \mathbb{1}, 0) = \frac{1}{N}\sum_{t=0}^{N-1} |\langle\psi_0|U(\varepsilon t)|\psi_0\rangle|^2. \tag{13}$$

It is clear that while $\widetilde{\mathcal{L}}(\psi_0)$ can technically be computed with the circuits in Figs. 4 and 5, this requires expanding $O_1 = |\psi_0\rangle\langle\psi_0|$ into a linear combination of unitaries, and such summation will generally contain exponentially many terms. To mitigate this issue, we also present two results that allow us to estimate Eq. (13) by either sequential- or parallel-in-time simulations.

First, let us consider the following proposition.

Proposition 2. The circuit in Fig. 6, which requires $(2n)$ qubits, can be used to estimate the quantity $\widetilde{\mathcal{L}}(\psi_0)$ of Eq. (13) up to δ accuracy with $\mathcal{O}(N/\delta^2)$ experiments. Proposition 2 simply follows from applying a SWAP test [45,78–80] (or more specifically, the state-overlap algorithm of Ref. [45]) between $U(\varepsilon t)\,|\psi_0\rangle$ and $|\psi_0\rangle$ for $t = 1, \ldots, (N-1)$. We note that the case $t = 0$ is trivial as $|\langle\psi_0|U(0)|\psi_0\rangle|^2 = 1$. When using the discrete history state, we can prove the following theorem.

Theorem 2. The circuit in Fig. 7, which requires $(2n + \log(N))$ qubits, can be used to estimate the quantity $\widetilde{\mathcal{L}}(\psi_0)$ of Eq. (9) up to δ accuracy with $\mathcal{O}(1/\delta^2)$ experiments. Again, we can see from Theorem 2 that performing a parallel-in-time simulation allows us to exponentially reduce the experiment complexity (from linear in N to being N-independent) at the cost of $\log(N)$ ancillas. Similarly to Proposition 2, the proof of Theorem 2 simply spans from computing the overlap between the discrete history state $|\Psi\rangle\langle\Psi|$ and $\mathbb{1}_T\otimes|\psi_0\rangle\langle\psi_0|$. Explicitly, we have

$$\begin{aligned}
&\mathrm{Tr}[|\Psi\rangle\langle\Psi|\,(\mathbb{1}_T\otimes|\psi_0\rangle\langle\psi_0|)] \\
&\quad= \langle\Psi|(\mathbb{1}_T\otimes|\psi_0\rangle\langle\psi_0|)|\Psi\rangle \\
&\quad= \frac{1}{N}\sum_{t=0}^{N}\langle\psi(\varepsilon t)\,|\psi_0\rangle\langle\psi_0|\,\psi(\varepsilon t)\rangle \\
&\quad= \widetilde{\mathcal{L}}(\psi_0).
\end{aligned} \tag{14}$$

IV. ACCESSING DYNAMICAL INFORMATION VIA SYSTEM-TIME ENTANGLEMENT

Thus far, we have seen that using the history state allows us to push the complexity of running multiple experiments onto ancillary clock-qubit requirements. However, as we will now show, the entanglement present between the time and system qubits in the history state has operational meaning and contains information that we can use to learn about the system's dynamics. Moreover, we will unveil a rigorous and explicit connection between these correlations and the equilibration problem. Protocols for obtaining these quantities from variations of the previous circuits are also provided in this section.

A. Properties, relation to the problem of equilibration and to temporal fluctuations of observables

First, let us again recall that the discrete history state is a bipartite state between the system Hilbert space $\mathcal{H}_S$ and the time, or clock Hilbert space $\mathcal{H}_T$. That is, $|\Psi\rangle \in \mathcal{H}_T \otimes \mathcal{H}_S$. Moreover, it is apparent from Fig. 2 and Eq. (4) that history states are in general entangled across the system-time partition. We will henceforth refer to the correlations between the system qubits and the clock qubits as system-time entanglement (following Ref. [8]).

It is important to note that, in general, Eq. (4) is not in the Schmidt's decomposition [81] of $|\Psi\rangle$ (as the states $|\psi(t)\rangle$ are not necessarily orthogonal). However, there exists a basis in which we can write the history state as

$$|\Psi\rangle = \sum_l \sqrt{p_l}|l\rangle_T|l\rangle_S, \tag{15}$$

where $\sqrt{p_l}$ are the so-called Schmidt coefficients, and $\{|l\rangle_S\}$ and $\{|l\rangle_T\}$ are orthonormal sets of states in $\mathcal{H}_S$ and $\mathcal{H}_T$, respectively. A simple way to quantify the system-time entanglement is through the linear entropy, defined as

$$E_2 = 1 - \mathrm{Tr}\big[\rho_T^2\big] = 1 - \mathrm{Tr}\big[\rho_S^2\big] = 1 - \sum_l p_l^2, \tag{16}$$

where $\rho_{T(S)} = \mathrm{Tr}_{S(T)}[|\Psi\rangle\langle\Psi|]$ is the reduced state of the history state in the clock (system) qubits. Here, we denote as $\mathrm{Tr}_{S(T)}$ the partial trace over the system (clock) qubits. In principle, one can also consider other entropies such as the von Neumann entropy. However, the linear entropy has the desirable property of being efficiently computable in a quantum device (see below).

There is a deep connection between the system-time entanglement and dynamical properties of the system, in particular to the problem of its equilibration: Let us recall first that given an arbitrary (for simplicity) pure state $|\psi\rangle = \sum_k c_k|k\rangle$ the infinite-time average of the associated density matrix is

$$\bar{\rho} = \int \frac{dt}{T} \sum_{k,k'} c_k c_{k'}^* e^{-i(E_k - E_{k'})t} |k\rangle\langle k'| \to \sum_k |c_k|^2 |k\rangle\langle k|, \tag{17}$$

where one assumes large (infinite) T and with $H|k\rangle = E_k|k\rangle$. In other words, if the state of the system is averaged over large enough times it loses all coherences in the energy basis. Under experimentally realistic conditions, it is feasible to identify this state with the stationary equilibrium state [4]. For "most" observables, this actually holds for short times T [59], meaning that a finite time window average of observables is also an interesting quantity in general. The quantum time formalism gives a new interpretation to the loss of coherences induced by a time average: Since the system is "entangled with time," we lose information by ignoring the "clock qubits." This loss induces precisely the (dephasing) quantum channel $\rho \to \bar{\rho}$ in the large T and small ε limit, a result that can be derived directly from a continuum quantum time formalism [82]. For discrete time, the following result holds.

Theorem 3. Let $|\Psi\rangle$ be the discrete history state in Eq. (4). The partial trace over the clock induces a quantum channel that in the large time limit implies $\rho_S \to \bar{\rho}$. Moreover, for any ε and N the following majorization relation holds:

$$\bar{\rho} \prec \rho_S = \tilde{\rho}, \tag{18}$$

with $\tilde{\rho}$ a discretization of Eq. (17). Furthermore, for a periodic evolution with period τ generated by a Hamiltonian with M distinct eigenvalues (i.e., $e^{-iH\tau} = \mathbb{1}$) and given a history state with $\log(M)$ clock qubits and time window $T = \tau$, we have

$$\rho_S = \bar{\rho}. \tag{19}$$

While phrased in a rather abstract way, this result has many interesting corollaries with clear operational meaning. The reason for this is that roughly speaking the history state is providing a way to prepare the equilibrated state of a quantum system: One simply needs to prepare the history state and ignore the clock qubits. In fact, this is the reason why the previous for evaluating time averages work. Moreover, the system-time entanglement entropies are in fact a lower bound to the entropies of the state in equilibrium, as it follows directly from Theorem 3 and basic majorization properties. Furthermore, we show in Appendix B that one can *rediscover* the quantum time formalism from the natural purification of this approximate dephasing channel: The history state arises from a simple isometric extension $U[K_t]$ of the channel as $|\Psi\rangle = U[K_t]|\psi_0\rangle$ with K_t the Krauss operators, $K_t = e^{-iHt\epsilon}/\sqrt{N}$. The interested reader can refer to Appendix B where the proof of Theorem 3 is provided together with a more detailed discussion.

With the previous in mind, let us consider again the task of estimating the infinite-time Loschmidt echo average in Eq. (11). We recall that $\bar{\mathcal{L}}(\psi_0)$ quantifies the degree of reversibility of the time evolution and is an indicator of the stability of the quantum system. Moreover, it is easy to see that

$$\bar{\mathcal{L}}(\psi_0) = \mathrm{Tr}[\bar{\rho}^2]; \tag{20}$$

i.e., the infinite-time average of the Loschmidt echo is the purity of the dephased state $\bar{\rho}$. We can now use these considerations and Theorem 3 to obtain the following result.

Corollary 1. Let $|\Psi\rangle$ be the discrete history state in Eq. (4), and let E_2 be the linear entropy of the system-time partition. Then, for any T and N we have

$$E_2 \leqslant (1 - \bar{\mathcal{L}}(\psi_0)). \tag{21}$$

Corollary 1 has several important implications. First, it bounds the amount of entanglement between the system and the clock qubits. In particular, it shows that the system-time entanglement can only be large if the infinite-time average of the Loschmidt echo value is small. Conversely, if $\bar{\mathcal{L}}(\psi_0)$ is large, E_2 has to be small. Second, let us remark that Eq. (21) is valid for all values of T, but most notably, also for all values of N. For large N and T, the equality is reached asymptotically, and we have that Eq. (21) becomes $\mathrm{Tr}[\rho_T^2] \equiv \bar{\mathcal{L}}(\psi_0)$. Moreover, as we will see below, our numerical analysis shows that $\mathrm{Tr}[\rho_T^2]$ can provide a better approximation to $\bar{\mathcal{L}}(\psi_0)$ than $\widetilde{\mathcal{L}}(\psi_0)$, implying that there exists no simple general relation between E_2 and $\widetilde{\mathcal{L}}(\psi_0)$.

We can understand the intuition behind Corollary 1 as follows. Let $|\psi_0\rangle$ be a stationary state of the unitary evolution. For instance, let $|\psi_0\rangle$ be an eigenstate of H with eigenenergy

E_0, so that $U(\varepsilon t)\,|\psi_0\rangle = e^{-i\varepsilon t E_0}\,|\psi_0\rangle$. Then, the discrete history state becomes

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_{t=1}^{N-1} e^{-i\varepsilon t E_0}\,|t\rangle \otimes |\psi_0\rangle\,. \tag{22}$$

Equation (22) reveals that $|\Psi\rangle$ is separable. It is also not hard to verify that in this case $\widetilde{\mathcal{L}}(\psi_0) = 1$. On the other hand, if $|\psi_0\rangle$ evolves through N orthogonal states $\langle\psi(\varepsilon t)|\,|\psi(\varepsilon t')\rangle = \delta_{tt'}$ then Eq. (4) is already the Schmidt decomposition of $|\Psi\rangle$ and the state is maximally entangled. The previous toy model shows that if the state is quasistationary (i.e., large Loschmidt echo), we can expect small values of entanglement. Similarly, if the state is significantly changing during the evolution (e.g., small Loschmidt echo value), then the history state will likely possess large amounts of entanglement. We note that the relation between the distinguishability of the evolved state and the system-time entanglement was first reported in Ref. [8]. However, the connection with the Loschmidt echo was not explored therein.

The result in Corollary 1 can be further strengthened for the special case where the time evolution is periodic. That is, when

$$e^{-iH\tau} = \mathbb{1}, \tag{23}$$

for some τ, and where we assume that H has M distinct eigenvalues, for M being a power of two. Now, we find that the following result holds.

Corollary 2. For a periodic evolution with period τ generated by a Hamiltonian with M distinct eigenvalues, as in Eq. (23), then for a history state with $\log(M)$ clock qubits and time window $T = \tau$, we have

$$E_2 = (1 - \bar{\mathcal{L}}(\psi_0)) = (1 - \widetilde{\mathcal{L}}(\psi_0)). \tag{24}$$

Corollary 2 shows that for periodic Hamiltonians the system-time entanglement is exactly the same as the infinite-time average of the Loschmidt echo $\bar{\mathcal{L}}(\psi_0)$, as well as the discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$. As shown in Appendix B, tracing out induces now a completely dephasing channel in the energy eigenbasis so that $\rho_S = \tilde{\rho} = \bar{\rho}$.

The previous results connecting the system-time entanglement with the Loschmidt echo allow us to derive even more operational meaning to E_2 as a bound for temporal fluctuations of observable. In Ref. [4], it was shown that given an observable O, $\bar{\mathcal{L}}(\psi_0)$ provides a bound on temporal fluctuations of observables as

$$\sigma_O^2 \leqslant \Delta_O^2 \bar{\mathcal{L}}(\psi_0), \tag{25}$$

with $\Delta_A^2 = \lambda_{\max}[O] - \lambda_{\min}[O]$ (the difference between the largest and smallest eigenvalues of O in the subspace of states satisfying $\langle n|\psi\rangle \neq 0$) and where σ_O^2 denotes the temporal variance

$$\begin{aligned}\sigma_O^2 &:= \overline{F(O)^2} - \overline{F}(O)^2 \\ &= \lim_{T\to\infty} \int_0^T \frac{dt}{T} \langle O(t)\rangle_{\psi_0}^2 - \left(\lim_{T\to\infty} \int_0^T \frac{dt}{T} \langle O(t)\rangle_{\psi_0}\right)^2.\end{aligned} \tag{26}$$

Here, we have used the notation defined in Eq. (7) with $F(O) \equiv \langle O\rangle$ (at a given time) while the "overline" denotes temporal average. Equation (25) shows that small temporal Loschmidt echo averages imply a small temporal variance of the observable O, and vice versa. In other words, a system with a small $\bar{\mathcal{L}}(\psi_0)$ can only exhibit smaller temporal fluctuations in its observables compared to a system with a large Loschmidt echo.

It should be clear to see that Theorem 1 readily implies the following corollary.

Corollary 3. Let O be an observable, and let σ_O^2 denote its temporal variance as in Eq. (26). The system-clock entanglement provides bound on temporal fluctuations as

$$\sigma_O^2 \leqslant \Delta_O^2 (1 - E_2) = \Delta_O^2 \mathrm{Tr}\big[\rho_S^2\big]. \tag{27}$$

Corollary 3 shows a clear physical meaning of the system-time entanglement. Namely, if E_2 is small, then the system is stable and predictable. This follows from the fact that the temporal variances of expectation values will be small. Conversely, if the system-time entanglement is large, then the system can be unstable and unpredictable, as evidenced by potentially large observable fluctuations.

B. Protocols for computing the system-time entanglement

The previous theorems and corollary shed light on the exciting possibility of understanding the dynamics of the system through the system-time entanglement. However, in order for these results to be truly useful, one needs to be able to measure E_2 from the history state. As we can see in Eq. (16), we need to estimate $\mathrm{Tr}[\rho_S^2]$ or $\mathrm{Tr}[\rho_T^2]$. While mathematically it makes no difference whatsoever which subsystem we focus on, as their purity is the same [see Eq. (15)], in practice it can be substantially easier to work with one system or the other.

As heuristically evidenced by our numerics (see below), the discrete history state with a number of clock qubits $\log(N)$ much smaller than the system size n produces results that accurately reproduce the infinity time average properties of the system dynamics. Thus, we will henceforth assume that $\log(N) \ll n$. This assumption implies that we can compute E_2, and therefore learn about the system, by just looking at the clock qubits. We now present two methods for estimating $\mathrm{Tr}[\rho_T^2]$.

Theorem 4. The quantity E_2 of Eq. (16) can be estimated up to δ accuracy with the circuit in Fig. 8, which requires $(2n + 2\log(N))$ qubits with $\mathcal{O}(1/\delta^2)$ experiments. Similarly, it can also be estimated with the circuit in Fig. 9, which requires $(n + \log(N))$ qubits with $\mathcal{O}(N/\delta^2)$ experiments. When using the circuit in Fig. 8, one prepares two copies of the history state $|\Psi\rangle$ and then performs the state-overlap circuit of Ref. [45]. On the other hand, when using the circuit in Fig. 9 one can estimate E_2 with a single copy of $|\Psi\rangle$ by using classical shadows, or randomized measurements [46,47]. For instance, one can prepare the history state and perform a random unitary on each qubit, followed by a measurement on the computational basis. The measurement outcomes are stored and then combined classically to estimate $\mathrm{Tr}[\rho_T^2]$.

To finish this section, we note that by comparing Proposition 2, Theorem 2, and Theorem 4, the method to estimate either $\widetilde{\mathcal{L}}(\psi_0)$ or E_2 with the least computational requirement [assuming $\log(N) \ll n$] is that of Fig. 9. Namely, here we can compute E_2 up to δ precision with a quantum computer with

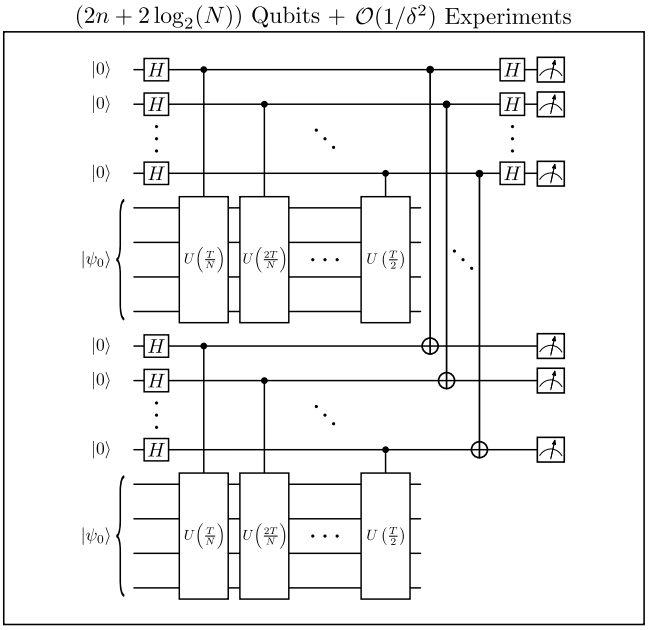


FIG. 8. Algorithm for estimating E_2 via state overlap. Here, we consider the task of evaluating Eq. (16). By taking two copies of the history state, we can estimate $\mathrm{Tr}[\rho_T^2]$ up to δ precision via the state-overlap circuit in Ref. [45]. This approach requires a quantum device with $(2n + 2\log(N))$ qubits and $\mathcal{O}(1/\delta^2)$ different experiments.

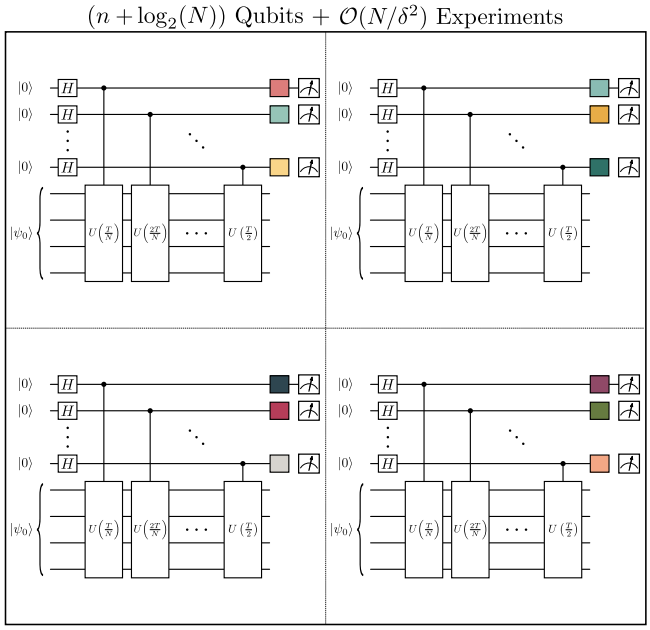


FIG. 9. Algorithm for estimating E_2 via randomized measurements. Here, we consider the task of evaluating Eq. (16). We start with a copy of the history state, and we apply a random unitary (indicated by a colored gate) to each qubit. Then, we measure each qubit in the computational basis and record the measurement outcome. These constitute the so-called "classical shadows" of ρ_T. As shown in Ref. [46], this procedure allows us to estimate $\mathrm{Tr}[\rho_T^2]$ up to δ precision with a quantum device with $(n + \log(N))$ qubits and $\mathcal{O}(N/\delta^2)$ different experiments.

