Quantum dynamics of thermalizing systems

Christopher David White,[1] Michael Zaletel,[2] Roger S. K. Mong,[3] and Gil Refael[1]

[1]*Institute for Quantum Information and Matter, Caltech**

[2]*Department of Physics, Princeton University, Princeton, NJ 08544, USA*

[3]*Department of Physics and Astronomy, University of Pittsburgh*

We introduce a method "DMT" for approximating density operators of 1D systems that, when combined with a standard framework for time evolution (TEBD), makes possible simulation of the dynamics of strongly thermalizing systems to arbitrary times. We demonstrate that the method performs well for both near-equilibrium initial states (Gibbs states with spatially varying temperatures) and far-from-equilibrium initial states, including quenches across phase transitions and pure states.
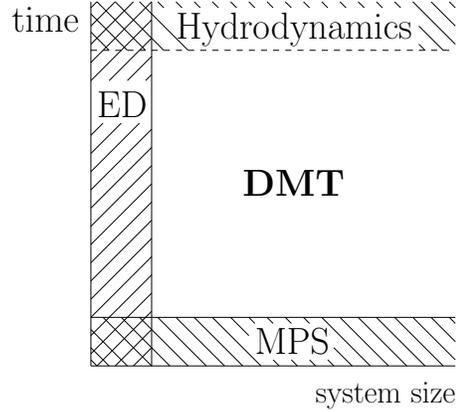
CONTENTS

* cdwhite@caltech.edu

FIG. 1. For small systems, exact diagonalization and related methods can treat time evolution of small ETH systems to long times. Matrix product state methods, on the other hand, can treat time evolution of large ETH systems, but only to short times, and hydrodynamic effective theories can phenomenologically describe the long-time limit, but not intermediate times. Our method, DMT, can treat large ETH systems at all times.

I. INTRODUCTION

Questions about how (and whether) hydrodynamic behavior emerges from microscopic quantum physics arise frequently in condensed matter physics. The exploration of this physics is hampered by the limitations of existing numerical methods (cf Fig. 1). Numerically exact methods (like exact diagonalization and Krylov subspace methods) can treat the dynamical properties of small systems at arbitrary times, but require memory and computation time exponential in system size. Matrix product state methods, on the other hand, can treat large systems—but only when the systems have little entanglement entropy. This means that for thermalizing systems, whose entanglement entropy grows linearly with time, MPS methods can only treat short-time behavior (see Figure 1).

We introduce a numerical method ("density matrix truncation" or "DMT") based on matrix product representations of density operators. This algorithm can accurately simulate not only short-time, low-entanglement

behavior and long-time, hydrodynamic behavior, but also the complex intermediate-time behavior from which the hydrodynamics emerges. While using matrix-product representations of mixed states is not new, the core of our algorithm is a new method for truncating matrix product density operators (MPDOs). This truncation exactly preserves the energy density of the system and other local conserved quantities, with the aim of leaving the hydrodynamics unaffected. It also avoids generating unphysical density matrices with negative entropy. At the same time, it is efficient enough that by taking large (but constant-in-system-size) bond dimension one can capture the thermalization process.

We first (in Sec. II) offer some background on matrix product state methods and intuition for why a method using MPDOs should be able to efficiently simulate time evolution governed by Hamiltonians satisfying the eigenstate thermalization hypothesis (ETH) to arbitrary times. We also motivate certain properties of our method. We then describe (in Sec. III) our algorithm for time evolution. This algorithm consists of a standard time-evolution framework very much like that in TEBD (described in Appendix A), paired with a novel scheme for truncating MPDOs. We then apply our algorithm to time evolution starting from a pure state (Sec. IV B) and find that it qualitatively improves upon existing methods. Applied to a mixed state (Sec. IV C and App. D), we find that DMT matches or exceeds the state of the art. We conclude with directions we hope will improve on the method.

II. BACKGROUND AND INTUITION

A. Background: matrix product state methods

Simulating real-time evolution of many-body quantum-mechanical systems is hard: a system of L sites generically requires storage and computation time exponential in L. One line of attack, e.g. time-evolving block decimation (TEBD) [1–3], proceeds by representing unitary time evolution as a series of small time-steps applied to a matrix-product state (MPS) representation of a pure state. These matrix-product structures offer efficient representations of certain states (broadly speaking, "low-entanglement states") in the sense that typical operations require polynomial time and memory.

MPS simulations of time-evolving pure states are stymied by the fast increase of entanglement entropy with time, which grows linearly in time for a typical global quench. When one compresses a pure state as an MPS, memory and computation time requirements grow exponentially in the entanglement entropy among the subsystems, so this linear growth in entanglement entropy sets a hard upper limit on the timescales on which MPSs are useful, though a variety of methods have been used in attempts to circumvent this limit. [4–15]. One

case in which entanglement growth does not limit the useful timescale of MPS simulations is many-body localized (MBL) dynamics, which exhibits a modest logarithmic entanglement growth [16, 17]. On the thermal side of the MBL transition, however, entanglement grows quickly, and even the transition itself is expected to show extensive (volume-law) entanglement [18]—consequently, pure-state time evolution under Hamiltonians remains restricted to short times.

We study time evolution under Hamiltonians that are not many-body localized, for which the long time behavior shows extensive (volume-law) entanglement. We note a very interesting perspective on this barrier was investigated in the recent work of Leviatan et al. [19]. They found that by continuing an energy-conserving version of the time evolution despite the large errors associated with staying in the MPS manifold, certain of the statistical aspects of the thermalization process were still correctly captured.

Research into mixed-state time evolution and Lindblad dynamics has also progressed. It has been proven that density matrices (and purifications) of Gibbs states of local Hamiltonians have efficient matrix product representations [20–23]. Two schools of thought have used this insight to develop a series of methods for simulating time evolution. One school employs density matrices [20, 24–38]. They note that the space of operators on a spin chain is the tensor product of onsite operator spaces, just as the space of many-body pure states being a tensor product on onsite Hilbert spaces; the chief difference (in this view) is merely the dimensionality of the onsite space. For example, on a spin-half chain, the space of onsite operators is four-dimensional, while the space of pure states is two dimensional. This school then applies familiar pure state methods, including the creation and truncation of matrix product states and time evolution by TEBD, to density matrices—which are, after all, merely vectors in a larger space. The resulting truncation algorithms minimize the error according to the Hilbert-Schmidt (Frobenius) norm. In certain situations—in particular, dynamics near thermal equilibrium or a non-equilibrium steady state—this approach works well. In other situations, however—in particular, time evolution starting from a pure state—the density matrices suffer from a catastrophic loss of positivity. (Even checking positivity is NP-hard in the system size [39].)

The second school [9, 40–48] uses purifications instead of density matrices to represent mixed states. They pair each site in the system with an ancilla, a notional site representing the bath. The mixed nature of the system is represented by entanglement between sites and their ancillae, so the system and ancillae together are in a pure state. Grouping each site and its ancilla into a larger onsite Hilbert space, one can write a matrix product representation for this pure state and apply the usual methods (truncation, TEBD, etc.) This solves the positivity problem: unlike density matrices, where many operators with reasonable matrix product representations are not

positive and hence are invalid as density matrices, every representable vector is a valid state. Moreover, since one can act with a unitary on the space of ancillae without changing the physical state, one can try to exploit this freedom to reduce the MPS bond dimension of the purification [43, 44]. There is also a hybrid approach which locally unzips a density matrix into a purification, which preserves positivity by construction [49]. These purification methods employ truncations which minimize error according to the inner product $\langle \cdot | \cdot \rangle$ defined on the whole (system with ancillae) state.

We argue that neither the Frobenius norm on density matrices nor the quantum-mechanical norm on purifications is the correct notion of error. In the case of density matrices, the Frobenius norm fails to account for the fact that truncations that change the component of the density-matrix vector along the identity (e.g., which are not trace-preserving) are disastrous, because they can lead to loss of positivity. Moreover, neither notion of error captures spatial locality: a good notion of error should prioritize short-range properties of the model and guarantee that certain quantities (the local conserved quantities of the model under consideration, like energy density or spin) are unchanged. Since the methods of both the density-matrix and purification schools generically change the model's conserved quantities at every gate application, they are unable in principle to approach the known "hydrodynamic" long-time behavior of systems which thermalize. This may be the reason that existing density-matrix methods lose accuracy over time, even though one would expect the accuracy of the matrix-product representation to improve as the state approaches equilibrium.

In this work we propose a truncation of density matrices that ameliorates the positivity problem of Frobenius truncation and exactly preserves the expectation values of all operators on all regions of up to three sites in diameter.

B. Intuition: thermalization and computation

Why should one be able to efficiently simulate the dynamics of a local ETH Hamiltonian? In the long-time limit, the system is well described (as far as local observables are concerned) by a Gibbs state, which has an efficient matrix product density operator (MPDO) representation [21, 22]. Moreover, the system will (one expects) locally thermalize before it reaches global equilibrium, and indeed after some short local thermalization time t_{therm} expectation values of local operators will be well approximated by the expectation values of those operators in a Gibbs state with spatially varying temperature, chemical potential, etc. Heuristically, one can imagine keeping the state exactly out to the local thermalization time and then truncating to an efficient representation. This would require a maximum bond dimension ca. $(d^2)^{vt_{\text{therm}}}$, where v is some entanglement speed and

d is the dimension of the onsite Hilbert space. If vt_{therm} is not too large, this approach itself may be workable.

In practice, however, one will wish to efficiently represent the state even at early and intermediate times $t < t_{\text{therm}}$—and also to avoid dependence on the hard-to-define and likely-unknown constant t_{therm}. Having decided upon an MPDO representation, then, one is faced with the problem of writing a truncation algorithm: an algorithm that will approximate a given MPDO by another, more compact MPDO.

The natural approach, by analogy with matrix product states, is to discard low-weight Schmidt vectors. (This approach turns out to be an imperfect solution, but it is a useful first step.) We call this truncation "Frobenius truncation." A density operator is a vector in a space with the same tensor-product structure as a state, but a larger onsite dimension. We can therefore cut the chain at bond j into two sections L and R, Schmidt decompose, and truncate it:

$$\rho = \sum_{\alpha=0}^{\chi-1} \hat{x}_{L\alpha} s_\alpha \hat{x}_{R\alpha} \mapsto \sum_{\alpha=0}^{\chi'-1} \hat{x}_{L\alpha} s_\alpha \hat{x}_{R\alpha}, \qquad \chi' < \chi \quad (1)$$

where $\hat{x}_{L\alpha}, \hat{x}_{R\alpha}$ are operators supported on L and R, respectively, and $\operatorname{tr} x_{L\alpha}^\dagger x_{L\beta} = \operatorname{tr} x_{R\alpha}^\dagger x_{R\beta} = \delta_{\alpha\beta}$. Explicitly, one starts with a matrix-product representation of the density operator