$(n + \log(N)) \ll 2n$ qubits and with $\mathcal{O}(N/\delta)$ experiments. This result then showcases the power of using the history state as it allows us to study physical properties of the system [such as bounding $\bar{\mathcal{L}}(\psi_0)$ or the temporal variances ΔO^2] with less requirements than we would otherwise need.

V. DEPTH-ESTIMATION AND PARALLEL-IN-TIME ADVANTAGES

In the previous sections, we have presented several methods where we used the history state to study temporal averages of quantities of the form of $\widetilde{F}(O_1, O_2, \omega)$ in Eq. (9). At the same time, we have shown how to compute the system-time entanglement, a quantity with many interesting applications. Crucially, these techniques require being able to implement the phase estimation-like circuit for preparing history states in Fig. 2 as a subroutine. As it is well known, under the assumption of black boxes implementing the controlled unitaries only $\log(N)$ steps are necessary [81] for such protocols, which seems to indicate an exponential advantage over sequential approaches. The most famous example of such an advantage is provided by Shor's algorithm [83] where the controlled unitaries can be efficiently prepared via modular exponentiation. In this section, we discuss the depth scaling under the more realistic scenario where the controlled unitaries associated with quantum evolution need to be explicitly implemented. We provide two different strategies: a direct Lie-Trotter product formula [84] approach and a Hamiltonian diagonalization scheme that we implemented variationally [48]. The second strategy, when successful, maintains the black box scaling leading to an exponential advantage. Instead, the Trotterization scheme leads to a more modest quadratic advantage for large enough N.

A. Direct Trotterization approach

We start by recalling that $U(\varepsilon t) = e^{-iH\varepsilon t}$ for some Hamiltonian of interest H. Usually, if one wishes to implement $U(\varepsilon t)$, the standard approach is to break the evolution into smaller, easier to implement evolutions $U(\varepsilon)$, and then repeat it t times. That is, one has $U(\varepsilon t) = U(\varepsilon)^t$. In this way, the depth of the circuit needed to implement $U(\varepsilon t)$ grows with t.

To be more specific, let us assume that one is employing a Lie-Trotter decomposition-based product formula. This was the first example of quantum advantage for quantum simulations [84] and remains a relevant and straightforward technique to this day [85]. The basic idea is to decompose a given Hamiltonian

$$H = \sum_{j=1}^{l} h_j \tag{28}$$

as $U(\varepsilon) \approx \prod_j e^{-ih_j\varepsilon}$. We will further assume that each h_j is local; i.e., it acts nontrivially on at most $\mathcal{O}(1)$ qubits (similar considerations hold if l grows polynomically with n). Importantly, note that the locality condition implies that l is linear in the system size. The evolution operator up to time $\tau = t\varepsilon$ can be approximated with t copies of these gates as $U(\tau) \approx [\prod_j e^{-ih_j\varepsilon}]^t$, which is also logarithmic in the system size.

A rough estimation of the error involved was provided in Ref. [84] by assuming that the main contribution to the error comes from the second-order term in the Lie-Trotter formula. Under this assumption, the number of times steps required for guaranteeing a fixed precision grows as $t \propto \tau^2$. Then, the total number of gates, denoted as $\#_g(\tau)$, scales as $\#_g(\tau) = \gamma l \tau^2$ for γ a constant dependent on the precision and the particular Hamiltonian. More general bounds were found in Ref. [86], which allows us to write $\#_g(\tau) = \gamma l \tau^\alpha$. This is the estimation we will be using. Under certain scenarios, this bound can be improved, e.g., by using the Lie-algebraic structure associated with the given Hamiltonian [87], and in general, the "actual" error scaling of such product formulas remains poorly understood [85]. For our purpose, the previous bound will suffice: We want to compare the total number of gates required in the sequential- (Figs. 4 and 6) versus the parallel-in-time approach (Figs. 5 and 7) assuming the same Trotterization scheme is applied to both. The interest in this quantity relies on the fact that the *total* number of gates employed in each simulation protocol is what determines the total time span required to complete the computation [88]. By estimating this quantity in each protocol, we can establish whether it is more convenient to use a sequential- or parallel-in-time approach.

Theorem 5. Consider the total number of gates required for implementing the evolution in the sequential approaches $\#^{\text{seq}}$, and the total number for the parallel-in-time approaches $\#^{\text{par}}$. They scale as $\#^{\text{seq}} \in \mathcal{O}(lN^{\alpha+1})$, $\#^{\text{par}} \in \mathcal{O}(l^2 N^\alpha)$ yielding

$$\#^{\text{par}} \sim \frac{\beta l}{N} \#^{\text{seq}}, \tag{29}$$

for $\beta \in \mathcal{O}(1)$, a constant independent both of the system and clock size.

Theorem 5 is a consequence of the fact that $\#^{\text{seq}}$ is given by the sum over the amount of gates of each run so that $\#^{\text{seq}} \sim N^{\alpha+1}$. Instead, in the parallel approach the total number gates only involves a sum over the $\log(N)$ gates of the same run, thus giving $\#^{\text{seq}} \sim N^\alpha \beta l$. The extra factor βl comes from the fact that those gates need to be controlled. Remarkably, the depth scaling with the number of times of the parallel-in-time approach is the same as the one of a single Trotter evolution up to time $\tau \equiv \varepsilon N$. See Appendix D for the details and the proof of the previous theorem.

Something really interesting has happened: In the parallel approach, we have an increase in depth that is *linear* in the system size, but we have reduced the total number of gates *exponentially* in the number of clock qubits with respect to a sequential approach (equivalent to a quadratic improvement in the total number of times N). We can then state the following:

Proposition 3. Given $\log(N)$ clock qubits and system of Hilbert space dimension d, the parallel-in-time approach outperforms the computational times of the sequential approach for

$$\log(N) \gtrsim \log(\beta l) \sim \log(\beta \log(d)). \tag{30}$$

Remarkably, the condition for a convenient clock size is doubly logarithmic in the system's Hilbert space dimension d. Typically, a modest number of qubits for the clock, much smaller than the system size, is sufficient to improve computational times. As an example for $n = 200$ system qubits and for $\beta = 5$, the inequality $\log(N) > \log(\beta \log(d))$ is achieved when $m \geqslant 10$, while for $n = 400$ one needs $m = 11$ clock qubits. Let us also remark that in most applications it is natural to scale N with the system size to properly capture the system behavior. In this sense, a quadratic speed up in N could indicate a substantial improvement when measured by the dimension of the system.

Finally, let us remark that Proposition 3 will hold under rather general conditions, since it is based on the fact that a sum over N terms is involved in the estimation of $\#^{\text{seq}}$, while a sum over $\log(N)$ terms is required for $\#^{\text{par}}$ (see proof in Appendix F). However, the scaling of the number of gates $\#^{\text{par}}$ we provided in Theorem 5 is still based on a pessimistic bound and linked to product formulas: The generic bounds we used for Trotterization can overestimate by far [89,90] the actual errors, which depend on the specific initial states and observables involved in the complete protocols. This means that actual implementations of the parallel-in-time protocols, whether based on product formulas or more advanced methods, might be much more efficient. The important message is that the advantages over sequential-in-time protocols, as stated in Proposition 3, hold more generally (see also below and the discussion).

B. Hamiltonian diagonalization and Cartan decomposition approach

In this section, we repeat the circuit depth analysis in another relevant scheme, namely, assuming one has access to a diagonalization of the Hamiltonian. In particular, we will also discuss how one can obtain such diagonalization variationally via the algorithm presented in Ref. [48].

Let us recall that there always exists a unitary W (whose columns are the eigenvectors of H) and a diagonal matrix D (whose entries are the eigenvalues of H) such that

$$H = WDW^\dagger. \tag{31}$$

Without loss of generality, we can expand D in some basis of mutually commuting operators

$$D = \sum_\mu c_\mu h_\mu, \tag{32}$$

where $[h_\mu, h_{\mu'}] = 0$ for all μ, μ'. If one has access to W and D, then the unitary evolution can be expressed as

$$\begin{aligned} U(\varepsilon t) &= U(\varepsilon)^t \\ &= W e^{-iD\varepsilon t} W^\dagger \\ &= W e^{-i\sum_\mu c_\mu \varepsilon t h_\mu} W^\dagger \\ &= W \left(\prod_\mu e^{-ic_\mu \varepsilon t h_\mu} \right) W^\dagger. \end{aligned} \tag{33}$$

The power of Eq. (33) can be seen from Fig. 10, where it is shown that we can use the diagonalization of H to implement $U(\varepsilon t)$ at fixed depth. Namely, the circuit depth for $\prod_\mu e^{-ic_\mu \varepsilon t h_\mu}$ and for $\prod_\mu e^{-ic_\mu \varepsilon t h_\mu t'}$ is exactly the same; we just change the parameters associated with each time evolution generated by h_μ. This means that in a sequential approach, each run requires $\mathcal{O}(n)$ gates independently of the evolution time and assuming for simplicity that each h_μ acts

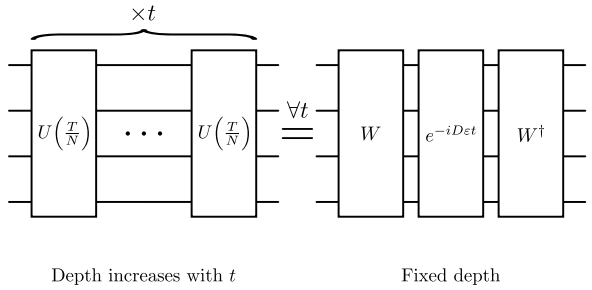


FIG. 10. Depth reduction via Hamiltonian diagonalization. On the left, we implement $U(\varepsilon t)$ by expanding the evolution into shorter-time evolutions $U(\varepsilon)$ (which we can then implement via Trotterization), and then perform $U(\varepsilon t) = U(\varepsilon)^t$. This comes at the cost of increasing the depth as t is increased. On the right, we use the diagonalization of H as in Eq. (31) to express $U(\varepsilon t) = We^{-iD\varepsilon t}W^\dagger$. Moreover, it can be seen by expanding D in a basis of mutually commuting operators [as in Eq. (32)] that the circuit implementation of $e^{-iD\varepsilon t}$ has the same depth for any t [see also Eq. (33)].

as a one-body operator. The total number of gates is then $\#^{\text{seq}} = \mathcal{O}(nN)$.

Moreover, the benefits of diagonalizing the Hamiltonian H are amplified when using this technique in the circuit for preparing the history state. As shown in Fig. 11(a), we can see that instead of controlling $\log N$ gates $U(2^{j-1}\varepsilon)$ (where for $j = 1, \ldots, \log N$), the history state can be prepared by first acting on the system qubits with the noncontrolled unitary W, followed by $\log N$ controlled gates $e^{-iD2^{j-1}\varepsilon}$, and finally by implementing a noncontrolled unitary $W^\dagger$. This further reduces the depth required to prepare the history state. First, we do not need to control W, nor $W^\dagger$. Second, we note that controlling $e^{-iD\varepsilon t}$ (for any t) is equivalent to controlling each term $e^{-ic_\mu \varepsilon t h_\mu}$ (since h_μ are mutually commuting). For instance, we can see in Fig. 11(b) that if h_μ are single-qubit Pauli operators acting on each qubit, then implementing a controlled $e^{-iD\varepsilon t}$ gate just requires controlling n single-qubit rotations. This gives the following theorem.

Theorem 6. By replacing the history state preparation subroutine with the diagonalized Hamiltonian as in Eq. (31) and Fig. 11, we can prepare the history state with a total number of gates

$$\#^{\text{par}} \in \mathcal{O}(\log(N)n), \tag{34}$$

i.e., logarithmic in both the number of times and system size.

The parallel-in-time advantage over the sequential approach condition now becomes $N > \beta \log(N)$, which is independent of the system size and is virtually always reached. Of course, one could argue that if we have classical access to the diagonalized Hamiltonian, then we could just expand the initial state state and the measured operators in the energy eigenbasis to compute any expectation value. However, this kind of expansion will not be tractable for large problem sizes. Instead, we will show below that if W is accessible in a quantum computer, then one can still leverage the diagonalization for depth reduction.

Finally, we note that if we want to study the entanglement in ρ_T as in Theorems 1 and 4 (see also Figs. 8 and 9), then the final unitary $W^\dagger$ in Fig. 11(a) can be omitted. This is due to the fact that the entanglement is invariant under local unitaries [91], and hence $W^\dagger$ cannot change the entanglement nor the spectral properties of ρ_T.

It is worth highlighting the fact that the main challenge for using Eq. (33) is that it requires access to the decomposition in Eq. (31), and that W and D might not be readily accessible. However, one can still attempt to variationally learn them [50]. For example, one can use the Variational Hamiltonian Diagonalization (VHD) algorithm in Ref. [48], which is aimed at training a parametrized ansatz for the diagonalization of H. The ansatz is composed of two parts: (1) a parametrized unitary $W(\boldsymbol{\alpha})$ and (2) a diagonal Hamiltonian $D(\boldsymbol{\beta})$ such that

$$\widehat{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = W(\boldsymbol{\alpha})D(\boldsymbol{\beta})W^\dagger(\boldsymbol{\alpha}). \tag{35}$$

One can quantify how much $\widehat{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ approximates the target Hamiltonian H by defining the cost function

$$C(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{\|H - \widehat{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\|_{\text{HS}}^2}{2^n}, \tag{36}$$

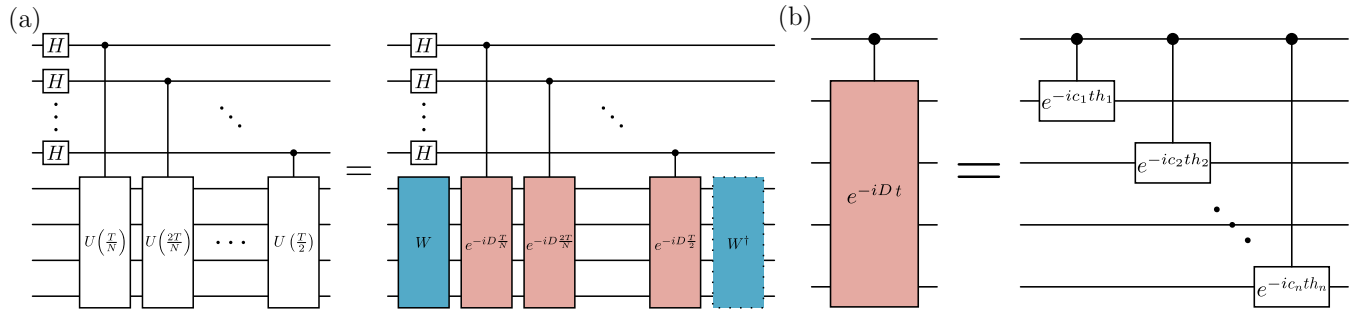


FIG. 11. History state via Hamiltonian diagonalization. (a) Here, we show how the circuit in Fig. 2 for preparing the history state changes when using the Hamiltonian diagonalization. In particular, we see that instead of controlling the $\log N$ gates $U(2^{j-1}\varepsilon)$ for $j = 1, \ldots, \log N$ we now need to implement noncontrolled unitaries W and $W^\dagger$, and simply control the fixed-depth $\log N$ gates $e^{-iD2^{j-1}\varepsilon}$ for $j = 1, \ldots, \log N$. We note that if we only care about the entanglement in ρ_T as in Figs. 8 and 9, then the final unitary $W^\dagger$ can be omitted (the entanglement is invariant under local unitaries [91]). (b) For the special case when the diagonal Hamiltonian D is expressed as a sum of single-qubit operators acting on each qubit [i.e., h_μ in Eq. (32) is a one-body operator acting on the μth qubits], then controlling the gates $e^{-iD2^{j-1}\varepsilon}$ simply requires controlling single-qubit gates.

where $\|X\|_{\rm HS} = \sqrt{{\rm Tr}[X^\dagger X]}$ is the Hilbert-Schmidt norm. Clearly, the cost is equal to zero if $H = \widehat{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Thus, the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are trained by solving the optimization task

$$\arg\min_{\boldsymbol{\beta},\boldsymbol{\alpha}} C(\boldsymbol{\alpha}, \boldsymbol{\beta}). \tag{37}$$

Here, a quantum computer is used to estimate the term in $C(\boldsymbol{\alpha}, \boldsymbol{\beta})$ [48], while classical optimizers are used to train the parameters.

In this variational setting, it is extremely important to pick an ansatz [i.e., a given unitary $W(\boldsymbol{\alpha})$, and diagonal Hamiltonian $D(\boldsymbol{\beta})$] that does not lead to trainability issues such as barren plateaus [56,92–95], where the cost function gradients are exponentially suppressed with the problem size. One of the leading strategies to mitigate such issues is to use the so-called problem-inspired ansätze, where one creates ansätze with strong inductive biases [96–98] based on the problem at hand. Recently, one such method was developed that exploits the Cartan decomposition of the Lie algebra generated by the target Hamiltonian H [49,99,100]. Below we explain such a method.

Consider the Hamiltonian of interest H. Then, without loss of generality we assume that it can be expressed as a sum of Hermitian traceless operators $\{H_i\}$ as

$$H = \sum_i a_i H_i, \tag{38}$$

where $a_i \in \mathbb{R}$. Then, let $\mathfrak{g} = \langle\{iH_i\}\rangle_{\rm Lie}$ be the Lie closure of the set of operators [101]. Note that, by definition, $iH \in \mathfrak{g}$. The result in Ref. [49] provides an efficient-in-dim($\mathfrak{g}$) circuit for the simulation of any e^{-iHt} for any set of coefficients $\{a_i\}$. To understand the technique of Ref. [49], we recall that a Cartan decomposition of the Lie algebra $\mathfrak{g}$ refers to the decomposition of $\mathfrak{g}$ into two orthogonal subspaces $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$, where $\mathfrak{k}$ is a Lie subalgebra, i.e., $[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}$, whereas $\mathfrak{m}$ is not: $[\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{k}$. Moreover, these two orthogonal subspaces satisfy $[\mathfrak{k}, \mathfrak{m}] = \mathfrak{m}$, and $\mathfrak{m}$ contains the maximal commutative subalgebra, also known as the Cartan subalgebra, $\mathfrak{h}$ of $\mathfrak{g}$. Note that for any pair of elements ih_1 and ih_2 in $\mathfrak{h}$, we have $[h_1, h_2] = 0$.

The Cartan decomposition provides us an ansatz to diagonalize H as follows. First, we note that H always admits a decomposition of the form

$$H = WhK^\dagger, \tag{39}$$

with $K \in e^{\mathfrak{k}}$ and $h \in \mathfrak{h}$. Reference [49] provides us with an ansatz for Eq. (35) as we can now parametrize and optimize over the Lie group $e^{\mathfrak{k}}$ and the algebra $\mathfrak{h}$. That is, we simply pick

$$W(\boldsymbol{\alpha}) = \prod_\nu e^{i\alpha_\nu B_\nu}, \tag{40}$$

where iB_ν belongs to a basis of $\mathfrak{k}$, and

$$D(\boldsymbol{\beta}) = \sum_\mu \beta_\mu h_\mu, \tag{41}$$

with ih_μ belonging to a basis of $\mathfrak{h}$. Taken together, Eqs. (40) and (41) provide a problem-inspired ansatz for the diagonalization of H, which we can use to solve Eq. (37).