$$\rho = \sum_{\alpha=0}^{\chi-1} \left[A_1^{\mu_1} \cdots A_j^{\mu_j} \right]_\alpha s_\alpha \left[B_{j+1}^{\mu_{j+1}} \cdots B_L^{\mu_L} \right]_\alpha \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_L^{\mu_L}, \tag{2}$$

where $A_l^{\mu_l}, B_l^{\mu_l}$ are $\chi \times \chi$ matrices—with the exception of $A_1^{\mu_1}$ and $B_L^{\mu_L}$, which are $1 \times \chi$ and $\chi \times 1$ respectively. We suppress for compactness a sum on μ (that is, on a basis for the space on onsite operators, here the Pauli matrices σ^μ, with $\sigma^0 = I$). The truncation (1) is then

$$\rho = \sum_{\alpha=0}^{\chi-1} \left[\cdots A_j^{\mu_j} \right]_\alpha s_\alpha \left[B_{j+1}^{\mu_{j+1}} \cdots \right]_\alpha \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_L^{\mu_L}$$
$$\mapsto \sum_{\alpha=0}^{\chi'-1} \left[\cdots A_j^{\mu_j} \right]_\alpha s_\alpha \left[B_{j+1}^{\mu_{j+1}} \cdots \right]_\alpha \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_L^{\mu_L}, \tag{3}$$

with $\chi' < \chi$. This approximation minimizes the Frobenius (Hilbert-Schmidt) norm distance—but a priori that is not the only norm one could use.

The Frobenius approximation scheme gives poor results for initial states far from equilibrium. One can see why by considering the expectation values of the operator $O_{lt}^y = U(t) \sigma_l^y U(t)^\dagger$ for a system that starts at $t = 0$ in a product of σ^y eigenstates. (Note that $O_{l,t}^y$ is a Schrödinger-picture operator parametrized by time t. We work in the Schrödinger picture throughout, except where noted.) At time t, $\langle \psi(t) | O_{l,t}^y | \psi(t) \rangle = \pm 1$—but generically $O_{l,t}^y$ will be a large, complicated operator (if we choose t larger than the whole-system entanglement time, as we are free to do, it will have support throughout the system) and essentially unrelated to the local

operators we wish to measure. There are 2^L such operators $O^y_{l_1 l_2 \ldots, t} = U(t) \sigma^y_{l_1} \sigma^y_{l_2} \cdots U(-t)$, all corresponding to long-range operators with expectation value ± 1. These operators $O^y_{l_1 l_2 \ldots, t}$ form part of an orthonormal basis for the space of operator space. Errors along the dimensions $O^y_{l_1 l_2 \ldots}$ will be penalized by the Frobenius-norm notion of distance with precisely the same severity as errors along more physically relevant dimensions, like σ^y_l. A more reasonable metric for truncation error should be willing to "forget" this information, in favor of more accurately keeping local operators, once they are no longer expected to feedback into the hydrodynamics. (There are more worrying problems still with the naïve Frobenius truncation, which rapidly leads to a dramatic loss of positivity in the supposed density matrix for many far-from-equilibrium initial-conditions, but these problems can be remedied by considering purifications.)

A BBGKY-like hierarchy for the dynamics of reduced density matrix of a spin chain offers a clue as to how to proceed. In a system governed by a Hamiltonian that is the sum of two-site terms $H = \sum_j h_{j,j+1}$, the dynamics of one-site reduced density matrices (say ρ_j, the reduced density matrix on site j) depends on the two site density matrices:

$$\frac{d}{dt} \rho_j = -i \operatorname{tr}_{\{j' \neq j\}} [H, \rho]$$
$$= -i [h_{j,j+1} \rho_{j,j+1}] - i [h_{j-1,j} \rho_{j-1,j}] \qquad (4)$$

where we write $\rho_{j,j+1}$ for the two-site reduced density matrix on sites $j, j+1$. Meanwhile the two-site reduced density matrices depend on three-site density matrices, the three-site on four-site, and so on up the sequence. One can imagine truncating this hierarchy at some length l—that is, tracking l-site reduced density matrices for some l, say 2 or 3 or 6, and writing the dynamics for the l-site density matrices in terms of some approximation for the $l+1$-site density matrices. A natural choice for such an approximation is to replace the $l+1$-site matrices by their disconnected component, e.g., for $l = 1$, take $\rho_{j,j+1} \sim \rho_j \rho_{j+1}$. The problem of the operators $O^y_{l,t}$ then never arises.

The BBGKY-like approach fixes certain problems with Frobenius truncation, but comes with its own set of problems. It is not obvious that the l-site density matrices will even be consistent, in the sense that it may not be possible to write them as reduced density matrices for a density matrix on the whole system. (Checking this, for a given set of reduced density matrices, is QMA-hard [50].) Moreover, the hard truncation at l sites may not be appropriate to capture the dynamics of the system. Longer-range operators may feed back into the dynamics of few-site density matrices via the hierarchy starting with (4)—and conversely, some short-range operators may have a negligible effect on the dynamics of operators of physical interest.

The Frobenius norm attempts to keeps all operators with equal weight, while BBGKY keeps connected components only up to a hard cutoff. A natural compro-
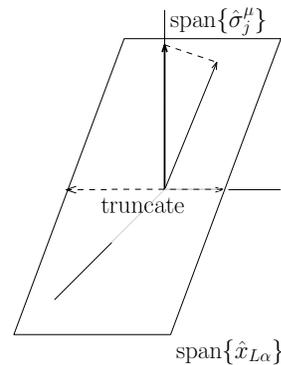

FIG. 2. DMT, viewed in the space of operators on the left half of the chain. We truncate perpendicular to certain physically-relevant operators (the σ^μ_j).

mise is to interpolate between the two by weighting the connected components of an operator according to some measure of locality. In the current work, we take the first step in this direction: we approximate the whole-system density matrices in such a way that the dynamics of reduced density matrices on up to three sites matches the BBGKY hierarchy, but instead of straightforwardly closing the BBGKY hierarchy at that level, we approximate larger connected components using a method similar in spirit to the Frobenius truncation (1). Our method zeros out long-range correlations, replacing entanglement between different parts of the system by entanglement with a notional bath.

Although we have used thermalizing Hamiltonians to motivate our method, our method does not in fact assume that the Hamiltonian governing the dynamics is thermalizing. We expect to be able to use it to treat MBL Hamiltonians, with a more stringent accuracy vs. bond dimension tradeoff.

III. METHOD: TRUNCATION OF MPDOS

Given an MPDO and a particular bond j, we wish to truncate the rank of the MPDO. How can we modify the Frobenius truncation (1) in such a way that it does not change the trace of the density matrix, nor the expectation values of local operators? The trick is to start by Schmidt decomposing the whole density matrix, and then cleverly choose basis changes on the spaces of left and right Schmidt vectors that put the data we want to avoid changing in specific, easily-understood locations (see Fig. 2). We can then slot our new truncation into a (slight modification of a) standard time-evolution framework, TEBD (see Appendix A).

Our truncation algorithm guarantees that the following will not change, up to the precision of the numerical linear algebra involved, in a truncation on bond j (cf Fig. 3):

1. the trace of the density matrix, $\operatorname{tr} \rho$;

2. the reduced density matrix $\rho_{1\cdots j+1}$ on sites $1, \ldots, (j+1)$; and

3. the reduced density matrix $\rho_{j\cdots L}$ on sites $j, \ldots, L$.

Consequently, no truncation will change the expectation of any operator on three contiguous sites, because any such operator is always contained within one of the guaranteed-preserved density operators.

These guarantees do not fully specify our truncation method. To do so, define a matrix of connected correlators across the cut j:

$$\tilde{M}_{\alpha\beta} = \langle \hat{y}_{L\alpha} \hat{y}_{R\beta} \rangle - \langle \hat{y}_{L\alpha} \rangle \langle \hat{y}_{R\beta} \rangle \tag{5}$$

with $\hat{y}_{L\alpha}$ and $\hat{y}_{R\beta}$ operators supported on sites $1, \ldots, j$ and $j+1, \ldots, L$ respectively. (The $\hat{y}$'s spans the set of observables and are defined below.) We wish to replace this matrix $\tilde{M}$ by another $\tilde{M}'$ with lower rank such that $\mathrm{tr}(\tilde{M} - \tilde{M}')^\dagger (\tilde{M} - \tilde{M}')$ is minimized subject to the constraints above. The focus on connected components is itself an important improvement.

A. Setting, notation, and tools

The concept of the method may be straightforward, but it is obscured by a flurry of notation. We start truncation on bond j with an MPDO of the form

$$\rho = \sum_{\alpha=0}^{\chi-1} \sum_{\{\mu\}} \left[A_1^{\mu_1} \cdots A_j^{\mu_j} \right]_\alpha s_\alpha \left[B_{j+1}^{\mu_{j+1}} \cdots B_L^{\mu_L} \right]_\alpha \\ \times \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_L^{\mu_L}, \tag{6}$$

on an L-site chain. The $A_l^{\mu_l}, B_l^{\mu_l}$ are $\chi \times \chi$ matrices—with the exception of $A_1^{\mu_1}$ and $B_L^{\mu_L}$, which are $1 \times \chi$ and $\chi \times 1$ respectively. (χ, called the bond dimension, will, in fact, vary between bonds and between steps of time-evolution and truncation, but for the moment we suppress this variation.)

Truncation at one bond preserves

this density matrix

this density matrix

Truncation at each bond in succession preserves
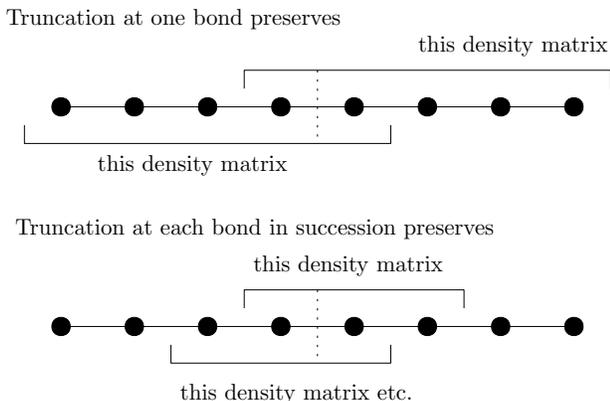
this density matrix

this density matrix etc.

FIG. 3. The reduced density matrices that are guaranteed to be preserved under truncation.

In writing our truncation algorithm, we hat our operators [51]. We use Roman letters (frequently j, l) to index sites and bonds; a bond inherits the index of the site to its left. We use Greek letters (frequently α, β, γ—but excepting μ and χ) for the virtual index labeling the Schmidt vectors. The Greek letter $\mu = 0, 1, 2, 3$ is used to label Pauli matrices, i.e., $\hat{\sigma}_j^\mu$ is an operator at site j (with $\hat{\sigma}^0 = I, \hat{\sigma}^1 = \hat{\sigma}^x$, etc.).