1. Example: XY model

Let us here exemplify the Cartan decomposition-based method. Consider a general XY Hamiltonian of the form

$$H = \sum_{j=1}^{n-1} a_j^x X_j X_{j+1} + a_j^y Y_j Y_{j+1} + \sum_j a_j^z Z_j. \tag{42}$$

Here, one can prove that $\mathfrak{g} = \langle i\{X_jX_{j+1}, Y_jY_{j+1}, Z_j\}\rangle_{\rm Lie} \cong \mathfrak{so}(2n)$. Thus, the ensuing Cartan decomposition is

$$\mathfrak{k} = {\rm span}\{\widehat{X_iY_j}, \widehat{Y_iX_j} \mid 1 \leqslant i < j \leqslant n\} \cong \mathfrak{so}(n) \oplus \mathfrak{so}(n), \tag{43}$$

$$\mathfrak{m} = {\rm span}\{Z_j, \widehat{X_iX_j}, \widehat{Y_iY_j} \mid 1 \leqslant i < j \leqslant n\}, \tag{44}$$

$$\mathfrak{h} = {\rm span}\{Z_j \mid 1 \leqslant i \leqslant n\} \cong \oplus_{i=1}^n \mathfrak{u}(1). \tag{45}$$

where we used the notation $\widehat{A_iB_j} = A_iZ_{i+1}\cdots Z_{j-1}B_j$. Let us mention that we can map all the elements in $\mathfrak{g}$ to operators quadratic in fermionic operators. In fact, all the gates obtained via exponentiation of this algebra are matchgates [102–104].

Here, we can see that the ansatz for the diagonal part of the ansatz in Eq. (35) is simply

$$D(\boldsymbol{\beta}) = \sum_{j=1}^n \beta_\mu Z_j. \tag{46}$$

On the other hand, it is clear from Eqs. (40) and (43) that a drawback in the proposal of Ref. [49] is that it requires us to implement gates that are obtained by exponentiation of highly nonlocal operators (e.g., $e^{-i\alpha\widehat{X_1Y_n}}$), a task that can be hard to implement and lead to deep circuits. Hence, we propose a different parametrization for $W(\boldsymbol{\alpha})$. Consider the following set of local operators:

$$\mathcal{G} = \{X_jY_{j+1}, Y_jX_{j+1}\}_{j=1}^{n-1} \subset \mathfrak{k}. \tag{47}$$

We can prove that the following proposition holds.

Proposition 4. $i\mathcal{G}$ is a generating set of the algebra $\mathfrak{k}$. That is, $\langle i\mathcal{G}\rangle_{Lie} = \mathfrak{k}$.

The key implication of Proposition 4 is that we can generate any unitary in the unitary subgroup $e^{\mathfrak{k}}$ by exponentiating only the local operators in $\mathcal{G}$. This means, that one can diagonalize H using an ansatz of the form

$$W(\boldsymbol{\alpha}) = \prod_{l=1}^{L}\prod_{j=1}^{n-1} e^{i\alpha_{j,l}X_jY_{j+1}} \prod_{i=1}^{n-1} e^{i\alpha_{i,l}Y_iX_{i+1}}. \tag{48}$$

We explicitly show the form of this ansatz in Fig. 12. Note that since the operators in Eq. (48) are two-body, then the circuit for $W(\boldsymbol{\alpha})$ only requires local two-qubit gates. Hence, such a construction significantly reduces the circuit requirements over that in Ref. [49].

The question still remains as to how large L needs to be. Here, we can leverage recent results from the quantum machine learning literature that state that by taking $L \geqslant \dim(\mathfrak{k})/(2n-2)$ it is generally sufficient to guarantee that any $K \in e^{i\mathfrak{k}}$ will be expressible [97]. Moreover, in this regime the ansatz is said to be overparametrized. In this overparametrization regime, the optimization of Eq. (37) becomes much easier to solve as many spurious local minima disappear [97,105].

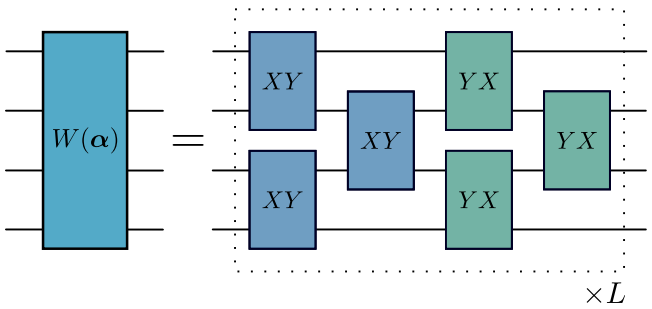


FIG. 12. Ansatz for $W(\boldsymbol{\alpha})$. By using Proposition 4, we propose an ansatz for the diagonalizing unitary $W(\boldsymbol{\alpha})$, which only uses local gates acting on neighboring qubits. We shown here a single "layer" of an $n = 4$ ansatz that is repeated L times.

Putting the previous results together and assuming we can efficiently solve Eq. (37), we can derive the following theorem.

Theorem 7. Let H be an XY Hamiltonian of the form in Eq. (42). Then, let $D(\boldsymbol{\beta})$ be a diagonal operator as in Eq. (46), and let $W(\boldsymbol{\alpha})$ be a unitary as in Eq. (48). By replacing the history state preparation subroutine with the trained diagonalized Hamiltonian as in Eq. (35) and Fig. 11, we can implement the circuits used in Theorems 1, 2, and 4 with circuit depths in $\mathcal{O}(\log(N)n)$.

The results in Theorem 7 showcase the extreme power of diagonalizing H via its Cartan decomposition as we can implement all the circuits in Figs. 5 and 7–9 with a depth that only scales as the product of the number of system and clock qubits.

Let us remark that one key aspect of our example is the polynomial size of the Lie algebra $\mathfrak{g}$, which allows us to efficiently diagonalize the considered Hamiltonian. More complicated Hamiltonians might lead to exponentially large algebras that for deep circuits (as those needed in the overparametrized regime) naturally induce BPs. We refer the reader to Refs. [95,104,106] for a recent discussion of these types of challenges faced by quantum variational protocols. Importantly, for the purpose of our present work, the use of VHD is not central, as the idea of parallelization and the benefits offered by using our proposed methods do not rely on solving a variational optimization problem. As previously mentioned, one can use parallel-in-time techniques with Trotterization, VHD, or any compilation method available.

VI. NUMERICAL SIMULATIONS

In this section, we first provide numerical simulations that showcase how the discrete-time approximations (computable via our algorithms) can capture the behavior of their continuum time counterparts. Similarly, we also show numerically that the system-time entanglement provides a alternative way to understand dynamical properties of the system. Next, we will demonstrate how the variational Hamiltonian diagonalization algorithms can be used to reduce the depth of the history state preparation circuit, as discussed in Sec. V B 1.

In all of our experiments, we consider a system of n qubits evolving by a unitary generated by the time-independent nonuniform XXZ Heisenberg model, whose Hamiltonian reads

$$H = \sum_{j=1}^{n} \left[\frac{J}{4}(X_jX_{j+1} + Y_jY_{j+1}) + \Delta Z_jZ_{j+1} + h_j(Z_j + \mathbb{1}) \right], \tag{49}$$

where we define either periodic boundary conditions as $h_{n+1} \equiv h_1$ (for $h = X, Y, Z$) or open boundary conditions $h_{n+1} \equiv 0$ and $h_j = \frac{\lambda}{2}\cos(2\pi\alpha j)$. For $\alpha = \frac{\sqrt{5}-1}{2}$ and $\Delta = 0$, one can use the Jordan-Wigner (JW) transformation [107] (see details in Appendix G) to show that in the thermodynamic limit this model exhibits a delocalization-localization transition at the critical point $\lambda = J$. Indeed, it is well known that such a transition induces sharp changes in long-time dynamical properties such as the Loschmidt echo average [65]. Our goal is then to use this paradigmatic model as a test bed to show that our proposed discrete-time average of the Loschmidt echo can capture the behavior of their continuum time counterparts. Analogously, for $\Delta \neq 0$ it is expected that this model exhibits a many-body localization transition with localization surviving the presence of interactions [57,108].

A. Single-particle localization, discrete-time averages, and system-time entanglement

Let us focus on the $\Delta = 0$ case first. To study the discrete-time average of the Loschmidt's echo, we have considered a chain of $n = 200$ sites with $J = 2$, $\alpha = \frac{\sqrt{5}-1}{2}$, and a number of clock qubits ranging from 1 to 10, corresponding to a maximum number of $N = 1024$ times. Note that with this choice, the system dimension is equal to 2^{200}, and hence much larger than the clock Hilbert space dimension, N. To study the effects of the window size, we have also considered values of ε spanning from 0.05 up to 1.95 with spacing 0.1 (see Fig. 3). The initial state of our simulations is $|\psi_0\rangle = (s^+_{99} + s^+_{100} + s^+_{101})|\downarrow\downarrow\cdots\downarrow\rangle/\sqrt{3}$, where s^+_j denotes the rising spin operator at site j. As such, at $t = 0$, the state is only partially delocalized in the middle of the chain. All simulations, including the computation of the exact infinite-time average $\bar{\mathcal{L}}(\psi_0)$, were performed via Jordan-Wigner diagonalization, and we refer the reader to Appendix G for additional details.

In Fig. 13, we first present a two-dimensional plot of the error between the infinite-time average of the Loschmidt echo $\bar{\mathcal{L}}(\psi_0)$ and its discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$ ($|\bar{\mathcal{L}}(\psi_0) - \widetilde{\mathcal{L}}(\psi_0)|$) averaged over $\lambda \in (0.1, 3.5)$ (with spacing $\Delta\lambda = 0.05$), for different values of ε and n. Here, we can see that, as expected, the error is reduced by increasing the number of clock qubits. The improvement follows two tendencies. First, there is an overall improvement when increasing T (i.e., when moving up in the $\log(N)$ axis for fixed ε), as this corresponds to better accuracy. On the other hand, for constant T it is beneficial to reduce ε (i.e., increase resolution), as shown by the blue solid curves.

We further explore the effect of fixing ε and increasing $\log(N)$ in Fig. 14(a). Therein, we show $\bar{\mathcal{L}}(\psi_0)$, as well as its discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$, for different number of clock qubits as a function of λ for fixed resolution $\varepsilon = 0.45$ (vertical dashed line in Fig. 13). First, we note that the infinite-time Loschmidt echo captures the delocalization-localization transition occurring at $\lambda = J = 2$. In particular, for $\lambda < 2$ we see that $\bar{\mathcal{L}}(\psi_0)$ is small, indicating a delocalized phase. On the

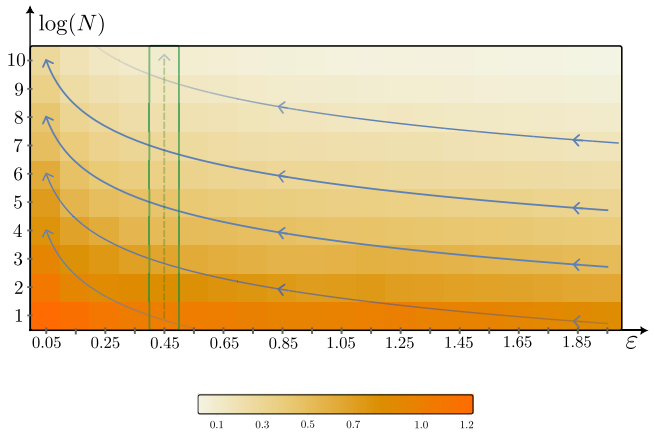


FIG. 13. Average approximation error $|\bar{\mathcal{L}}(\psi_0) - \widetilde{\mathcal{L}}(\psi_0)|$ for a chain of $n = 200$ sites. The vertical axis indicates the number of clock qubits, while the horizontal axis is the window size ε. The blue (solid) curves represent constant values of T with the arrows pointing toward the direction of greater resolution. Green arrows indicate a cross cut of fixed ε and increasing $\log(N)$ that is shown in further detail in Fig. 14.

other hand, for $\lambda > 2$ the evolved state is localized as $\bar{\mathcal{L}}(\psi_0)$ is large. Next, let us note that as $\log(N)$ increases, $\widetilde{\mathcal{L}}(\psi_0)$ quickly becomes a good approximation for its infinite-time counterpart (as expected from Fig. 13). However, Fig. 14 also reveals that $\widetilde{\mathcal{L}}(\psi_0)$ capture the delocalization-localization transition even for a small number of clock qubits. Already for $\log(N) = 6$, the inflection point of $\widetilde{\mathcal{L}}(\psi_0)$ approaches the critical value $\lambda = 2$.

Next, we study how the system-time entanglement, as measured through the subsystem purity $\mathrm{Tr}[\rho_S^2]$ for $\rho_S = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|]$, approximates the infinite-time average Loschmidt echo (see Corollary 1). In Fig. 14(b), we plot $\bar{\mathcal{L}}(\psi_0)$, as well as $\mathrm{Tr}[\rho_S^2]$, for different number of clock qubits as a function of λ. Again, we see a clear convergence toward $\bar{\mathcal{L}}(\psi_0)$ as the number of clock qubits is increased. This result shows that the subsystem purity provides an excellent approximation of $\bar{\mathcal{L}}(\psi_0)$. Moreover, one can also observe that the system-time entanglement clearly captures the delocalization-localization transition. This fact can be readily understood from the fact that in the localized phase the state does not change considerably with time, and hence a small amount of entanglement is expected. This example perfectly exemplifies the fact that the system-time entanglement in the history state carries valuable information about the system dynamics. Moreover, since we know that $\mathrm{Tr}[\rho_S^2] = \mathrm{Tr}[\rho_T^2]$, then one can estimate the reduced state tomography by studying only the reduced state on the $\log(N)(\ll n)$ clock qubits.

Figures 14(a) and 14(b) show that both the discrete-time Loschmidt echo and the subsystem purity provide good approximations of $\bar{\mathcal{L}}(\psi_0)$. To better compare their performances, we show in Fig. 14(c) curves for $\bar{\mathcal{L}}(\psi_0)$, $\widetilde{\mathcal{L}}(\psi_0)$, and $\mathrm{Tr}[\rho_S^2]$ for the same chain of $n = 200$ spins, but for $\varepsilon = 1.25$, i.e., for less accuracy (see Fig. 3). In this regime, one can see that while $\widetilde{\mathcal{L}}(\psi_0)$ suffers from undesired oscillations, $\mathrm{Tr}[\rho_S^2]$ can still provide a good approximation for the same number of qubits. In particular, Fig. 14(c) shows that $\widetilde{\mathcal{L}}(\psi_0)$ can be smaller than $\bar{\mathcal{L}}(\psi_0)$ in unpredictable ways (due to insufficient resolution), meaning that $\widetilde{\mathcal{L}}$ cannot be strictly used to provide strict bounds such as the one in Corollary 1. While $\mathrm{Tr}[\rho_S^2]$ oscillates as well, this quantity never crosses the black points, in agreement with our bounds. Here, we also observe that the system-time entanglement provides a better convergence in the localized region. On the other hand, the entanglement curves are above the $\widetilde{\mathcal{L}}$ curves in the delocalized sector. Notice, however, that this discrepancy can be mitigated by increasing the number of qubits. Finally, we note that in Fig. 14(c) we also depict the differences $(\widetilde{\mathcal{L}}(\psi_0) - \bar{\mathcal{L}}(\psi_0))$ and $(\mathrm{Tr}[\rho_S^2] - \bar{\mathcal{L}}(\psi_0))$, which confirm that $\mathrm{Tr}[\rho_S^2]$ is always strictly larger than $\bar{\mathcal{L}}(\psi_0)$, whereas $\widetilde{\mathcal{L}}(\psi_0)$ can indeed be smaller than the infinite-time average.

Finally, as an example of Corollary 3 we also numerically show how the system-time entanglement provides a bound for the fluctuation of observables. We use as an example the observable $O = s^+_{L/2}s^-_{L/2+1} + s^+_{L/2+1}s^-_{L/2}$ and as the initial state $|\psi_0\rangle = (s^+_{L/2} + s^+_{L/2+1})|\downarrow\downarrow\cdots\downarrow\rangle/\sqrt{2}$. In this case, the bounds of Eq. (27) become

$$\sigma_O^2 \leqslant 4\bar{\mathcal{L}}(\psi_0) \leqslant 4\mathrm{Tr}\big[\rho_S^2\big] \tag{50}$$

since $\Delta_O = 2$. In Fig. 15, we plot the numerical results for a chain of $n = 100$ sites and $\log(N) = 9$ qubit clocks (i.e.,

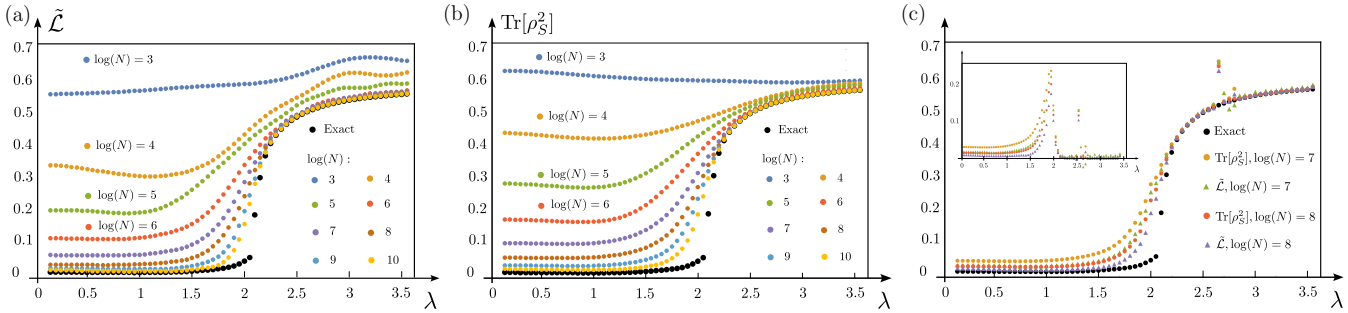


FIG. 14. Discrete-time approximations to the Loschmidt echo $\bar{\mathcal{L}}(\psi_0)$ (black dots) as a function of λ for a chain of $n = 200$ sites. In panel (a), we depict the exact infinite-time average, as well as the discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$ for different values of clock qubits $\log(N)$ and for $\varepsilon = 0.45$. In panel (b), we depict the exact infinite-time average $\bar{\mathcal{L}}(\psi_0)$ (black dots), as well as the purity of the reduced state $\rho_S = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|]$ for different values of clock qubits $\log(N)$ and for $\varepsilon = 0.45$. In panel (c), we depict the exact infinite-time average $\bar{\mathcal{L}}(\psi_0)$ (black dots), its discrete-time approximation $\widetilde{\mathcal{L}}(\psi_0)$, and the purity of the reduced state $\rho_S = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|]$ for different values of clock qubits $\log(N)$ and for $\varepsilon = 1.25$. The inset corresponds to the difference between each approximation and $\bar{\mathcal{L}}(\psi_0)$.