Following the standard notation [52], the MPDO is in mixed-canonical form with an orthogonality center at site j—that is, for any $j_1 \leq j$ and $j_2 \geq j$, the operators

$$\hat{x}_{L\alpha}[j_1] = \sum_{\{\mu\}} \left[A_1^{\mu_1} \cdots A_{j_1}^{\mu_{j_1}} \right]_\alpha \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_{j_1}^{\mu_{j_1}}, \\ \hat{x}_{R\alpha}[j_2] = \sum_{\{\mu\}} \left[B_{j_2+1}^{\mu_{j_2+1}} \cdots B_L^{\mu_L} \right]_\alpha \hat{\sigma}_{j_2+1}^{\mu_{j_2+1}} \cdots \hat{\sigma}_L^{\mu_L} \tag{7}$$

are orthogonal with respect to the Frobenius inner product

$$\mathrm{tr}[x_{L\alpha}^\dagger x_{L\beta}] = \mathrm{tr}[x_{R\alpha}^\dagger x_{R\beta}] = \delta_{\alpha\beta}.$$

This mixed-canonical form gives the Schmidt decomposition of the density matrix ρ at bond j:

$$\rho = \sum_{\alpha=0}^{\chi-1} \hat{x}_{L\alpha}[j] \; s_\alpha \; \hat{x}_{R\alpha}[j]. \tag{8}$$

Henceforth, we will implicitly always be working with Schmidt vectors at bond j, and drop the bond label as follows:

$$\hat{x}_{L\alpha} = \hat{x}_{L\alpha}[j], \\ \hat{x}_{R\alpha} = \hat{x}_{R\alpha}[j]. \tag{9}$$

The two vector spaces $\mathrm{span}\{\hat{x}_{L\alpha}\}$ and $\mathrm{span}\{\hat{x}_{R\alpha}\}$ are the setting in which we work.

We frequently abuse notation by replacing s for a diagonal matrix whose entries are s_α. This allows us to shorten Eq. (8) into

$$\rho = \hat{x}_L s \hat{x}_R, \tag{10}$$

where $\hat{x}_{L,R}$ are row and column vectors of operators, respectively.

With stage set and notation defined, we can walk through the algorithm.

B. Re-writing the MPDO to expose properties whose preservation we guarantee

We wish to take the MPDO

$$\rho = \sum_{\alpha=0}^{\chi-1} \hat{x}_{L\alpha} s_\alpha \hat{x}_{R\alpha} \tag{11}$$

(cut along bond j) and re-write it as

$$\rho = \sum_{\alpha,\beta=0}^{\chi-1} \hat{y}_{L\alpha} \, M_{\alpha\beta} \, \hat{y}_{R\beta}, \tag{12}$$

with the new bases $\{\hat{y}_{L\alpha}\}, \{\hat{y}_{R\alpha}\}$. The bases $\{\hat{y}_{L\alpha}\}, \{\hat{y}_{R\alpha}\}$ and M are chosen such that the properties we wish to avoid changing are characterized by certain easily-identifiable blocks of M:

1. $\operatorname{tr}\rho$ is independent of $M_{\alpha\beta}$ for $\alpha, \beta \neq 0$.

2. The reduced density matrix on sites $1, \ldots, (j+1)$, $\rho_{1\cdots j+1} = \operatorname{tr}_{\{(j+2)\cdots L\}} \rho$, is independent of $M_{\alpha\beta}$ for $\beta \geq 4$.

3. The reduced density matrix on sites $j, \ldots, L$, $\rho_{j\cdots L} = \operatorname{tr}_{\{1\cdots(j-1)\}} \rho$, is independent of $M_{\alpha\beta}$ for $\alpha \geq 4$.

Once we have made this basis change we will be able to modify $M_{\alpha\beta}, \alpha, \beta \geq 4$ with impunity: no such modification will violate our guarantees.

Consider a change of basis

$$\hat{y}_{L\beta} \equiv (\hat{x}_L Q_L^*)_\beta$$
$$\equiv \sum_{\alpha=0}^{\chi-1} \hat{x}_{L\alpha} Q_{L\alpha\beta}^*$$
$$= \sum_{\alpha,\{\mu\}} \left[A_1^{\mu_1} \ldots A_j^{\mu_j} \right]_\alpha Q_{L\alpha\beta}^* \, \hat{\sigma}_1^{\mu_1} \cdots \hat{\sigma}_j^{\mu_j},$$
$$\hat{y}_{R\beta} \equiv (Q_R^\dagger \hat{x}_R)_\beta \tag{13}$$
$$\equiv \sum_{\alpha=0}^{\chi-1} \hat{Q}_{R\alpha\beta}^* x_{R\alpha}$$
$$= \sum_{\alpha,\{\mu\}} Q_{R\alpha\beta}^* \left[B_{j+1}^{\mu_{j+1}} \ldots B_L^{\mu_L} \right]_\alpha \hat{\sigma}_{j+1}^{\mu_{j+1}} \cdots \hat{\sigma}_L^{\mu_L}.$$

with $Q_{L,R}$ unitary $\chi \times \chi$ matrices. Now write

$$\rho = \hat{x}_L s \hat{x}_R = [\hat{x}_L Q_L^*][Q_L^T s Q_R][Q_R^\dagger \hat{x}_R] = \hat{y}_L M \hat{y}_R, \tag{14}$$

we can see that M is related to s via

$$M = Q_L^T s Q_R. \tag{15}$$

The requisite basis transformations $Q_{L,R}$ are given by QR decompositions

$$Q_{L\alpha\beta} R_{L\beta}{}^\mu = \operatorname{tr}[\hat{x}_{L\alpha} \hat{\sigma}_j^\mu]$$
$$\propto \left[A_1^0 \cdots A_{j-1}^0 A_j^\mu \right]_\alpha \in \mathbf{C}^{\chi \times 4},$$
$$Q_{R\alpha\beta} R_{R\beta}{}^\mu = \operatorname{tr}[\hat{x}_{R\alpha} \hat{\sigma}_{j+1}^\mu] \tag{16}$$
$$\propto \left[B_{j+1}^\mu \cdots B_{L-1}^0 B_L^0 \right]_\alpha \in \mathbf{C}^{\chi \times 4}.$$