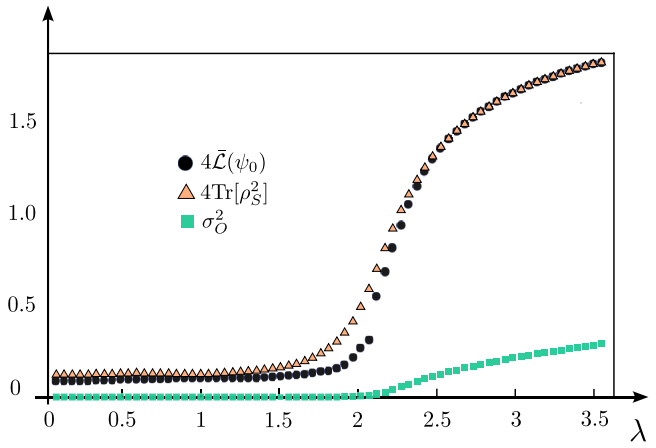


FIG. 15. Observable fluctuations as a function of λ for a chain of $n = 100$ sites. We depict the observable fluctuations σ_O^2, the infinity-time average of the Loschmidt echo $\bar{\mathcal{L}}(\psi_0)$, and the reduced subsystem purity $\mathrm{Tr}[\rho_S^2]$. We consider $\log(N) = 9$ qubit clocks, and we take $\varepsilon = 0.5$

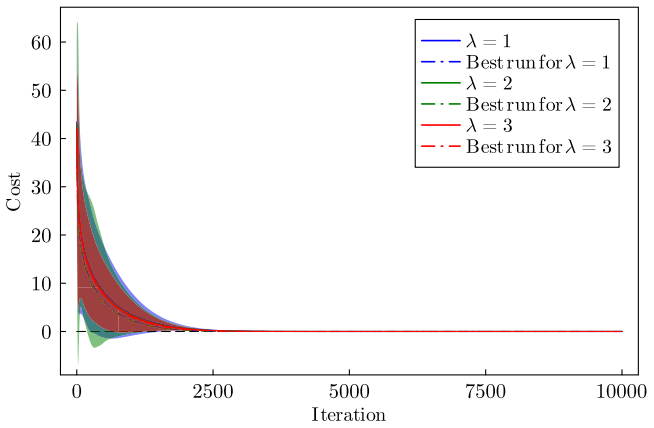


FIG. 16. Variational Hamiltonian diagonalization training curves. The plot shows the decay of the loss function in Eq. (36) versus the number of training iterations. Solid lines correspond to the average over 10 different initializations for each value of $\lambda \in \{1, 2, 3\}$. Shaded regions indicate the variance over the run sample, while dash-dot lines show the best run for each batch of training runs, i.e., the one that reached the stopping value first.

512 times). We see that while the bound is not tight, both $\bar{\mathcal{L}}(\psi_0)$ and $\mathrm{Tr}[\rho_S^2]$ are capable of clearly separating the different phases. As expected from our bounds, the system-time entanglement provides a less tight but strict bound. However, given that one can experimentally compute the system-time entanglement efficiently in quantum computers, this bound is still useful for practical purposes. Moreover, it is important to highlight again the fact that the system-time entanglement is obtained from a discrete-time formalism [in contrast to $\bar{\mathcal{L}}(\psi_0)$ that requires infinite-time averages]. As such, our notion of system-time entanglement provides valuable and strict information about the system's observable dynamics and its eventual equilibration [a feature not available for the discrete-time Loschmidt echo $\tilde{\mathcal{L}}(\psi_0)$].

B. Diagonalization via Cartan decomposition

In this section, we show how one can use the variational Hamiltonian diagonalization [to diagonalize the Hamiltonian in Eq. (49) and thus reduce the depth of the history state preparation circuit]. We will take the ansatz for $D(\boldsymbol{\beta})$ and $W(\boldsymbol{\alpha})$ as appearing in Eqs. (46) and (48). Thus, as depicted in Fig. 12, $W(\boldsymbol{\alpha})$ consists of L layers of two-qubit gates generate by XY and YX arranged in a brick wall fashion, whereas the diagonal part $D(\boldsymbol{\beta})$ is just a sum of Pauli Z operators on each qubit. To train the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we will optimize the Hilbert-Schmidt cost function defined in Eq. (36). Details of the simulation can be found in Appendix G.

We considered a chain of $n = 6$ sites, and set $J = 2$ and $\alpha = \frac{\sqrt{5}-1}{2}$. Moreover, we diagonalized the Hamiltonian for $\lambda = 1, 2, 3$, thus allowing us to show the success of the algorithm in each important region of the phase diagram.

In Fig. 16, we show the training curves (loss function versus iteration step) for an ansatz with $L = 18$ layers the three different values of λ that we considered. Here, we can clearly see that as the number of iterations increases, the cost function value goes to zero, indicating that we can accurately diagonalize the target Hamiltonian. Notably, we can see that all the trained curves converged to the solution, meaning that the optimizer did not get stuck in a local minimum. Such extremely high optimization success rate can be understood from the fact that the circuit is overparametrized [109]; i.e., it contains enough parameters to explore all relevant directions. In fact, using the results from Ref. [109], we know that a circuit with a set of generators $\mathcal{G}$ will be overparametrized if the number of parameters is $\dim(\langle i\mathcal{G}\rangle_{\mathrm{Lie}})$. Importantly, we can use Proposition 4 to know that $\dim(\langle i\mathcal{G}\rangle_{\mathrm{Lie}}) = n(n-1) = 30$. Since the ansatz $W(\boldsymbol{\alpha})$ contains $2(n-1)$ parameters per layer (see Fig. 12), then we can overparametrize it with $L = \lceil n/2 \rceil$. Indeed, in Fig. 17 we show the minimal cost achieved versus the number of layers L, and as expected we see a computational phase transition at $L = 3$: For smaller number of layers, the ansatz is underparametrized and can get stuck in local minima, but for $L \geqslant 3$ it is overparametrized and training becomes easier. As the plot shows, once the model is overparametrized, further increasing the number of layers does not lead to any improvement in the minimum loss value achievable. Finally, in Fig. 18 we show the success of the diagonalization by applying a successfully trained diagonalizing unitary $W(\boldsymbol{\alpha})$ to its target Hamiltonian H, obtaining a perfectly diagonal matrix and verifying the success of the algorithm.

C. Many-body localization and system-time entanglement

Consider now the full $\Delta \neq 0$ model in an open chain. As shown in the Appendixes, the Hamiltonian is essentially mapped to a Hubbard-like interacting model via the JW transformation, with Δ the interacting strength. It has been argued [108] that the localization phenomenon in this system could survive the interaction, at least for small Δ, with the transition point shifting to higher values of the critical λ as Δ increases (see the proposed phase diagram in Ref. [108]).

In this section, we numerically show how our parallel-in-time algorithms can be used to capture this behavior. We considered a chain of $n = 54$ spins and an initial state

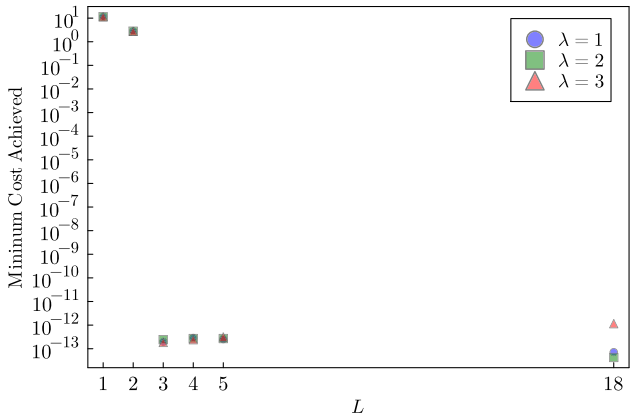


FIG. 17. Overparametrization transition. We show the minimum cost achieved when training the variational Hamiltonian diagonalization algorithm for a problem system of $n = 6$ sites. The overparametrization regime is reached with a number of layers $L \geqslant 3$.

$|\psi_0\rangle = s^+_{n/2} s^+_{n/2+1} |\downarrow\downarrow \cdots \downarrow\rangle$, corresponding to two particles localized in the middle of the chain. In Fig. 19(a), we show a contour plot of the time average of the Loschmidt's echo for $N = 2^{15}$ corresponding to 15 clock qubits, and with $\epsilon = 0.05$. One can qualitatively appreciate the tendency of the localization sector λ moving to higher values of λ as the interaction strength increases. Similarly, the system-time entanglement, measured by the purity $\mathrm{Tr}[\rho_S^2]$, is depicted in Fig. 19(b), capturing the same behavior. Let us notice that many "fast oscillations" are present. While these may be smoothed out by considering averages over an additional offset added to the argument of the cosine function defining h_j and over the initial state [108], a single run of our algorithms would correspond to have these quantities as fixed, as in our example. At the same time, these oscillations are clearly present, as shown in Fig. 19(c), through the exact infinite-time average $\bar{\mathcal{L}}(\psi_0)$. As shown therein, in order to fully capture them via entanglement, a sufficiently large number of click qubits is required, with 18 clock qubits already capturing most of this subtle behavior. We remark also how the bound of Theorem 1

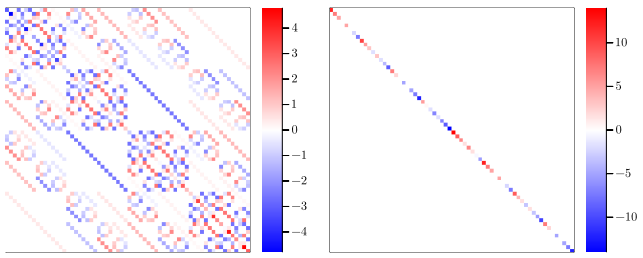


FIG. 18. Testing the success of the Hamiltonian diagonalization algorithm. The left panel shows the initial "diagonalized" Hamiltonian $\hat{D}(\boldsymbol{\alpha}) = W(\boldsymbol{\alpha}) H W^\dagger(\boldsymbol{\alpha})$ where $\boldsymbol{\alpha}$ is a vector of initial random parameters $\alpha_i \in [0, 2\pi)$. The right panel instead features $\hat{D}(\boldsymbol{\alpha}^*)$ obtained by applying the trained diagonalizing unitary to its target Hamiltonian H. Here, $\boldsymbol{\alpha}^*$ are the optimal parameters of the random successful run at hand.

is strictly satisfied, with the entanglement approaching the "dips" always from above as N increases.

Let us now briefly discuss our choice of the initial state. Notice that we are considering "low-energy" states. This allows us to consider large temporal averages and modestly large chains (see details in the Appendix C for a description of our methods). The more usual assumption made in the literature [57,108] of an initial half-filled state renders the time averages we considered practically inaccessible for classical computations except for small chains. At the same time, even for this simple state, the system exhibits an interesting behavior as shown in Fig. 19, thus providing us with a nontrivial example of the ability of our algorithms to capture many-body properties, even for a relatively small number of clock qubits. In Appendix G, we discuss how to simulate the interacting model both via exact diagonalization after mapping it to a fermionic model and via standard tensor network (TN) techniques. Particularly, since the system at hand is one dimensional, in the latter case we resort to matrix product state (MPS) algorithms [110], for which we provide scaling of complexity and deviation from the exact results. The analysis of bond dimension confirms the nontrivial behavior of the system for our choice of initial state. The MPS perspective also allows us to discuss the effects of Trotter error (see Appendix G for details).

VII. DISCUSSIONS

The simulation of quantum systems has widely been considered the most important application of quantum computation since its conception [111]. Traditionally, the focus of quantum simulations has revolved around computing quantum states and physical quantities at a given time, harnessing the exponential growth of the Hilbert space of qubits to mimic the behavior of many-body systems. However, many fundamental quantities, such as correlation functions or the equilibrium state of a quantum system, are associated with large temporal sums of the previous. In this manuscript, we have shown that by treating time itself quantum mechanically, which in a computational scheme corresponds to using clock qubits, those quantities become readily accessible.

This result arises from a fruitful analogy between the recent quantum time discussions in quantum foundations and quantum gravity fields, and the fields of quantum information and computation. More importantly, by developing quantum time inspired algorithms, we have disclosed important connections between the correlations contained in history states and the problem of equilibration of an isolated quantum system, thus unveiling a link to statistical mechanics as well. In particular, we have shown that the system-time entanglement is a good measure of equilibration, and as a by-product, how the formalism provides a way to prepare approximate equilibrium states. Whether under proper conditions this can provide a useful scheme for studying thermalization as well is left for future investigations. These considerations show that, in addition to the practical applications of the various proposed algorithms, the framework we presented offers insights that can be applied to diverse areas of many-body physics and potentially to other scenarios where one needs to simulate nonunitary quantum processes (e.g., via linear combination of unitaries).

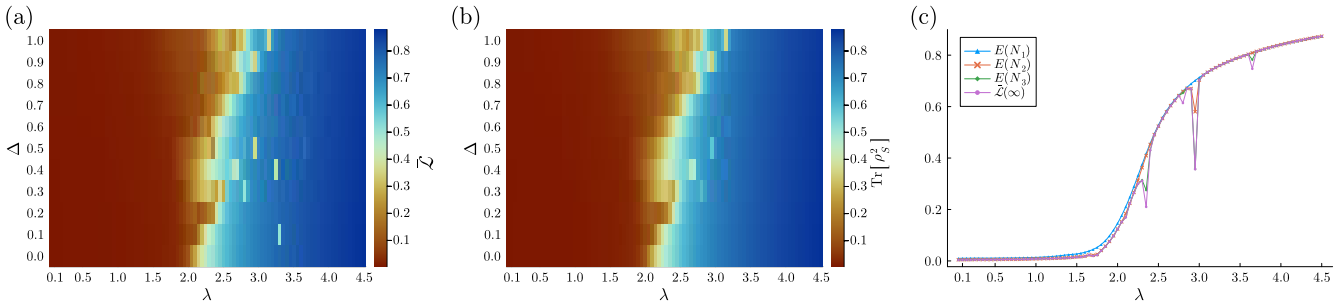


FIG. 19. System-time entanglement and Loschmidt echo for Hubbard-like interacting model. In all three panels, we consider $|\psi_0\rangle = s^+_{n/2}s^+_{n/2+1}|\downarrow\downarrow\cdots\downarrow\rangle$ as the initial state of a chain of $n = 54$ spins. (a) Contour plot of the discrete-time approximation of $\bar{\mathcal{L}}$ as a function of the "disorder" strength $\lambda \in (0.05, 4.5)$ with spacing 0.05 and the interaction coupling $\Delta \in (0, 1)$ with spacing 0.1. Here, the number of clock qubits is 15 ($N = 2^{15}$), while the time step is taken as $\epsilon = 0.05$. (b) Contour plot of the discrete-time approximation of $\mathrm{Tr}[\rho_S^2]$ under the same conditions of panel (a). (c) System-time entanglement ($E = \mathrm{Tr}[\rho_S^2]$) vs exact Loschmidt echo $[\bar{\mathcal{L}}(\infty)]$ for fixed $\Delta = 0.05$, $\epsilon = 0.05$, and time registers of sizes $N_1 = 2^{12}$, $N_2 = 2^{15}$, $N_3 = 2^{18}$.

A key advantage of our framework is that whatever compilation method one would use to transform e^{-iHt} into a sequence of gates for sequential deployment can also be readily adapted to parallel deployment. While here we discuss Trotterization approaches and the VHD algorithm, we note that parallel-in-time can be also combined with any form of state-of-the for compilation and compression. For instance, one could instead use tensor-network compilation techniques [112–114] based on out-of-distribution generalization [115,116] and further reduce the depth of the circuit via parallelization. Such improvement, while quadratic in the circuit depth, should not be readily disregarded as quantum circuit runs are expected to be (1) expensive and (2) slow in the foreseeable future, meaning that the reduction in the number of circuit runs our techniques offers could lead to significant financial and time savings.

Next, let us note that the system-time entanglement is also indicative of the need for entangling gates to prepare history states. In this work, we have considered a fixed architecture that even for a direct Trotterization approach provides a running time advantage over sequential-in-time schemes. We have also considered a further simplification via a variational Hamiltonian diagonalization approach. Interestingly, the fact that the entanglement is state and evolution dependent suggests that one can consider more adaptive preparation schemes, where the entanglement is used as a measure of "compressibility" (or complexity) of the history state. While we leave the study of these possibilities for future investigations, we should mention that a global variational scheme has been proposed recently [32] based on the Feynman-Kitaev Hamiltonian [6]. As with any variational protocol, knowledge of the structure of the solution is valuable for choosing proper ansätze, thus rendering our efficient protocol particularly relevant for any near-term implementation as well. At the same time, all of our protocols can be easily updated by replacing the history state preparation subroutine. The disclosed parallel-in-time advantages clearly hold and can be readily extended to any related proposal including [32].

In all the protocols we have considered, the circuits give information about the history of a closed system. However, these protocols can be extended to consider the history of states that at some points in time are being subject to measurements. This becomes feasible by following the recent treatment of Ref. [7] that incorporates ancillary memories (following von Neumann) and describes the history of the whole (history of the system + ancillas). This framework opens many additional interesting possibilities for quantum algorithms. A similar treatment may be applied to quantum evolutions associated with nonunitary channels (open systems).

On the other hand, even if we focus on closed systems there is much yet to explore: Most of the parallel protocols in the paper involve measurements on the system side. We also know that adding a projective measurement in the time basis of the clock qubits yields predictions at a given time. However, the most characteristic features of quantum mechanics arise when one is considering measurements in different bases, meaning that the full potential of quantum time effects remains unexplored. In addition, multiple copies of history states may be used to study higher momenta of mean values, thus opening even more possibilities. Protocols exploiting time-reversed evolution [117,118] might also be considered. Going even further, it has been recently discussed [11–14,33,119,120] that the PaW mechanism alone is not enough for achieving a fully symmetric version of QM. This has led to nontrivial extensions of the original formalism of Page and Wootters (see also Appendix A). We can speculate that in the near future, motivated by the current results, these extensions may provide further informational and computational insights related to the time domain.

ACKNOWLEDGMENTS

The authors would like to thank Luis Pedro Garcia Pintos, Lodovico Scarpa, Lucia Vilchez, Ivan Chernyshev, and Hasan Sayginel for fruitful discussions. N.L.D, P.B., and M.C. were supported by the Laboratory Directed Research and Development (LDRD) program of Los Alamos National Laboratory (LANL) under Projects No. 20210116DR, No. 20230049DR, and No. 20230527ECR, respectively. P.B. was also supported by the U.S. Department of Energy (DOE) through a quantum computing program sponsored by the LANL Information Science & Technology Institute. N.L.D. was also supported by the Center for Nonlinear Studies (CNLS) at LANL. M.L. was initially supported by the Center for Nonlinear Studies

at LANL and by the U.S. DOE, Office of Science, Office of Advanced Scientific Computing Research, under the Accelerated Research in Quantum Computing (ARQC) program. M.C. was also supported by LANL's ASC Beyond Moore's Law project. N.L.D. and J.M.M acknowledge support from CONICET, and R.R. from CIC Argentina. This research used resources provided by the Los Alamos National Laboratory Institutional Computing Program and by the Darwin test bed at Los Alamos National Laboratory (LANL) that is funded by the Computational Systems and Software Environments subprogram of LANL's Advanced Simulation and Computing program (NNSA/DOE).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIXES

Here, we present derivations, additional details for the main results, and further general discussions on quantum time notions. The Appendixes are organized as follows: In Appendix A, we give an overview of quantum time approaches and describe connections between our work and the literature. In Appendix B, we explain in more detail the relation between quantum time and the equilibration problem and prove the related Theorems in the main body. In Appendix C, we give detailed proofs of all the different circuits we proposed in the main body. In Appendix D, we prove Theorem 5 about circuit depths in the sequential- and parallel-in-time schemes. In Appendix E, we prove Proposition 4 about the algebra of the XY model. In Appendix F, we prove Theorem 7 regarding the depth of the total protocols involving both Hamiltonian diagonalization and history states. Finally, in Appendix G we provide details about the methods employed in the numerics (Sec. VI).

APPENDIX A: CONNECTION TO LITERATURE AND OVERVIEW OF QUANTUM TIME APPROACHES

Here, we briefly discuss the broader conceptual picture of "quantum time-related proposals." This should also clarify the range of applicability of our current ideas. Regarding the Page and Wootters approach [40], it is worth noting that their original idea was to replace dynamics with quantum correlations. The physical picture was a static universe from which evolution *emerges* from a convenient separation between the system and the rest. In this sense the "time" Hilbert space corresponds to a "cosmological" clock. One main motivation for these ideas was the discussion about time in quantum gravity [38], where the Wheeler-deWitt equation [123] suggests a static universe. Such an equation appears as a constraint induced by the "gauge" freedom in the choice of coordinates in general relativity and can be understood within the framework of Dirac's generalization of Hamiltonian dynamics to constrained systems [124,125].

However, there are important conceptual and mathematical differences between the PaW's and Dirac's approaches, which are at the core of our ability to leverage the first and not the second to develop useful computational schemes. To clarify this statement, let us briefly discuss a simple but archetypal [126] application of Dirac's approach: In classical mechanics, one can accommodate time itself and its conjugate variable p_t in an extended phase space by introducing a variable τ that parametrizes phase-space variables, including $t = t(\tau)$ and $p_t = p_t(\tau)$. The Poisson brackets are also extended by imposing $\{t, p_t\} = 1$ with $p_t \neq H$ that for a single particle (sp) and $x^0 \equiv t$, $p_t \equiv p_0$ completes the algebra $\{x^\mu, p_\nu\} = \delta^\mu_\nu$, which now is explicitly covariant. Part of the quantization scheme now consists of the replacement

$$\{x^\mu, p_\nu\} = \delta^\mu_\nu \to [x^\mu, p_\nu] = i\hbar\delta^\mu_\nu, \tag{A1}$$

which for $\mu = \nu = 0$ is just $[T, P_T] = i\hbar$ (when the system is a particle), i.e., $x^0 \equiv T$ with $[T, H] = 0$ for H some Hamiltonian of the particle. On the other hand, the independence of physical quantities on the way the τ parametrization is chosen leads to $\mathcal{J}|\Psi\rangle = 0$ after quantization. This reparametrization invariance is analogous to the general covariance of general relativity, and the constraint is analogous to the Wheeeler-deWitt equation.

However, contrary to our main body discussion, in Dirac's approach the constraint equation defines the so-called physical Hilbert space that is regarded as different from the previous "kinematic" Hilbert space. Moreover, the physical Hilbert space is not treated as a subspace of the latter, instead the kinematical one appears only as an auxiliary step in the whole quantization process but not in the final construction. As a consequence, the time operator is ruled out in the final formalism for not being a physical observable (see, e.g., Ref. [127] for discussions about this procedure and the related "Hilbert space problem"). As we have shown in the main body, in the PaW approach the interpretation and treatment of the constraint are completely different: The time operator is not disregarded since it corresponds to an actual observable of the clock system. The complete Hilbert space, e.g., as defined by Eq. (A1) is preserved, while the state of the system is recovered by conditioning. The theoretical ideas underlying our manuscript clearly exploit the "extended" Hilbert space associated with the PaW approach and are mostly inspired by the recent developments in the context of quantum information [7,8,11,12], and in particular in the discrete-time formulation [8], which was further extended in Ref. [9].

Let us also stress that for the purposes of our manuscript it is not relevant whether explicit covariance is achieved via the definition of a time operator. This means that we can consider arbitrary many-body systems as we have shown in the main body. This also includes relativistic systems (properly discretized quantum field theories) but their simulation corresponds to the evolution as seen from a fixed reference frame: In this case, there is no simple rule to relate a history state in a given reference frame to another. In this sense, we regard the PaW formalism as nonexplicitly relativistic. There is, however, a simple situation where such a rule is straightforward: For a single relativistic particle, Lorentz transformations can be introduced explicitly and geometrically (as space-time "rotations" independently of the theory) [11,12]. For this to be actually achieved, one needs to preserve the kinematical space and deal with an extended inner product, a result that is again highlighting the advantages of the extended scheme. This approach was recently developed in Ref. [11] for Dirac's

particles and in Ref. [12] for scalar particles where the inner products were successfully related to the "physical" ones while preserving the extended Hilbert spaces. As a consequence, the associated history states can still be realized by the means presented in this manuscript (adopting some discretization of space-time), and in principle, Lorentz transformations may be introduced as nonlocal gates acting on both the system and time qubits.

Another interesting feature of these single-particle quantum time formalisms (which preserve the extended Hilbert space) is that their "second quantization", as introduced in Refs. [12–14] naturally leads to an approach to many-body scenarios and to quantum field theories in which space and time are *explicitly* on equal footing (see also the closely related "quantum mechanics of events" proposal [33]). In particular, the approach in Ref. [13] also leads to a redefinition of the Path Integral formulation [14]. In fact, the real challenge to treat space and time on equal footing at the Hilbert space level in relativistic settings has to do with another more subtle asymmetry in the treatment of time [13,14,42,128]: joint systems separated in space are described by the tensor product of the corresponding Hilbert spaces. No such rule is applied to time. This asymmetry is particularly evident in the case of quantum field theories, in which case space is treated as a site, which, e.g., for bosons is equivalent to a tensor product structure in space. While time is a parameter, it is not a site index and there is no associated tensor product structure. This is manifest in the *equal time* canonical algebra imposed on the fields which requires a fixed foliation. We notice also that this is an obstacle in defining a notion of "timelike entanglement" (see, however, Refs. [13,14,129–131] for recent related discussions), a consideration which applies to all QM: the previous asymmetry is present in any quantum mechanical system, as also discussed in Refs. [10,13,14,42,44,132].

For these reasons, it is not sufficient to define a quantum time operator to solve "the problem of time" or more precisely to have an explicit space-time symmetric version of QM. This renders the aforementioned second quantization of "PaW particles" particularly interesting and relevant in the context of *quantum field theories*, as shown recently in Refs. [119,120] (for a different extension of the PaW scheme, developed along the lines of parametrized field theories, see Ref. [133]). While the description of the results relating the PaW mechanism with field theories exceeds the purpose of this manuscript, we can speculate that these developments may provide additional quantum computational and informational tools, just as the PaW approach inspired the various algorithms of our manuscript.

APPENDIX B: MORE ON THE RELATION BETWEEN QUANTUM TIME AND THE EQUILIBRATION PROBLEM: PROOF OF THEOREM 3

In this section, we highlight the relations between the current quantum time approach (the PaW-inspired formalism), its associated concept of system-time entanglement, and the problem of equilibration of an isolated quantum system. In particular, we provide a prove and analysis of Theorem 3. Let us mention that this connection has not been explored in the literature (see, however, the recent article [28]).

Let us recall first that a system is said to equilibrate if it evolves toward some particular, in general, mixed state and remains in that state or close to it for almost all times [58]. More precisely, this holds for either a small subsystem of a large quantum system evolving unitarily [58] (subsystem equilibration) or for expectation values associated with "experimentally realistic conditions" [4] (observable equilibration). The last interpretation refers to the fact that the probability that mean values computed with $|\psi(t)\rangle\langle\psi(t)|$ differ from mean values computed with the equilibrium state becomes exponentially vanishing for typical experimentally accessible observables [4,134]. In both scenarios, the dephased state $\bar{\rho}$ defined in Eq. (17) plays a central role since it is identified with the equilibrium state (equivalently, the subsystems equilibrium state is a partial trace of this global state). Moreover, the theoretical bound [4] used to prove this is related to $\bar{\mathcal{L}}$ precisely because $\mathrm{Tr}[\bar{\rho}^2] = \bar{\mathcal{L}}$; i.e., the purity of the time-averaged state is $\bar{\mathcal{L}}$. The same quantity provides a bound for the (time average of) the distance between the subsystem state and (the partial trace of) $\bar{\rho}$ [58].

Now returning to the history state, we have seen that the quadratic entanglement involves the purity of $\rho_S = \mathrm{Tr}_T[|\Psi\rangle\langle\Psi|]$. The explicit expansion of ρ_S is

$$\rho_S = \frac{1}{N}\sum_t |\psi(\varepsilon t)\rangle\langle\psi(\varepsilon t)|. \tag{B1}$$

In order to relate this state with $\bar{\rho}$, we can use the energy basis $|k\rangle$ so that $H|k\rangle = E_k|k\rangle$. Any initial state has an expansion $|\psi_0\rangle = \sum_k c_k|k\rangle$. Notice that even if there is degeneracy we can always write a fixed pure state as before, with $|c_k|^2 := \sum_j |\psi_{kj}|^2$, $|k\rangle := c_k^{-1}\sum_j \psi_{kj}|kj\rangle$, and $H|kj\rangle = E_k|kj\rangle$; i.e., j is a degeneracy index. The quantity $c_k|k\rangle$ is the projection of $|\psi\rangle$ onto the subspace of states with energy E_k. Thus, we obtain

$$\rho_S = \frac{1}{N}\sum_t \sum_{k,k'} c_k c_{k'}^* e^{-i(E_k - E_{k'})t\varepsilon}|k\rangle\langle k'|, \tag{B2}$$

which is precisely the discrete-time version of $\bar{\rho}$ as defined in Eq. (17), i.e.,

$$\rho_S = \tilde{\rho}. \tag{B3}$$

Notice that while the purity of $\bar{\rho}$ coincides with $\bar{\mathcal{L}}$, the purity of ρ_S is not $\tilde{\mathcal{L}}$. Remarkably, the natural discrete-time generalization of $\bar{\mathcal{L}} = \mathrm{Tr}[\bar{\rho}^2]$ is $\mathrm{Tr}[\rho_S^2]$ and not $\tilde{\mathcal{L}}$: The system-time entanglement provides the proper discrete-time version of time average related bounds (it is worth remarking that the *discrete-time entanglement* provides strict bounds to the *infinite and continuum* time averages). This is captured by Theorem 3 and its corollaries.

We are now in a position to give the proof of Theorem 3:

Proof. Let us first rewrite ρ_S as

$$\rho_S = \sum_{k,k'} c_k c_{k'}^* \Delta_{kk'}|k\rangle\langle k'|, \tag{B4}$$

with $\Delta_{kk'} := \frac{1}{N}\sum_t e^{-i(E_k - E_{k'})t\varepsilon}$. Notice that $\Delta_{kk} = 1$. Being a quantum state, we can diagonalize ρ_S in some basis as

$$\rho_S = \sum_l p_l|l\rangle\langle l|. \tag{B5}$$

In fact, it is clear that this is what is obtained by tracing over the clocks with $|\Psi\rangle$ in its Schmidts decomposition (15).

We can combine these two expansions to write

$$\langle k|\rho|k\rangle = |c_k|^2 = \sum_l p_l |\langle l|k\rangle|^2. \tag{B6}$$

The quantity $|\langle l|k\rangle|^2$ defines a double stochastic matrix, thus yielding the desired majorization relation

$$\{|c_k|^2\} \prec \{p_l\}, \tag{B7}$$

thus implying the desired majorization relation between states (we recall that $\bar{\rho} = \sum_k |c_k|^2 |k\rangle\langle k|$). Notice that similar ideas have been used in Ref. [8].

It is also clear that by taking the limits $\varepsilon \to 0$ and then $T \to \infty$, the energy dephasing is recovered: The first limit can be considered by writing

$$\Delta_{kk'} = \frac{1}{\varepsilon N} \sum_t \varepsilon\, e^{-i(E_k - E_{k'})t\varepsilon} \to \int_0^T \frac{dt}{T} e^{-i(E_k - E_{k'})t},$$

with $dt \equiv \varepsilon$. The large T limit now is the familiar limit used conventionally (which yields a delta), thus leading again to the asymptotic relation

$$\rho_S \to \bar{\rho}. \tag{B8}$$

Instead, in the periodic case we can use the fact that the periodicity condition requires $E_k \equiv E_{l_k} = 2\pi l_k/T$ for l_k an integer. This means that

$$\Delta_{kk'} = \frac{1}{N} \frac{1 - e^{i(E_k - E_{k'})T}}{1 - e^{i(E_k - E_{k'})\varepsilon}} = \delta_{kk'},$$

with the last equality holding only in the periodic case. By replacing this in Eq. (B5), we obtain precisely $\rho_S = \bar{\rho}$. ■

Notice that the Corollaries 1 and 2 follow immediately from the Schur concavity of functions defining any entropy. These corollaries are the particular case of the linear entropy. Interestingly, bounds can be obtained by just considering other entropies.

Let us also make more precise the main body statement that the quantum time formalism gives an interpretation to the loss of coherences induced by a time average: As we said, since the system is "entangled with time," by ignoring the "clock qubits" we lose information. For discrete time, the information loss induces $\rho \to \tilde{\rho}$ that is a *quantum channel* with Krauss operators $K_t := e^{-iHt\varepsilon}/\sqrt{N}$ such that

$$\tilde{\rho} = \sum_t K_t^\dagger \rho K_t. \tag{B9}$$

As usual, this quantum channel can be purified by using the isometric extension [135] $U[K_t] := \sum_t |t\rangle \otimes K_t$ so that the channel is recovered by tracing over the "environment" (here the clock) in a global state $U[K_t]|\psi\rangle$. This global state is precisely the history state of Eq. (4), i.e.,

$$|\Psi\rangle = U[K_t]|\psi_0\rangle. \tag{B10}$$

We see that one can "rediscover" the quantum time formalism from the natural purification of the channel (B9). Just as in the general case, where a nontrivial channel is induced by correlations between the system and an environment, the system is correlated (entangled) with time. Moreover, the quantum channel's theory [135] implies that there is a unitary V such that

$$|\Psi\rangle = V|0\rangle \otimes |\psi\rangle. \tag{B11}$$

One possible unitary is provided by the circuit of Fig. 2. We should remark, however, that the operation of tracing over the clock degrees of freedom is very different from measuring on the clock register and conditioning the state of the system. When conditioning, one has access to the clock, as it is required for implementing a projection at a given time state. In summary, the history state contains both information about evolution at specific times, recovered from conditioning, and about the "equilibration channel" (B9), recovered by ignoring the clock.

APPENDIX C: PROOFS INVOLVING CIRCUITS

1. Proof of Proposition 1

Proof. Consider the circuit in Fig. 4. For ease of calculation, we will assume that $\rho = |\psi_0\rangle\langle\psi_0|$, for some initial state $|\psi_0\rangle$. Moreover, we recall that both O_1 and O_2 are Pauli operators. The input state to the circuit is $|\psi_0\rangle|0\rangle$, where the single-qubit state initialized in the zero state is an ancilla used to perform a Hadamard test. The action of the Hadamard gates is to map

$$|\psi_0\rangle|0\rangle \to |\psi_0\rangle|+\rangle. \tag{C1}$$

Assuming that the colored dashed gate in Fig. 4 is an identity, we next have a controlled O_2 operation. This produces the state

$$|\psi_0\rangle|+\rangle \to \frac{1}{\sqrt{2}}(|\psi_0\rangle|0\rangle + O_2|\psi_0\rangle|1\rangle). \tag{C2}$$

Next, in the tth experiment, we evolve the state with a unitary $U(\varepsilon t)$, leading to

$$\frac{1}{\sqrt{2}}(|\psi_0\rangle|0\rangle + O_2|\psi_0\rangle|1\rangle) \to \frac{1}{\sqrt{2}}(U(\varepsilon t)|\psi_0\rangle|0\rangle + U(\varepsilon t)O_2|\psi_0\rangle|1\rangle). \tag{C3}$$

The next controlled O_1 gate leads to

$$\frac{1}{\sqrt{2}}(U(\varepsilon t)|\psi_0\rangle|0\rangle + U(\varepsilon t)O_2|\psi_0\rangle|1\rangle) \to \frac{1}{\sqrt{2}}(U(\varepsilon t)|\psi_0\rangle|0\rangle + O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle). \tag{C4}$$

The final Hadamard gate produces the state

$$\begin{aligned}
&\frac{1}{\sqrt{2}}(U(\varepsilon t)|\psi_0\rangle|0\rangle + O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle) \\
&\quad\to \frac{1}{2}(U(\varepsilon t)|\psi_0\rangle(|0\rangle + |1\rangle) \\
&\qquad + O_1U(\varepsilon t)O_2|\psi_0\rangle(|0\rangle - |1\rangle)) \\
&\quad= \frac{1}{2}((U(\varepsilon t)|\psi_0\rangle + O_1U(\varepsilon t)O_2|\psi_0\rangle)|0\rangle \\
&\qquad + (U(\varepsilon t)|\psi_0\rangle - O_1U(\varepsilon t)O_2|\psi_0\rangle)|1\rangle).
\end{aligned} \tag{C5}$$

Then, the probability of measuring the ancilla qubit in the zero state is

$$\begin{aligned} p(0) &= \tfrac{1}{4}((\langle\psi_0| U^\dagger(\varepsilon t) + \langle\psi_0| O_2 U^\dagger(\varepsilon t) O_1))((U(\varepsilon t)|\psi_0\rangle \\ &\quad + O_1 U(\varepsilon t) O_2 |\psi_0\rangle)) \\ &= \tfrac{1}{4}(2 + \langle\psi_0| U^\dagger(\varepsilon t) O_1 U(\varepsilon t) O_2 |\psi_0\rangle \\ &\quad + \langle\psi_0| O_2{}^\dagger U^\dagger(\varepsilon t) O_1{}^\dagger U(\varepsilon t)|\psi_0\rangle) \\ &= \tfrac{1}{2}\big(1 + \mathrm{Re}\big[\langle O_1(t) O_2\rangle_{\psi_0}\big]\big). \end{aligned} \tag{C6}$$

Similarly, the probability of measuring the ancilla qubit in the one state is

$$p(1) = \tfrac{1}{2}\big(1 - \mathrm{Re}\big[\langle O_1(t) O_2\rangle_{\psi_0}\big]\big). \tag{C7}$$

Combining Eqs. (C6) and (C7) shows that the expectation value of the Z operator on the ancilla qubit is

$$\langle Z\rangle = p(1) - p(0) = \mathrm{Re}\big[\langle O_1(t) O_2\rangle_{\psi_0}\big]. \tag{C8}$$

By adding an $S^\dagger$ gate in place of the colored dashed gate in Fig. 4, one finds

$$\langle Z\rangle = \mathrm{Im}\big[\langle O_1(t) O_2\rangle_{\psi_0}\big]. \tag{C9}$$

Since this procedure needs to be repeated N times, and since we want to estimate the expectation values up to precision δ, then one needs to perform $\mathcal{O}(N/\delta^2)$ experiments. ■

2. Proof of Theorem 1

Proof. Consider the circuit in Fig. 5. For ease of calculation, we will assume that $\rho = |\psi_0\rangle\langle\psi_0|$ for some initial state $|\psi_0\rangle$. We also recall that both O_1 and O_2 are Pauli operators. The input state to the circuit is $|0\rangle^{\log N}|\psi_0\rangle|0\rangle$, where we recall that $|0\rangle^{\log N}$ is the initial state of the clock qubits. The action of the Hadamard gates is to map

$$|0\rangle^{\log N}|\psi_0\rangle|0\rangle \to \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}|t\rangle|\psi_0\rangle|+\rangle\,. \tag{C10}$$

Here, we have used the identity $|+\rangle^{\log(N)} = \frac{1}{\sqrt{N}}\otimes_{j=1}^{\log N}(|0_j\rangle + |1_j\rangle) = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}|t\rangle$ that follows by expressing t in its binary form $t = \sum_{j=1}^{\log N} t_j 2^{j-1}$. Assuming the colored dashed gate in Fig. 5 is an identity, we next have a controlled O_2 operation. This produces the state

$$\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}|t\rangle|\psi_0\rangle|+\rangle \to \frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(|\psi_0\rangle|0\rangle + O_2|\psi_0\rangle|1\rangle). \tag{C11}$$