(here we use the Einstein summation convention). In this context, the fact that the fact that the $R_{L\beta}{}^\mu$ is upper triangular is exactly the statement that $R_{L\beta}{}^\mu = 0$ for $\beta > \mu$. (Similarly for $R_{R\beta}{}^\mu$.)

To see that this is in fact the basis change we seek, first note the trace relations

$$\operatorname{tr}[\hat{\sigma}_j^\mu \hat{y}_{L\beta}] = \sum_\alpha \operatorname{tr}[\hat{\sigma}_j^\mu \hat{x}_{L\alpha}] Q_{L\alpha\beta}^* = R_{L\beta}{}^\mu,$$
$$\operatorname{tr}[\hat{\sigma}_{j+1}^\mu \hat{y}_{R\beta}] = \sum_\alpha Q_{R\beta\alpha}^\dagger \operatorname{tr}[\hat{\sigma}_{j+1}^\mu \hat{x}_{L\alpha}] = R_{R\beta}{}^\mu. \tag{17}$$

The trace of the density matrix is

$$\operatorname{tr}\rho = \sum_{\alpha,\beta=0}^{\chi-1} (\operatorname{tr}\hat{y}_{L\alpha}) M_{\alpha\beta} (\operatorname{tr}\hat{y}_{R\beta}) = R_{L0}{}^0 M_{00} R_{R0}{}^0. \tag{18}$$

(Recall that $\hat{\sigma}_j^0 = \hat{\sigma}_{j+1}^0 = I$ and that R's are upper triangular.) This shows that $\operatorname{tr}\rho$ is independent of the majority of M, as desired. Similarly, the density matrices on sites $1, \ldots, j+1$ and $j, \ldots, L$ are

$$\rho_{1\cdots j+1} = \sum_{\alpha,\beta=0}^{\chi-1} \hat{y}_{L\alpha} M_{\alpha\beta} \sum_{\mu=0}^{3} \frac{\operatorname{tr}[\hat{y}_{R\beta} \hat{\sigma}_{j+1}^\mu] \hat{\sigma}_{j+1}^\mu}{2}$$
$$= \frac{1}{2} \sum_{\mu=0}^{3} \hat{\sigma}_{j+1}^\mu \sum_{\alpha,\beta=0}^{\chi-1} \hat{y}_{L\alpha} M_{\alpha\beta} R_{R\beta}{}^\mu, \tag{19a}$$

and

$$\rho_{j\cdots L} = \sum_{\alpha,\beta=0}^{\chi-1} \sum_{\mu=0}^{3} \frac{\hat{\sigma}_j^\mu \operatorname{tr}[\hat{\sigma}_j^\mu \hat{y}_{L\alpha}]}{2} M_{\alpha\beta} \hat{y}_{R\beta}$$
$$= \frac{1}{2} \sum_{\mu=0}^{3} \hat{\sigma}_j^\mu \sum_{\alpha,\beta=0}^{\chi-1} R_{L\alpha}{}^\mu M_{\alpha\beta} \hat{y}_{R\beta}. \tag{19b}$$

That is, they depend only on the left four columns and top four rows of M, again as desired.

C. Modifying the MPDO

Working in the $\{y\}$ bases, we can modify $M_{\alpha\beta}$ for $\alpha, \beta \geq 4$ at will without violating our guarantees. We wish to do so in a way that reduces the rank of M while doing the least violence, in some sense, to the connected components of correlations across the cut. Explicitly, we wish to change the quantities $C(A_L, B_R) = \langle \hat{A}_L \hat{B}_R \rangle - \langle \hat{A}_L \rangle \langle \hat{B}_R \rangle$, where $\hat{A}_L, \hat{B}_R$ have support on the left, right portions of the chain respectively, as little as possible.

First, let us see what the correlator involves. Using the definition $\langle \hat{A} \rangle = \frac{1}{Z} \operatorname{tr} \hat{A}\rho$, with $Z \equiv \operatorname{tr}\rho = R_{L0}{}^0 R_{R0}{}^0 M_{00}$, it is easy to see that

$$\langle \hat{A}_L \rangle = \frac{1}{Z} \operatorname{tr}[\hat{y}_{L\alpha} \hat{A}_L] M_{\alpha\beta} R_{R\beta}{}^0 = \frac{1}{Z} a_\alpha M_{\alpha 0} R_{R0}{}^0 \tag{20a}$$

and

$$\langle \hat{B}_R \rangle = \frac{1}{Z} R_{R\alpha}{}^0 M_{\alpha\beta} \operatorname{tr}[\hat{y}_{R\beta} \hat{B}_R] = \frac{1}{Z} R_{L0}{}^0 M_{0\beta} b_\beta. \tag{20b}$$