Next, the sequence of $\log N$ controlled gates $U(2^{j-1}\frac{T}{N}) = U(\frac{T}{N})^{2^{j-1}}$ for $j = 1,\ldots,\log N$ perform the operations

$$\begin{aligned} &\frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(|\psi_0\rangle|0\rangle + O_2|\psi_0\rangle|1\rangle) \\ &\quad\to \frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(U(\varepsilon t)|\psi_0\rangle|0\rangle + U(\varepsilon t)O_2|\psi_0\rangle|1\rangle). \end{aligned} \tag{C12}$$

The next controlled O_1 gate leads to

$$\begin{aligned} &\frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(U(\varepsilon t)|\psi_0\rangle|0\rangle + U(\varepsilon t)O_2|\psi_0\rangle|1\rangle) \\ &\quad\to \frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(U(\varepsilon t)|\psi_0\rangle|0\rangle + O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle), \end{aligned} \tag{C13}$$

while the controlled phase gates perform the map

$$\begin{aligned} &\frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}|t\rangle(U(\varepsilon t)|\psi_0\rangle|0\rangle + O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle) \\ &\quad\to \frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}(|t\rangle U(\varepsilon t)|\psi_0\rangle|0\rangle \\ &\qquad + e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle). \end{aligned} \tag{C14}$$

The final Hadamard gate gives rise to

$$\begin{aligned} &\frac{1}{\sqrt{2N}}\sum_{t=0}^{N-1}(|t\rangle U(\varepsilon t)|\psi_0\rangle|0\rangle + e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle|1\rangle) \\ &\quad\to \frac{1}{2\sqrt{N}}\sum_{t=0}^{N-1}(|t\rangle U(\varepsilon t)|\psi_0\rangle(|0\rangle + |1\rangle) \\ &\qquad + e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle(|0\rangle - |1\rangle)) \\ &\quad= \frac{1}{2\sqrt{N}}\sum_{t=0}^{N-1}(|t\rangle U(\varepsilon t)|\psi_0\rangle \\ &\qquad + e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle)|0\rangle \\ &\qquad + (|t\rangle U(\varepsilon t)|\psi_0\rangle - e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle)|1\rangle\,. \end{aligned} \tag{C15}$$

Thus, the probability of measuring the ancilla qubit in the zero state is

$$\begin{aligned} p(0) &= \frac{1}{4N}\sum_{t=0}^{N-1}\sum_{t'=0}^{N-1}(\langle t'|\langle\psi_0| U^\dagger(t') \\ &\quad + e^{i\omega\varepsilon t'}\langle t'|\langle\psi_0| O_2{}^\dagger U^\dagger(t') O_1{}^\dagger)(|t\rangle U(\varepsilon t)|\psi_0\rangle \\ &\quad + e^{-i\omega\varepsilon t}|t\rangle O_1U(\varepsilon t)O_2|\psi_0\rangle) \\ &= \frac{1}{4N}\sum_{t=0}^{N-1}(2 + e^{-i\omega\varepsilon t}\langle\psi_0| U^\dagger(\varepsilon t)O_1U(\varepsilon t)O_2|\psi_0\rangle \\ &\quad + e^{i\omega\varepsilon t}\langle\psi_0| O_2{}^\dagger U^\dagger(\varepsilon t)O_1{}^\dagger U(\varepsilon t)|\psi_0\rangle) \\ &= \frac{1}{2}(1 + \mathrm{Re}[\widetilde{F}(O_1, O_2, \omega)]). \end{aligned} \tag{C16}$$

Similarly, the probability of measuring the ancilla qubit in the one state is

$$p(1) = \tfrac{1}{2}(1 - \mathrm{Re}[\widetilde{F}(O_1, O_2, \omega)]). \tag{C17}$$

Combining Eqs. (C16) and (C17) shows that the expectation value of the Z operator on the ancilla qubit is

$$\langle Z\rangle = p(1) - p(0) = \mathrm{Re}[\widetilde{F}(O_1, O_2, \omega)]. \tag{C18}$$

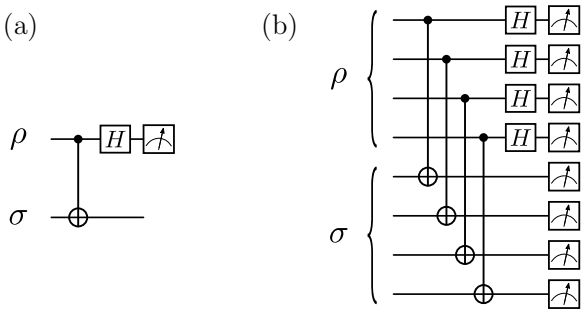


FIG. 20. Algorithm for computing the overlap between two quantum states. We show an algorithm that takes as input two arbitrary n-qubit quantum states ρ and σ and that estimates the overlap $\mathrm{Tr}[\rho\sigma]$. In panel (a), we show the algorithm for the case when ρ and σ are single-qubit states, and in panel (b) the generalization for larger qubit sizes. We can see that in all cases the circuit depth is equal to 2, and hence independent of n.

By adding an $S^\dagger$ gate in place of the colored dashed gate in Fig. 5, one finds

$$\langle Z\rangle = \mathrm{Im}[\widetilde{F}(O_1, O_2, \omega)]. \tag{C19}$$

If one wishes to estimate $\langle Z\rangle$ up to precision δ, then one needs to perform $\mathcal{O}(1/\delta^2)$ measurements. This results in $\mathcal{O}(1/\delta^2)$ experiments. ■

3. Proof of Proposition 2

Proof. Consider the circuit in Fig. 6. We can readily see that the circuit therein is nothing but the circuit for computing the overlap between two quantum states ρ and σ derived in Ref. [45] and also shown in Fig. 20.

To understand the algorithm of Ref. [45], let us consider first the case of ρ and σ being single-qubit states. We know that the following identity holds:

$$\mathrm{Tr}[\rho\sigma] = \mathrm{Tr}[(\rho\otimes\sigma)\mathrm{SWAP}], \tag{C20}$$

where SWAP denotes the SWAP operator, whose action can be defined in the computational basis as $\mathrm{SWAP}\,|ij\rangle = |ji\rangle$. Then, let us define the Bell basis:

$$\begin{aligned} |\Phi_1\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad |\Phi_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\ |\Phi_3\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad |\Phi_4\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \end{aligned} \tag{C21}$$

It is not hard to see that the SWAP operator is diagonal in the Bell basis, and that it can be expressed as

$$\mathrm{SWAP} = |\Phi_1\rangle\langle\Phi_1| + |\Phi_2\rangle\langle\Phi_2| + |\Phi_3\rangle\langle\Phi_3| - |\Phi_4\rangle\langle\Phi_4|\,. \tag{C22}$$

Thus, we can estimate the expectation value of the SWAP operator over the state $\rho\otimes\sigma$:

$$\begin{aligned} \langle\mathrm{SWAP}\rangle_{\rho\otimes\sigma} &= \mathrm{Tr}[(\rho\otimes\sigma)\mathrm{SWAP}] \\ &= P(\Phi_1) + P(\Phi_2) + P(\Phi_3) - P(\Phi_4), \end{aligned} \tag{C23}$$

where here we defined the probabilities $P(\Phi_i) = \langle\Phi_i|\rho\otimes\sigma|\Phi_i\rangle$. That is, $P(\Phi_i)$ denotes the probability of measuring the state $\rho\otimes\sigma$ in the Bell basis and obtaining the measurement outcome $|\Phi_i\rangle$. This result shows that we cane estimate $\mathrm{Tr}[\rho\sigma]$ by preparing $\rho\otimes\sigma$ and measuring in the Bell basis. Crucially, one can readily prove the such a measurement can be performed with the circuit in Fig. 20(a), which is composed of a cnot gate and a Hadamard gate in the first qubit (i.e., with the inverse of the circuit used to prepare a Bell state).

A similar result will follow when $\rho\otimes\sigma$ are n-qubit states, but now one has to measure the expectation value of the operator $\prod_{j=1}^n \mathrm{SWAP}_{j_A,j_B}$, where here SWAP_{j_A,j_B} denotes the operator that swaps the jth qubit of ρ with the jth qubit of σ. Since each operator SWAP_{j_A,j_B} can be expanded in a (local) Bell basis, the circuit used to measure in the eigenbasis of $\prod_{j=1}^n \mathrm{SWAP}_{j_A,j_B}$ and hence to estimate $\mathrm{Tr}[\rho\sigma]$ is precisely that which is shown in Fig. 20(b). Finally, since we are computing an expectation value, if we want to reach a precision δ for each N, we need to run $\mathcal{O}(N/\delta^2)$ experiments. ■

4. Proof of Theorem 2

Proof. The proof of this theorem follows very closely that of Proposition 2. In particular, we can see from Fig. 7 that we are performing a SWAP test, i.e., measuring the overlap via Bell-basis measurements, between the reduced state ρ_S of the history state on the system qubits and the state $|\psi_0\rangle\langle\psi_0|$. On the other hand, one can readily see by expanding ρ_S as in Eq. (B1) that the overlap between ρ_S and $|\psi_0\rangle\langle\psi_0|$ is precisely $\widetilde{\mathcal{L}}(\psi_0)$. Hence, we can use the circuit in Fig. 7 to estimate $\widetilde{\mathcal{L}}(\psi_0)$ up to δ precision with $\mathcal{O}(1/\delta^2)$ experiments. ■

APPENDIX D: PROOF OF THEOREM 5

Proof. Let us remark first that the gates we consider here are only those directly related to implementing the unitary evolution. The other parts of the circuit are common to both the sequential- and parallel-in-time approaches, meaning that it is advantageous to have less runs (thus the parallel approach is always more convenient). The only exception is when we add the phases in Fig. 5 but this only adds $\log(N)$ gates, which is negligible in this analysis. A fully explicit depth analysis taking into account all gates for each circuit is provided in Appendix F in the context of the variational diagonalization scheme.

Let us discuss first the conventional sequential approach. The maximum number of gates one requires in a single run is the "last one" corresponding to $U(T)$. In this case, one needs $\#_g(T) \in \gamma l T^\alpha$ gates. A more relevant quantity is the total number of gates $\#^{\mathrm{seq}}$ one needs to implement since this will determine the actual computational time. By summing over the amount of gates of each run, one obtains

$$\#^{\mathrm{seq}} = \gamma l \sum_{t=1}^{N-1} (\varepsilon t)^\alpha \in \gamma l \mathcal{O}(N^{\alpha+1}), \tag{D1}$$

where we retain the information on l for a proper comparison.

Now let us consider the parallel-in-time approach. In this scheme, each $U(t)$ is first "Trotterized" and each individual gate is controlled. This gives an additional overall multiplicative factor proportional to l, say βl. For local Hamiltonians, this is logarithmic in the system size. On the other hand, the parallelization reduces the number of evolution gates

exponentially. The total number of gates $\#^{\text{par}}$ is in fact now

$$\#^{\text{par}} = \gamma\beta l^2 \sum_{j=0}^{\log(N)-1} (\varepsilon 2^j)^\alpha \in \beta\gamma l^2 \mathcal{O}(N^\alpha). \tag{D2}$$

Notice that contrary to the sequential case, in the parallel-in-time scheme we sum over a logarithmic amount of times, with $U(T/2)$ the gate with the longest evolution in the whole protocol ($T/2 = \varepsilon 2^{\log(N)-1}$). By comparing the previous equations the last part of the Theorem immediately follows. ■

APPENDIX E: PROOF OF PROPOSITION 4

Proof. Note that $\mathfrak{k}$ admits a *bodyness-graded* basis

$$B = \bigcup_{k=2}^{n} B_k, \quad \text{where } B_k = \{\widehat{X_i Y}_{i+k-1}, \widehat{Y_i X}_{i+k-1}\}_{i=1}^{n-k}. \tag{E1}$$

Clearly, $|B_k| = 2(n-k)$ and $2\sum_{k=2}^{n}(n-k) = 2\binom{n}{2} = n(n-1) = \dim(\mathfrak{k})$. The claim is that the set $B_2 \subset B$ generates $\mathfrak{k}$.

We prove this by induction. First, we show B_2 generates B_3. For all $i \in 1, \ldots, n-2$, we obtain each element in B_3 from

$$[X_i Y_{i+1}, X_{i+1} Y_{i+2}] \propto X_i Z_{i+1} Y_{i+2}, \tag{E2}$$

$$[Y_i X_{i+1}, Y_{i+1} X_{i+2}] \propto Y_i Z_{i+1} Z_{i+2}. \tag{E3}$$

Now let us see that, in general, B_2 and B_k suffice to generate B_{k+1}. Specifically, each element in B_{k+1} follows from

$$[X_i Y_{i+1}, \widehat{X_{i+1} Y}_{i+k-1}] \propto \widehat{X_i Y}_{i+k}, \tag{E4}$$

$$[Y_i X_{i+1}, \widehat{Y_{i+1} X}_{i+k-1}] \propto \widehat{Y_i X}_{i+k}, \tag{E5}$$

for $i \in 1, \ldots, n-(k-1)$. ■

APPENDIX F: PROOF OF THEOREM 7

Proof. In what follows, we will assume that the history state preparation subroutine is replaced as in Fig. 11 where W and D are, respectively, given by the optimized ansätze $W(\boldsymbol{\alpha})$ and $W(\boldsymbol{\beta})$. Moreover, we recall from Ref. [97] that for $W(\boldsymbol{\theta})$ to be overparametrized it needs a number of parameters in $\mathcal{O}(\dim(\mathfrak{k})) = \mathcal{O}(n^2)$. Since the ansatz in Eq. (48) has $(2n-2)$ parameters per layer (see also Fig. 12), then we need L to be in $\mathcal{O}(\dim(\mathfrak{k}))/(2n-2)) = \mathcal{O}(n)$.

We begin with the history state preparation circuit in Fig. 5. Let us list the gates that we need to implement and their respective depths. Gates that are listed on the same list item can be parallelized.

(1) Hadamard gates on the clock qubits and $W(\boldsymbol{\alpha})$ on the system qubits. The depth is $\mathcal{O}(n)$ as it is dominated by that of $W(\boldsymbol{\alpha})$.

(2) $\log N$ controlled gates $e^{-iD2^{j-1}\varepsilon}$ for $j = 1, \ldots, \log N$. Each controlled gate has a depth in $\mathcal{O}(n)$ as we need to control single-qubit gates [see Fig. 11(b)]. The depth is then $\mathcal{O}(\log(M)n)$.

(3) A gate $W^\dagger(\boldsymbol{\alpha})$ on the system qubits. The depth is $\mathcal{O}(n)$.

Hence, the total depth to implement the circuit in Fig. 5 is $\mathcal{O}(\log(M)n + 2n) = \mathcal{O}(\log(M)n)$.

Next, let us consider the circuit in Fig. 7. Assuming as in Theorem 1 that O_1 and O_2 are Pauli operators we have.

(1) Hadamard gates on the clock and ancilla qubits and $W(\boldsymbol{\alpha})$ on the system qubits. The depth is $\mathcal{O}(n)$ as it is dominated by that of $W(\boldsymbol{\alpha})$.

(2) Apply a controlled O_2 gate. If O_2 acts nontrivially on at most k_2 qubits (with $k_2 \leqslant n$), then the controlled O_2 gate is implemented with $\mathcal{O}(k_2)$ gates, or more generally with depth in $\mathcal{O}(n)$.

(3) $\log N$ controlled gates $e^{-iD2^{j-1}\varepsilon}$ for $j = 1, \ldots, \log N$. Each controlled gate has a depth in $\mathcal{O}(n)$ as we need to control single-qubit gates [see Fig. 11(b)]. The depth is then $\mathcal{O}(\log(M)n)$.

(4) A gate $W^\dagger(\boldsymbol{\alpha})$ on the system qubits. The depth is $\mathcal{O}(n)$.

(5) Apply a controlled O_1 gate. If O_1 acts nontrivially on at most k_1 qubits (with $k_1 \leqslant n$), then the controlled O_1 gate is implemented with $\mathcal{O}(k_1)$ gates, or more generally with depth in $\mathcal{O}(n)$.

(6) Apply $\log(M)$ controlled phase gates. Their depth is $\mathcal{O}(\log(M))$.

The total depth of the circuit is then $\mathcal{O}(\log(M)n + 4n + \log(M)) = \mathcal{O}(\log(M)n)$. Note that if O_1 and O_2 are instead expressed as a sum of Pauli operators, then the experiment complexity changes, but not the circuit depth.

We now consider the circuit in Fig. 8. Here, we know that we need to prepare two copies of the history states, which can be done with depth in $\mathcal{O}(\log(M)n)$, plus implement a depth-two state-overlap circuit [45]. Hence, the depth is still in $\mathcal{O}(\log(M)n)$.

Finally, the circuit in Fig. 9 is composed of the history state preparation circuit plus a random unitary on each qubit. The depth is then in $\mathcal{O}(\log(M)n)$. ■

APPENDIX G: NUMERICAL METHODS

Here, we discuss the analytical techniques used to reduce the computational cost in the simulations of Sec. VI.

1. Simulations via Jordan Wigner

Consider first the nonuniform XX Hamiltonian of Eq. (42), for n sites and periodic boundary conditions. We have employed the well-known JW map $c_j^\dagger = s_j^+ \exp(-i\pi \sum_{k=1}^{j-1} s_k^+ s_k^-)$ [136] [with $s_j^\pm = (X_j \pm iY_j)/2$] in order to write

$$H_\sigma = \frac{J}{2}\sum_{j=1}^{n}(c_j^\dagger c_{j+1} + \text{H.c.}) + \lambda \sum_{j=1}^{n} \cos(2\pi\alpha j) c_j^\dagger c_j - J\delta_{\sigma 1}(c_n^\dagger c_{n+1} + c_{n+1}^\dagger c_n), \tag{G1}$$

which up to the border term is the Aubry-Andre Hamiltonian [107], with $\{c_j, c_{j'}^\dagger\} = \delta_{jj'}$ and the other anticommutators vanishing. Here, σ indicates the parity dependence of H_σ on the sector of states in which acts, with $\sigma = \pm 1 \equiv e^{i\pi N}$ [137] and with the convention $c_{n+1} = c_1$. As usual, the vacuum state $|0\rangle$ is mapped to $|\downarrow\cdots\downarrow\rangle$ with $Z_j = 2c_j^\dagger c_j - 1$.

Notice that for each parity we can write $H_\sigma = \mathbf{c}^\dagger M_\sigma \mathbf{c}$ with $\mathbf{c} = (c_1 c_2 \cdots)^t$ and M_σ an $n \times n$ matrix. For general "single-particle" states $|\psi\rangle = \sum_j \psi_j c_j^\dagger |0\rangle \equiv \psi_1|\uparrow\downarrow\cdots\rangle + \psi_2|\downarrow\uparrow\cdots\rangle + \cdots$, this allows us to write

$$\mathcal{L}(t) = |\langle\psi|e^{-iHt}|\psi\rangle|^2 = |\boldsymbol{\psi}^\dagger e^{-iM_{-1}t}\boldsymbol{\psi}|^2, \tag{G2}$$

with $\boldsymbol{\psi} = (\psi_1\psi_2\cdots)^t$. Thus, we only need to exponentiate a matrix of $n \times n$ dimensions, rather than the original Hamiltonian of size $2^n \times 2^n$. The states we employed in Sec. VI are mapped to sp states, allowing us to reach large values of n. In addition, one has to perform the sums and/or integrals in time. Let us also mention that the spectrum of the model exhibits remarkable (fractal) properties [107,138,139], rendering numerical treatments almost mandatory for our purposes.

To obtain $\widetilde{\mathcal{L}}(\psi_0)$, we basically summed the expression (G2) over discrete times. Instead, in order to obtain the exact Loschmidt's echo average we used

$$\overline{\mathcal{L}}(\psi_0) = \sum_k |\langle k|\psi\rangle|^4 \tag{G3}$$

for $|k\rangle$ of the eigenbasis H. Notice that we do not need the complete basis $|k\rangle$ but only those eigenstates corresponding to sp excitations so that $|k\rangle \equiv \boldsymbol{\phi}_k \boldsymbol{c}^\dagger|0\rangle$ implying $\langle k|\psi\rangle = \boldsymbol{\phi}_k^\dagger\boldsymbol{\psi}$, for $|\psi\rangle$ an sp state.

For computing the infinite temporal variance of observables, we used a similar strategy. One can prove that for a general observable and state [4] $\sigma_O^2 = \sum_{k\neq k'} |\rho_{kk'}|^2|O_{kk'}|^2$ with $O_{kk'} = \langle k|O|k'\rangle$, $\rho_{kk'} = \langle k|\rho|k'\rangle$. In the main body example, we used a single-particle operator that in fermionic notation can be written as $O = c^\dagger_{L/2}c_{L/2+1} + \text{H.c.} = \sum_{i,j} M_{ij}c_i^\dagger c_j$. Thus, one can replace $|O_{kk'}|^2$ with $|M_{kk'}|^2 := |\phi_k^\dagger M \phi_{k'}|^2$ and $|\rho_{kk'}|^2$ with $|\boldsymbol{\phi}_k^\dagger\boldsymbol{\psi}|^2|\boldsymbol{\phi}_{k'}^\dagger\boldsymbol{\psi}|^2$.

Regarding now the numerical computation of E_2, notice first that

$$\text{Tr}\rho_S^2 = \text{Tr}(\text{Tr}_T|\Psi\rangle\langle\Psi|)^2 = \frac{1}{N^2}\sum_{t,t'}|\langle\psi(\varepsilon t')|\psi(\varepsilon t)\rangle|^2; \tag{G4}$$

i.e., the quantity $\text{Tr}\rho_T^2 = \text{Tr}\rho_S^2$ also depends on the overlap of the state at different times. For sp states, these overlaps have the expression

$$\langle\psi(t')|\psi(t)\rangle = \boldsymbol{\psi}^\dagger e^{-iM(t-t')}\boldsymbol{\psi}, \tag{G5}$$

meaning again that we only need to exponentiate a matrix of $m \times m$ size. For time-independent Hamiltonians, only the time differences matter. This allows one to write the ρ_S purity as the single sum

$$\text{Tr}\rho_S^2 = \frac{2}{N^2}\sum_t (N-t)\mathcal{L}(\varepsilon t) - \frac{1}{N}, \tag{G6}$$

which holds for general time-independent Hamiltonians. For our numerics, we have combined Eqs. (G6) and (G2).

2. Variational Hamiltonian diagonalization

Here, we present the details of the variational Hamiltonian diagonalization algorithm. All the simulations were conducted with the Julia programming language [140] with the LINEARALGEBRA and SPARSEARRAYS packages.

We trained the parameters in the ansatz by using gradient-based minimization, resorting to the *ADAM* optimizer [141] for both the diagonalization unitary parameters $\boldsymbol{\alpha}$ and the diagonal, eigenvalues, matrix ones $\boldsymbol{\beta}$. We used two different instances of the ADAM optimizer, both initialized with the same learning rate $\eta_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0.1$. Although unusually large (the standard value for ADAM's initial learning rate is $\eta = 0.001$), this learning rate proved to be very effective. A training instance of our model comprises $n_r = 10$ randomly initialized runs, each lasting at most $n_I = 10^5$ training iterations. In practice, runs that achieve loss values less than 10^{-14} are stopped and deemed successful. We studied the $N = 6$ sites nonuniform XX model defined by the Hamiltonian in Eq. (42) for the value $\alpha = \frac{\sqrt{5}-1}{2}$ that makes it possible, in the thermodynamic limit, for a delocalization-localization transition at the critical point $\lambda = J$ to exist.

3. Simulation of the interacting case

Consider now the fully interacting model of Eq. (49). Under the JW map, and assuming open boundary conditions, the term $H_{\text{int}} = \sum_{j=1}^{n-1} Z_jZ_{j+1}$ is mapped to $H_{\text{int}} = 4\sum_{j=1}^{n-1}(c_j^\dagger c_j - 1/2)(c_{j+1}^\dagger c_{j+1} - 1/2)$ that from the fermionic perspective corresponds to interactions, essentially of the Hubbard model type [142]. The cost of exactly simulating the evolution of arbitrary states classically becomes exponential for this model. However, considering that the Hamiltonian preserves the number of fermions, if we restrict our study to a low number of particles we can diagonalize the Hamiltonian in polynomial subspaces.

We restrict our study to a total number of particles $N' = 2$. As noticed in the main body, this simple scenario already captures the phenomenon of single-particle localization surviving interactions. In this subspace, we can write

$$H_{\text{int}} \equiv 4\sum_{j=1}^{n-1} c_j^\dagger c_j c_{j+1}^\dagger c_{j+1} - c_1^\dagger c_1 - c_n^\dagger c_n. \tag{G7}$$

Now we introduce the notation $|k,l\rangle = c_k^\dagger c_l^\dagger|0\rangle$ and define the $n^2 \times n^2$ matrix $\bar{H}_{ij;kl} := \langle ij|H|kl\rangle$ with (i,j) and (k,l) joined indices leading to n^2 distinct values; i.e., we can write $(i,j) \to a = n(i-1)+j$ for $a = 1,\ldots,n^2$. The entries of $\bar{H}$, although cumbersome, are straightforwardly obtained via fermionic contractions, yielding

$$\langle i,j|\sum_{i',j'} h_{i'j'}c_{i'}^\dagger c_{j'}|k,l\rangle = \delta_{jl}h_{ik} - \delta_{il}h_{jk} - \delta_{jk}h_{il} + \delta_{ik}h_{jl}, \tag{G8}$$

$$\begin{aligned}\langle i,j|H_{\text{int}}|k,l\rangle &= 2\{\delta_{j1}\delta_{k1}\delta_{il} - \delta_{i1}\delta_{k1}\delta_{jl} - \delta_{j1}\delta_{l1}\delta_{ik} + \delta_{i1}\delta_{l1}\delta_{jk} \\ &\quad - (-\delta_{in}\delta_{jl}\delta_{kn} + \delta_{il}\delta_{jn}\delta_{kn} - \delta_{ik}\delta_{jn}\delta_{ln} + \delta_{in}\delta_{jk}\delta_{ln})\} \\ &\quad + 4\{\delta_{ji+1}(\delta_{ik}\delta_{i+1l} - \delta_{i+1k}\delta_{il}) \\ &\quad - \delta_{ij+1}(\delta_{jk}\delta_{j+1l} - \delta_{j+1k}\delta_{jl})\},\end{aligned} \tag{G9}$$

from which $\bar{H}$ is immediately obtained by choosing h_{ij} as in Eq. (49), adding the interacting part multiplied by Δ and properly reshaping the indices.

Given the polynomial size, one can diagonalize $\bar{H}$ exactly. One obtains numerically $\sum_b \bar{H}_{a;b}\Psi_b^{(\mu)} = \lambda^\mu\Psi_a^{(\mu)}$ and then aims to define fermionic states that are eigenstates of H. In order to do so, an additional projection is generally needed so that $\Psi_{kl}^\mu = -\Psi_{lk}^\mu$. Having projected the $n^2 \times 1$ vector $\Psi^{(\mu)}$ accordingly, it is reshaped to an $n \times n$ matrix allowing to

define

$$|\mu\rangle := \sum_{k,l} \Psi_{kl}^{(\mu)} |k, l\rangle, \tag{G10}$$

satisfying $H|\mu\rangle = \lambda_\mu |\mu\rangle$. It is now straightforward to write $\mathcal{L}(t) = \sum_\mu \langle\psi|\mu\rangle\langle\mu|\psi\rangle e^{iE\lambda_\mu} = \sum_\mu |\langle\psi|\mu\rangle|^2 e^{iE\lambda_\mu}$ since the states $|\mu\rangle$ provide a complete basis of the $N' = 2$ subspace. In addition, for $|\psi\rangle = |i, j\rangle$ one obtains $|\langle\psi|\mu\rangle|^2 = |\Psi_{ij}^{(\mu)}|^2$. With these results at hand, one can write the temporal sums involved in computing $\bar{\mathcal{L}}(i, j)$, namely, the average of the Loschmidt's echo for the initial state $|\psi\rangle = |i, j\rangle$, in closed form:

$$\widetilde{\mathcal{L}}(i, j) = \frac{4}{N} \sum_{\mu \leqslant \mu' = 1}^{n^2} \left|\Psi_{ij(\mu)}^{(\mu)}\right|^2 \left|\Psi_{ij(\mu')}^{(\mu')}\right|^2 \left[1 - \cos(T(\lambda_\mu - \lambda_{\mu'})) + \cot\left(\frac{\varepsilon}{2}(\lambda_\mu - \lambda_{\mu'})\right) \sin(T(\lambda_\mu - \lambda_{\mu'}))\right]. \tag{G11}$$

Similarly, and using (G6), we obtain

$$\mathrm{Tr}\left[\rho_S^2\right] = \frac{4}{N} \sum_{\mu \leqslant \mu' = 1}^{n^2} \left|\Psi_{ij(\mu)}^{(\mu)}\right|^2 \left|\Psi_{ij(\mu')}^{(\mu')}\right|^2 \left[\frac{\csc^2\left(\frac{\varepsilon}{4}(\lambda_\mu - \lambda_{\mu'})\right)\left(N\left(-\cos\left(\frac{\varepsilon}{2}(\lambda_\mu - \lambda_{\mu'})\right)\right) - \cos\left(\frac{T}{2}(\lambda_\mu - \lambda_{\mu'})\right) + N + 1\right)}{N^2} - \frac{1}{N}\right]. \tag{G12}$$

In these expressions, $ij(\mu)$ represents those ij for which $m = n(i-1) + j$. These are the final expressions we employed in Sec. VI, with the eigenvalues λ_μ and eigenfunctions $\Psi_{ij}^{(\mu)}$ obtained via numerical diagonalization as described above. Finally, notice that by similar means we can compute $\bar{\mathcal{L}}(\psi_0)$ using

$$\bar{\mathcal{L}}(\psi_0) = \sum_\mu |\langle\psi|\mu\rangle|^4 = 4 \sum_\mu \left|\Psi_{ij(\mu)}^{(\mu)}\right|^4, \tag{G13}$$

which may be recovered from Eq. (G11) and taking the limit $T \to \infty$.

Finally, we provide yet another method of simulation of our chain based on TN methods [110,143–145]. In the case of one-dimensional systems, as the one studied throughout our manuscript, MPS methods are the best tool TNs have to offer. MPSs are a powerful tool for simulating the dynamics of quantum many-body systems as they provide an efficient classical representation of quantum states with limited entanglement, which is often the case for ground states and low-energy excitations of local Hamiltonians. In a quantum system with N sites, the state $|\psi\rangle$ is typically described by a rank-N tensor $\Psi_{i_1 i_2 \ldots i_N}$, where each index i_k corresponds to a local basis state at site k. For a system of d-dimensional local spaces, this tensor has d^N elements, which becomes infeasible to store and manipulate directly for large N. The key idea of MPS is to express this large tensor as a contraction of a series of smaller tensors, each associated with a site

$$\Psi_{i_1 i_2 \cdots i_N} = \sum_{\alpha_1, \alpha_2, \ldots, \alpha_{N-1}} A_{i_1 \alpha_1}^{[1]} A_{\alpha_1 i_2 \alpha_2}^{[2]} \cdots A_{\alpha_{N-1} i_N}^{[N]}. \tag{G14}$$

Here, each $A^{[k]}$ is a tensor with three indices: i_k is the physical index corresponding to the local basis at site k, and α_{k-1} and α_k are the bond indices that connect neighboring sites. The bond dimension χ is the maximum dimension of the α indices, determining the amount of entanglement the MPS can represent. Larger χ values allow for the representation of more entangled states but increase computational costs.

To simulate the time evolution of an initial quantum state $|\psi_0\rangle$ under a given Hamiltonian H, one can use Trotterization. The goal is to approximate the time evolution operator e^{-iHt}, which describes how the state evolves over time t, as a sequence of local gates. This is achieved by breaking the evolution into smaller time steps δt and decomposing the Hamiltonian into local terms:

$$e^{-iHt} \approx (e^{-iH_1\delta t} e^{-iH_2\delta t} \cdots e^{-iH_N\delta t})^n, \tag{G15}$$

where $n = t/\delta t$ and $H = H_1 + H_2 + \cdots + H_N$. The second-order Trotter-Suzuki decomposition is commonly used due to its balance between accuracy and computational efficiency. The error for a single time step in second-order Trotterization scales as $\mathcal{O}((\delta t)^3)$, meaning that the total error over the entire simulation time is $\mathcal{O}(t(\delta t)^2)$. Notice that studying Trotterization and its performance is crucial, as it could play a key role in implementing quantum evolution algorithms on actual quantum computers.

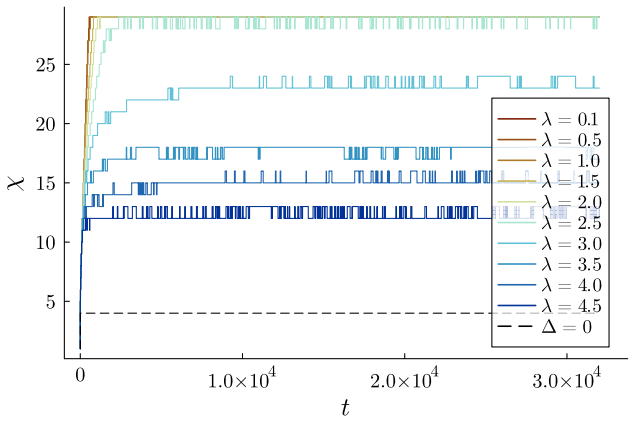


FIG. 21. Scaling of bond dimension with time. We report the scaling with time of the bond dimension χ of the MPS used to represent the interacting model state, with $\Delta = 0.05$. The black, dashed line shows the noninteracting model ($\Delta = 0$) bond dimension, which does not depend on λ and is equal to $\chi = 4$ throughout the whole system's evolution. Instead, in the interacting case, the maximum value of χ becomes smaller as the localization strength λ increases, confirming that the stronger the localization the easier the description of the system.

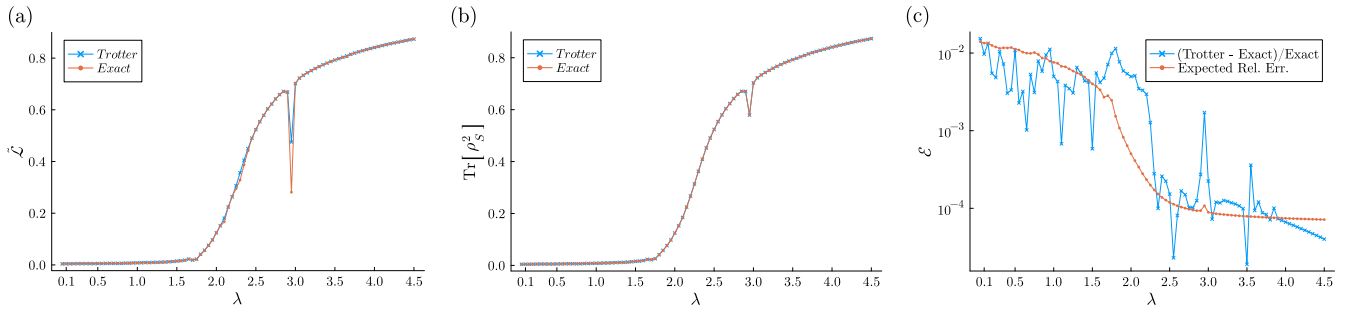


FIG. 22. Trotterization error for Loschmidt echo and entanglement. Comparison between the *exact* diagonalization technique and the Trotterized (*Trotter*) MPS method for the interacting model with $\Delta = 0.05$. Panels (a) and (b) show the Loschmidt echo $\tilde{\mathcal{L}}$ and system-time entanglement $\mathrm{Tr}[\rho_S^2]$, respectively. We can numerically observe how the finer information encoded in the former cannot be exactly caputerd via Trotterization, whereas we get perfet accordance in the entanglement behavior. Finally, panel (c) shows that the numerical relative error between the two techniques for computing the system-time entanglement matches the expected error introduced by second-order Trotterization.

Hence, one can encode the initial state $|\psi_0\rangle$ in an MPS, and repeatedly updating it via the gates resulting from Trotterizing the evolution operator. To apply a gate to an MPS, one can resort to the Time-Evolving Block Decimation (TEBD) algorithm. TEBD takes advantage of the local nature of interactions in the Hamiltonian to update the MPS tensors sequentially by first joining two adjacent tensors $A^{[k]}$ and $A^{[k+1]}$ into one, contracting along the shared bond index, contracting the resulting tensor with the gate, and ultimately performing singular value decomposition to split the updated tensor back into two local ones. The computational cost of TEBD is $\mathcal{O}(N\chi^3)$, where N is the number of sites and χ is the bond dimension. It is important to notice that the bond dimension χ is a critical parameter in MPS simulations. Besides determining the maximum amount of entanglement that the MPS can capture, it also controls the feasibility of the chosen algorithm. While in principle one could not constraint χ, letting it increase during the evolution to account for the possibly increasing correlations, most often one decides for a maximum value $\chi_{\max}$ after which the MPS method begins to be approximate. In summary, MPS combined with Trotterization and the TEBD algorithm provides an efficient framework for simulating time evolution in one-dimensional quantum systems. The choice of the bond dimension $\chi_{\max}$ and the time step δt are key factors in managing computational cost and error.

Here, we showcase how such an MPS based method fares when facing the interacting model studied in the main text. Particularly, we are interested in the scaling of the unconstrained χ with the number of time steps, and in how much the Trotterization error makes the results deviate from the ones obtained via exact diagonalization. Particularly, Fig. 21 shows how χ grows with time in the interacting model with $\Delta = 0.05$ for the different values of λ considered in the main text. Compared to the noninteracting case, for which we find $\chi = 4$ through the whole evolution, now the less the system is localized (small λ) the more the correlations grow. However, we can see that the bond dimension eventually caps off at values smaller than $\chi = 30$, ensuring that the method remains very efficient.

Finally, in Fig. 22 we show a comparison between the values of the Loschmidt echo and of the system-time entanglement obtained through exact diagonalization and through MPS simulation in the interacting case, where we keep $\Delta = 0.05$. As expected, the Loschmidt echo is more susceptible to errors, and while the trend is reproduced, Trotterization fails at capturing the finer details. This does not happen for the system-time entanglement, which gets exactly reproduced. The last panel of Fig. 22 confirms the predicted error coming from Trotterization, showing that no further sources of uncertainty are affecting the simulation.

[1] G. Auletta and G. Parisi, *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results* (World Scientific, Singapore, 2001).

[2] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. **75**, 715 (2003).

[3] M. Schlosshauer, Decoherence, the measurement problem, and interpretations of quantum mechanics, Rev. Mod. Phys. **76**, 1267 (2005).

[4] P. Reimann, Foundation of statistical mechanics under experimentally realistic conditions, Phys. Rev. Lett. **101**, 190403 (2008).

[5] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, Princeton, New Jersey, 2014), Vol. 53.

[6] J. R. McClean, J. A. Parkhill, and A. Aspuru-Guzik, Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics, Proc. Natl. Acad. Sci. USA **110**, E3901 (2013).

[7] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum time, Phys. Rev. D **92**, 045033 (2015).

[8] A. Boette, R. Rossignoli, N. Gigena, and M. Cerezo, System-time entanglement in a discrete-time model, Phys. Rev. A **93**, 062127 (2016).

[9] A. Boette and R. Rossignoli, History states of systems and operators, Phys. Rev. A **98**, 032108 (2018).