We define $a_\alpha = \mathrm{tr}[\hat{y}_{L\alpha}\hat{A}_L]$, $b_\alpha = \mathrm{tr}[\hat{y}_{R\alpha}\hat{B}_R]$ for convenience. (Throughout this subsection, we employ Einstein summation notation over the Greek indices.) The expectation value of the product is then:

$$\langle \hat{A}_L \hat{B}_R \rangle = \frac{1}{Z}\mathrm{tr}[\hat{y}_{L\alpha}\hat{A}_L]M_{\alpha\beta}\mathrm{tr}[\hat{y}_{R\beta}\hat{B}_R]$$
$$= \frac{1}{Z}a_\alpha M_{\alpha\beta}b_\beta. \tag{21}$$

Putting all these together we find:

$$C(\hat{A}_L, \hat{B}_R) = \frac{1}{Z}a_\alpha\left(M_{\alpha\beta} - \frac{M_{\alpha 0}M_{0\beta}}{M_{00}}\right)b_\beta. \tag{22}$$

Since we would like to only alter pieces of the density matrix which affect correlations accross the cut at bond j, we will modify the matrix in the parenthesis, and denote it by:

$$\tilde{M}_{\alpha\beta} = M_{\alpha\beta} - \frac{M_{\alpha 0}M_{0\beta}}{M_{00}}. \tag{23}$$

At this point we have pushed one step further the process of re-writing the density matrix so that its structure explicitly reflects the distinction between information we are willing to change and information we are not willing to change, but we still have not truncated it. To carry out the truncation, perform an SVD on the lower right block of $\tilde{M}$, writing $\tilde{M}_{\alpha\beta} = \sum_\gamma X_{\alpha\gamma}r_\gamma Y_{\gamma\beta}$ for $\alpha, \beta \geq 4$. Choose an integer χ' (we will see shortly how it relates to the bond dimension of the final truncated MPDO) and insert a projection $P^{(\chi')}$ onto the largest χ' elements of r to form a new matrix

$$\tilde{M}' = \begin{bmatrix} & | & \\ \hline & | & \\ & XrP^{(\chi')}Y & \end{bmatrix}. \tag{24}$$

$\tilde{M}'$ differs from $\tilde{M}$ only in those elements $\tilde{M}'_{\alpha\beta}$ with $\alpha, \beta \geq 4$—that is, in those elements that encapsulate correlations with range ≥ 2. Moreover, only small elements of M are changed, since we take small (connected) correlations and set them identically to zero.

This truncation results in a new matrix $M'_{\alpha\beta}$ to replace $M_{\alpha\beta}$:

$$M'_{\alpha\beta} = \tilde{M}'_{\alpha\beta} + \frac{M_{\alpha 0}M_{0\beta}}{M_{00}} \tag{25}$$

and then

$$\rho' = \hat{y}_{L\alpha}M'_{\alpha\beta}\hat{y}_{R\beta}. \tag{26}$$

The matrix M' has rank at most $8 + \chi'$ (see App. B).

Since we know M' and (matrix-product representations of) the $x_{L\alpha}$ and $x_{R\beta}$, putting this into MPDO form like (A7) with bond dimension $8 + \chi'$ is a matter of re-arrangement. To rearrange into MPDO form perform a

second singular value decomposition, this time on M', for

$$M' = Us'V. \tag{27}$$

Since M' has rank at most $\chi' + 8$, there will be at most $\chi' + 8$ nonzero singular values s. The density matrix after truncation is then

$$\rho \mapsto \sum_{\{\mu\}}\left[A_1^{\mu_1}\cdots A_j'^{\mu_j}\right]s'\left[B_{j+1}'^{\mu_{j+1}}\cdots B_L^{\mu_L}\right]\hat{\sigma}_1^{\mu_1}\cdots\hat{\sigma}_L^{\mu_L}, \tag{28}$$

with

$$A_j'^{\mu_j} = A_j Q_L^* U,$$
$$B_{j+1}'^{\mu_j} = V Q_L^* B_{j+1}; \tag{29}$$

and the rest of matrices $A_1, \ldots, A_{j-1}, B_{j+2}, \ldots, B_L$ are untouched. Note that, regardless of our choice of χ', the reduced density matrices on sites $1, \ldots, j+1, j, \ldots, L$ are exactly as they were before the truncation.

IV. RESULTS

We apply the method to a variety of initial states, pure and mixed. We time evolve by the boustrophedon Trotter decomposition (A6) of a Hamiltonian known to satisfy the eigenstate thermalization hypothesis with Trotter step $\delta t = 1.0$ (except where specified). We work at a maximum bond dimension cutoff (i.e., at each gate application we truncate to this cutoff using the algorithm described in section III) and measure performance by varying this cutoff.

A. Hamiltonian

We take as our Hamiltonian the spin-1/2 transverse-field Ising model on an L-site chain with open boundary conditions:

$$H = \sum_{j=1}^{L-1}S_j^z S_{j+1}^z + \frac{1}{2}h^x\sum_{j=1}^{L}S^x + \frac{1}{2}h^z\sum_{j=1}^{L}S^z. \tag{30}$$

At $h^z = 0$ the model is integrable (by Jordan-Wigner and Bogoliubov transformations); the longitudinal field $h^z\sum S^z$ breaks integrability, and at $h^z = 0.8090, h^x = 0.9045$ the model is known to satisfy the eigenstate thermalization hypothesis in a strong sense [53]. We work at onsite fields

$$h^z = 0.8090, \quad h^x = 0.9045 \tag{31}$$

(except where otherwise specified). Despite their ETH nature, TFIM Hamiltonians like this can display ill-understood pre-thermalization behavior, thought to be related to long-range emergent conserved or nearly-conserved quantities [54–56]. We do not expect DMT to be able to capture this emergent-integrable behavior (see Section II), so we choose our initial conditions to avoid it.
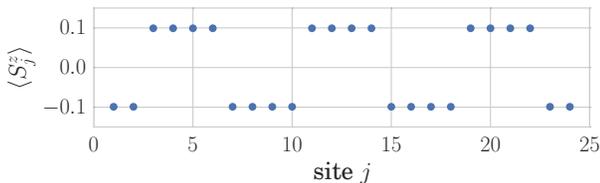
FIG. 4. $\langle S^z \rangle$ for the initial state described in (32) on a 24-site chain.