[10] D. Horsman, C. Heunen, M. F. Pusey, J. Barrett, and R. W. Spekkens, Can a quantum state over time resemble a quantum state at a single time? Proc. R. Soc. A **473**, 20170395 (2017).

[11] N. L. Diaz, J. M. Matera, and R. Rossignoli, History state formalism for scalar particles, Phys. Rev. D **100**, 125020 (2019).
[12] N. L. Diaz and R. Rossignoli, History state formalism for Dirac's theory, Phys. Rev. D **99**, 045008 (2019).
[13] N. L. Diaz, J. M. Matera, and R. Rossignoli, Spacetime quantum actions, Phys. Rev. D **103**, 065011 (2021).
[14] N. L. Diaz, J. M. Matera, and R. Rossignoli, Path integrals from spacetime quantum actions, Ann. Phys. **479**, 170052 (2025).
[15] D. Pabón, L. Rebón, S. Bordakevich, N. Gigena, A. Boette, C. Iemmi, R. Rossignoli, and S. Ledesma, Parallel-in-time optical simulation of history states, Phys. Rev. A **99**, 062333 (2019).
[16] L. R. Mendes and D. O. Soares-Pinto, Time as a consequence of internal coherence, Proc. R. Soc. A **475**, 20190470 (2019).
[17] T. Favalli and A. Smerzi, Time observables in a timeless universe, Quantum **4**, 354 (2020).
[18] A. Valdés-Hernández, C. G. Maglione, A. P. Majtey, and A. Plastino, Emergent dynamics from entangled mixed states, Phys. Rev. A **102**, 052417 (2020).
[19] E. Castro-Ruiz, F. Giacomini, A. Belenchia, and Č. Brukner, Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems, Nat. Commun. **11**, 2672 (2020).
[20] L. R. Mendes, F. Brito, and D. O. Soares-Pinto, Non-linear equation of motion for Page-Wootters mechanism with interaction and quasi-ideal clocks, arXiv:2107.11452.
[21] C. Foti, A. Coppo, G. Barni, A. Cuccoli, and P. Verrucchi, Time and classical equations of motion from quantum entanglement via the Page and Wootters mechanism with generalized coherent states, Nat. Commun. **12**, 1787 (2021).
[22] P. A. Höhn, A. R. Smith, and M. P. Lock, Trinity of relational quantum dynamics, Phys. Rev. D **104**, 066001 (2021).
[23] P. A. Höhn, A. R. Smith, and M. P. Lock, Equivalence of approaches to relational quantum dynamics in relativistic settings, Front. Phys. **9**, 587083 (2021).
[24] F. Lomoc, A. Boette, N. Canosa, and R. Rossignoli, History states of one-dimensional quantum walks, Phys. Rev. A **106**, 062215 (2022).
[25] N. P. D. Loc, Time-system entanglement and special relativity, Mod. Phys. Lett. A **39**, 2350183 (2024).
[26] I. L. Paiva, A. Te'eni, B. Y. Peled, E. Cohen, and Y. Aharonov, Non-inertial quantum clock frames lead to non-Hermitian dynamics, Commun. Phys. **5**, 298 (2022).
[27] I. L. Paiva, A. C. Lobo, and E. Cohen, Flow of time during energy measurements and the resulting time-energy uncertainty relations, Quantum **6**, 683 (2022).
[28] T. Favalli and A. Smerzi, Peaceful coexistence of thermal equilibrium and the emergence of time, Phys. Rev. D **105**, 023525 (2022).
[29] S. Rijavec, Heisenberg-picture evolution without evolution, Phys. Rev. D **108**, 063507 (2023).
[30] V. Baumann, M. Krumm, P. A. Guérin, and Č. Brukner, Noncausal Page-Wootters circuits, Phys. Rev. Res. **4**, 013180 (2022).
[31] L. Apadula, E. Castro-Ruiz, and Č. Brukner, Quantum reference frames for Lorentz symmetry, Quantum **8**, 1440 (2024).
[32] S. Barison, F. Vicentini, I. Cirac, and G. Carleo, Variational dynamics as a ground-state problem on a quantum computer, Phys. Rev. Res. **4**, 043161 (2022).
[33] V. Giovannetti, S. Lloyd, and L. Maccone, Geometric event-based quantum mechanics, New J. Phys. **25**, 023027 (2023).
[34] D. Chester, X. D. Arsiwalla, L. H. Kauffman, M. Planat, and K. Irwin, Quantization of a new canonical, covariant, and symplectic Hamiltonian density, Symmetry **16**, 316 (2024).
[35] J. J. Sakurai and E. D. Commins, *Modern Quantum Mechanics* (Cambridge University Press, Cambridge, UK, 1995), revised edition.
[36] C. J. Isham, Canonical quantum gravity and the problem of time, *Integrable Systems, Quantum Groups, and Quantum Field Theories* (Springer, Netherlands, 1993), p. 157.
[37] R. Gambini, R. A. Porto, J. Pullin, and S. Torterolo, Conditional probabilities with Dirac observables and the problem of time in quantum gravity, Phys. Rev. D **79**, 041501 (2009).
[38] K. V. Kuchař, Time and interpretations of quantum gravity, Int. J. Mod. Phys. D **20**, 3 (2011).
[39] W. Pauli, *Die allgemeinen prinzipien der wellenmechanik* (Springer, Berlin, 1933).
[40] D. N. Page and W. K. Wootters, Evolution without evolution: Dynamics described by stationary observables, Phys. Rev. D **27**, 2885 (1983).
[41] A. Connes and C. Rovelli, Von Neumann algebra automorphisms and time-thermodynamics relation in generally covariant quantum theories, Classical Quantum Gravity **11**, 2899 (1994).
[42] C. J. Isham, Quantum logic and the histories approach to quantum theory, J. Math. Phys. **35**, 2157 (1994).
[43] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, Sci. Rep. **5**, 18281 (2015).
[44] J. Cotler, C.-M. Jian, X.-L. Qi, and F. Wilczek, Superdensity operators for spacetime quantum mechanics, J. High Energy Phys. 09 (2018) 093.
[45] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, Learning the quantum algorithm for state overlap, New J. Phys. **20**, 113022 (2018).
[46] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).
[47] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, Science **364**, 260 (2019).
[48] B. Commeau, M. Cerezo, Z. Holmes, L. Cincio, P. J. Coles, and A. Sornborger, Variational Hamiltonian diagonalization for dynamical quantum simulation, arXiv:2009.02559.
[49] E. Kökcü, T. Steckmann, Y. Wang, J. Freericks, E. F. Dumitrescu, and A. F. Kemper, Fixed depth Hamiltonian simulation via Cartan decomposition, Phys. Rev. Lett. **129**, 070501 (2022).
[50] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).
[51] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Simulating large quantum circuits on a small quantum computer, Phys. Rev. Lett. **125**, 150504 (2020).
[52] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T.

Menke *et al.*, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[53] P. J. Ollitrault, S. Jandura, A. Miessen, I. Burghardt, R. Martinazzo, F. Tacchino, and I. Tavernelli, Quantum algorithms for grid-based variational time evolution, Quantum **7**, 1139 (2023).

[54] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

[55] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, Contemp. Phys. **56**, 172 (2015).

[56] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, Challenges and opportunities in quantum machine learning, Nat. Comput. Sci. **2**, 567 (2022).

[57] A. Štrkalj, E. V. Doggen, I. V. Gornyi, and O. Zilberberg, Many-body localization in the interpolating Aubry-André-Fibonacci model, Phys. Rev. Res. **3**, 033257 (2021).

[58] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[59] A. S. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, Quantum systems equilibrate rapidly for most observables, Phys. Rev. E **90**, 012121 (2014).

[60] G. Mussardo, Infinite-time average of local fields in an integrable quantum field theory after a quantum quench, Phys. Rev. Lett. **111**, 100401 (2013).

[61] L. C. Venuti and P. Zanardi, Universality in the equilibration of quantum systems after a small quench, Phys. Rev. A **81**, 032113 (2010).

[62] L. C. Venuti, N. T. Jacobson, S. Santra, and P. Zanardi, Exact infinite-time statistics of the Loschmidt echo for a quantum quench, Phys. Rev. Lett. **107**, 010403 (2011).

[63] A. Goussev, R. A. Jalabert, H. M. Pastawski, and D. A. Wisniacki, Loschmidt echo, Scholarpedia **7**, 11687 (2012), revision #127578.

[64] J. Yang and A. Hamma, Many-body localization transition, temporal fluctuations of the Loschmidt echo, and scrambling, arXiv:1702.00445.

[65] B. Zhou, C. Yang, and S. Chen, Signature of a nonequilibrium quantum phase transition in the long-time average of the Loschmidt echo, Phys. Rev. B **100**, 184313 (2019).

[66] G. Rickayzen, *Green's Functions and Condensed Matter* (Courier Corporation, London, 2013).

[67] E. Khatami, G. Pupillo, M. Srednicki, and M. Rigol, Fluctuation-dissipation theorem in an isolated system of quantum dipolar bosons after a quench, Phys. Rev. Lett. **111**, 050403 (2013).

[68] D. J. Luitz and Y. B. Lev, Anomalous thermalization in ergodic systems, Phys. Rev. Lett. **117**, 170404 (2016).

[69] E. Kökcü, H. A. Labib, J. K. Freericks, and A. F. Kemper, A linear response framework for quantum simulation of bosonic and fermionic correlation functions, Nat Commun **15**, 3881 (2024).

[70] Á. M. Alhambra, J. Riddell, and L. P. García-Pintos, Time evolution of correlation functions in quantum many-body systems, Phys. Rev. Lett. **124**, 110605 (2020).

[71] J. Pedernales, R. Di Candia, I. Egusquiza, J. Casanova, and E. Solano, Efficient quantum algorithm for computing *n*-time correlation functions, Phys. Rev. Lett. **113**, 020505 (2014).

[72] M. L. Baez, M. Goihl, J. Haferkamp, J. Bermejo-Vega, M. Gluza, and J. Eisert, Dynamical structure factors of dynamical quantum simulators, Proc. Natl. Acad. Sci. USA **117**, 26123 (2020).

[73] R. Coldea, D. Tennant, E. Wheeler, E. Wawrzynska, D. Prabhakaran, M. Telling, K. Habicht, P. Smeibidl, and K. Kiefer, Quantum criticality in an Ising chain: Experimental evidence for emergent E_8 symmetry, Science **327**, 177 (2010).

[74] C. Jia, E. Nowadnick, K. Wohlfeld, Y. Kung, C.-C. Chen, S. Johnston, T. Tohyama, B. Moritz, and T. Devereaux, Persistent spin excitations in doped antiferromagnets revealed by resonant inelastic light scattering, Nat. Commun. **5**, 3314 (2014).

[75] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Hybrid quantum-classical approach to correlated materials, Phys. Rev. X **6**, 031045 (2016).

[76] J. Kreula, S. R. Clark, and D. Jaksch, Non-linear quantum-classical scheme to simulate non-equilibrium strongly correlated fermionic many-body dynamics, Sci. Rep. **6**, 32940 (2016).

[77] R. Sakurai, W. Mizukami, and H. Shinaoka, Hybrid quantum-classical algorithm for computing imaginary-time correlation functions, Phys. Rev. Res. **4**, 023219 (2022).

[78] H. Buhrman, R. Cleve, J. Watrous, and R. De Wolf, Quantum fingerprinting, Phys. Rev. Lett. **87**, 167902 (2001).

[79] A. W. Harrow and A. Montanaro, Testing product states, quantum Merlin-Arthur games and tensor optimization, J. ACM **60**, 1 (2013).

[80] G. Gutoski, P. Hayden, K. Milner, and M. M. Wilde, Quantum interactive proofs and the complexity of separability testing, TheORY Comput. **11**, 59 (2015).

[81] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[82] For infinite T, one has to consider subtleties related to the normalization of states; see Refs. [7,11,12].

[83] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Rev. **41**, 303 (1999).

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

[85] 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).

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

[87] R. D. Somma, A Trotter-Suzuki approximation for Lie groups with applications to Hamiltonian simulation, J. Math. Phys. **57**, 062202 (2016).

[88] If the duration of time required to implement the gates depends on the interval ε, the reasoning can be adapted by changing α. For example, following Ref. [84] one may assume that implementing $e^{-ih_j\varepsilon}$ takes $\varepsilon = \tau/t$ time and we need $\propto t$ gates, meaning a total duration proportional to τ, rather than quadratic as the number of gates.

[89] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, Proc. Natl. Acad. Sci. USA **115**, 9456 (2018).

[90] M. Heyl, P. Hauke, and P. Zoller, Quantum localization bounds Trotter errors in digital quantum simulation, Sci. Adv. **5**, eaau8342 (2019).

[91] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[92] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 1 (2018).

[93] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[94] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).

[95] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Barren plateaus in variational quantum computing, Nat. Rev. Phys. **7**, 174 (2025).

[96] J. Kübler, S. Buchholz, and B. Schölkopf, The inductive bias of quantum kernels, Adv. Neural Inf. Process. Syst. **34**, 12661 (2021).

[97] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, Diagnosing barren plateaus with tools from quantum optimal control, Quantum **6**, 824 (2022).

[98] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, Group-invariant quantum machine learning, PRX Quantum **3**, 030341 (2022).

[99] D. Camps, E. Kökcü, L. Bassman Oftelie, W. A. De Jong, A. F. Kemper, and R. Van Beeumen, An algebraic quantum circuit compression algorithm for Hamiltonian simulation, SIAM J. Matrix Anal. Appl. **43**, 1084 (2022).

[100] E. Kökcü, D. Camps, L. B. Oftelie, J. K. Freericks, W. A. de Jong, R. Van Beeumen, and A. F. Kemper, Algebraic compression of quantum circuits for Hamiltonian evolution, Phys. Rev. A **105**, 032420 (2022).

[101] R. Zeier and T. Schulte-Herbrüggen, Symmetry principles in quantum systems theory, J. Math. Phys. **52**, 113510 (2011).

[102] R. Jozsa and A. Miyake, Matchgates and classical simulation of quantum circuits, Proc. R. Soc. A **464**, 3089 (2008).

[103] K. Wan, W. J. Huggins, J. Lee, and R. Babbush, Matchgate shadows for fermionic quantum simulation, Commun. Math. Phys. **404**, 629 (2023).

[104] N. L. Diaz, D. García-Martín, S. Kazi, M. Larocca, and M. Cerezo, Showcasing a barren plateau theory beyond the dynamical Lie algebra, arXiv:2310.11505.

[105] X. You, S. Chakrabarti, and X. Wu, A convergence theory for over-parameterized variational quantum eigensolvers, arXiv:2205.12481.

[106] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp *et al.*, Does provable absence of barren plateaus imply classical simulability? Nat. Commun. **16**, 7907 (2025).

[107] S. Aubry, Analyticity breaking and Anderson localization in incommensurate lattices, Ann. Isr. Phys. Soc. **3**, 133 (1980).

[108] S. Iyer, V. Oganesyan, G. Refael, and D. A. Huse, Many-body localization in a quasiperiodic system, Phys. Rev. B **87**, 134202 (2013).

[109] M. Larocca, N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo, Theory of overparametrization in quantum neural networks, Nat. Comput. Sci. **3**, 542 (2023).

[110] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, Ann. Phys. **349**, 117 (2014).

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

[112] J. Gibbs and L. Cincio, Deep circuit compression for quantum dynamics via tensor networks, Quantum **9**, 1789 (2025).

[113] Y. Zhang, R. Wiersema, J. Carrasquilla, L. Cincio, and Y. B. Kim, Scalable quantum dynamics compilation via quantum machine learning, arXiv:2409.16346.

[114] J. Gibbs and L. Cincio, Learning circuits with infinite tensor networks, arXiv:2506.02105.

[115] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, Generalization in quantum machine learning from few training data, Nat. Commun. **13**, 4919 (2022).

[116] M. C. Caro, H.-Y. Huang, N. Ezzell, J. Gibbs, A. T. Sornborger, L. Cincio, P. J. Coles, and Z. Holmes, Out-of-distribution generalization for learning quantum dynamics, Nat. Commun. **14**, 3751 (2023).

[117] S. Yoshida, A. Soeda, and M. Murao, Reversing unknown qubit-unitary operation, deterministically and exactly, Phys. Rev. Lett. **131**, 120602 (2023).

[118] Y.-A. Chen, Y. Mo, Y. Liu, L. Zhang, and X. Wang, Quantum algorithm for reversing unknown unitary evolutions, arXiv:2403.04704.

[119] N. L. Diaz, J. M. Matera, and R. Rossignoli, Spacetime quantum and classical mechanics with dynamical foliation, Phys. Rev. D **109**, 105008 (2024).

[120] N. L. Diaz and R. Rossignoli, Spacetime quantum mechanics for bosonic and fermionic systems, arXiv:2506.10250.

[121] C. Kiefer, *Quantum Gravity*, *International Series of Monographs on Physics* (Oxford University Press,, Oxford, 2012), 3rd ed.

[122] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory: Volume 2, Loop amplitudes, Anomalies and Phenomenology* (Cambridge University Press, Cambridge, 2012).

[123] B. S. DeWitt, Quantum theory of gravity. I. The canonical theory, Phys. Rev. **160**, 1113 (1967).

[124] P. A. M. Dirac, Generalized Hamiltonian dynamics, Can. J. Math. **2**, 129 (1950).

[125] P. A. M. Dirac, The theory of gravitation in Hamiltonian form, Proc. R. Soc. London A **246**, 333 (1958).

[126] This is a common example in both canonical quantum gravity [121] and string theory [122] textbooks.

[127] D. Marolf, Quantum observables and recollapsing dynamics, Classical Quantum Gravity **12**, 1199 (1995).

[128] C. Isham and N. Linden, Continuous histories and the history group in generalized quantum theory, J. Math. Phys. **36**, 5392 (1995).

[129] K. Doi, J. Harper, A. Mollabashi, T. Takayanagi, and Y. Taki, Timelike entanglement entropy, J. High Energy Phys. 05 (2023) 052.

[130] K. Narayan and H. K. Saini, Notes on time entanglement and pseudo-entropy, arXiv:2303.01307.

[131] C. S. Chu and H. Parihar, Time-like entanglement entropy in AdS/BCFT, J. Math. Phys. **2023**, 173 (2023).

[132] Z. Zhao, R. Pisarczyk, J. Thompson, M. Gu, V. Vedral, and J. F. Fitzsimons, Geometry of quantum correlations in spacetime, Phys. Rev. A **98**, 052312 (2018).

[133] P. A. Hoehn, A. Russo, and A. R. Smith, Matter relative to quantum hypersurfaces, arXiv:2308.12912.

[134] F. Anza, C. Gogolin, and M. Huber, Eigenstate thermalization for degenerate observables, Phys. Rev. Lett. **120**, 150603 (2018).

[135] M. M. Wilde, From classical to quantum Shannon theory, *Quantum Information Theory* (Cambridge University Press, 2017).

[136] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, Ann. Phys. **16**, 407 (1961).

[137] N. Canosa and R. Rossignoli, Entanglement between distant qubits in cyclic *xx* chains, Phys. Rev. A **75**, 032350 (2007).

[138] M. Wilkinson, Critical properties of electron eigenstates in incommensurate systems, Proc. R. Soc. London, Ser. A **391**, 305 (1984).

[139] Y. G. Sinai, Anderson localization for one-dimensional difference Schrödinger operator with quasiperiodic potential, J. Stat. Phys. **46**, 861 (1987).

[140] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, SIAM Rev. **59**, 65 (2017).

[141] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)* (San Diego, 2015).

[142] J. Hubbard, Electron correlations in narrow energy bands, Proc. R. Soc. London A **276**, 238 (1963).

[143] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[144] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, Phys. Rev. Lett. **91**, 147902 (2003).

[145] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863 (1992).