B. Application: pure-state evolution

We engineer an initial state with a long-wavelength energy variation by taking a product of σ^y eigenstates and rotating blocks of four spins alternately towards $+z$ and $-z$ (cf Fig. 4). The initial state is

$$|\text{near}Y\rangle = \prod_{j=1}^{L} \left[1 + i(1 + g_j)\sigma_j^+\right] |\downarrow\downarrow \ldots \downarrow\rangle \qquad (32)$$

(suitably normalized), where

$$g_j = 0.1 \times \begin{cases} -1 & j \bmod 8 = 1,\, 2,\, 7,\, \text{or } 0, \\ +1 & j \bmod 8 = 3,\, 4,\, 5,\, \text{or } 6. \end{cases} \qquad (33)$$

(We choose the state to be near the σ^y product state in order that we may avoid the pre-thermalization behavior found in [54–56].)

Since the initial state is a product state, it may be represented exactly as an MPO with bond dimension $\chi = 1$. Trotter time evolution increases the bond dimension with each time step δt, but truncation (whatever the algorithm) kicks in only at a time $t_{\text{trunc}}(\chi_{\max})$ when $\chi(t)$ reaches $\chi_{\max}$. Thus for each $\chi_{\max}$ the time evolution is semiexact (that is, exact up to error resulting from the Trotter decomposition of the Hamiltonian) for $t < t_{\text{trunc}}(\chi_{\max})$, at which time it begins to deviate from the semiexact value. This effect appears in all of our results; we also use it to benchmark our method (vide infra).

Figure 5 shows the normalization $Z = \frac{\text{tr}\,\rho}{\sqrt{\text{tr}\,\rho^2}}$ as a function of time. The normalization is related to the second Rényi entropy of the entire chain $S_2 \equiv -\ln \frac{\text{tr}[\rho^2]}{[\text{tr}\,\rho]^2}$ via $S_2 = 2\ln Z$. The DMT produces bath entropy for the system, and this is reflected in the increase of Z as a function of time. In contrast, we find that the ZV method produces non-physical states with $Z < 1$ over the course of time-evolution, which results from negative eigenvalues of the density matrix generated in the truncation. The observation that $Z \geq 1$ for DMT does not imply positive semi-definitity, but suggests that any error arising from the negative eigenvalues is small and well-controlled.

Figure 6 shows the second Rényi entropy of the left half of the chain—that is, the subsystem consisting of sites 1 to $L/2 = 32$. (The von Neumann entropy is difficult to calculate for MPDOs, while second Rényi entropies are
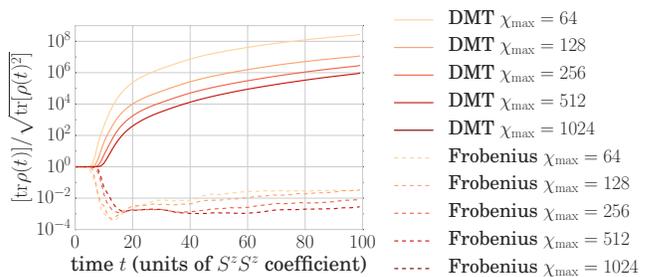

FIG. 5. Normalization $Z = [\text{tr}\,\rho]/\sqrt{\text{tr}[\rho^2]}$ as a function of time, comparing DMT (solid) and Frobenius (dashed), for the initial pure state (32) evolving under the Hamiltonian (30). Note that the second Rényi entropy of the whole chain (64-sites long) is $S_2 = 2\ln Z$. As $Z = 1$ for a pure state, any deviations from $Z = 1$ result from the truncation.
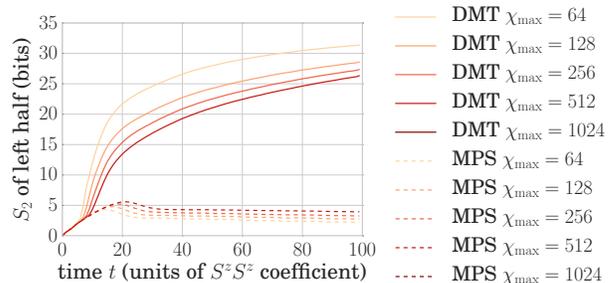

FIG. 6. Second Rényi entropy (in units of bits) of left half of a 64-site chain for an initial pure state (32) evolving under the Hamiltonian (30), in MPS and DMT simulations. The largest entropy we see ($\chi = 64$ at $t = 100$) is $S_2 \approx 31.4\,\text{bits}$, very close to the theoretical maximum of $32\,\text{bits}$.

nearly trivial.) In contrast to MPSs, MPDOs can represent states with arbitrarily large entropy by replacing system entanglement entropy with bath entropy.

Note that once truncation starts, the entropy in the DMT simulation increases above that in MPS. This is not unexpected: in ordinary MPS TEBD the entanglement entropy of the left half is exactly its entropy with the right half and is a property of the MPS at bond $\lfloor L/2 \rfloor$, so it can only increase when we apply a gate at bond $\lfloor L/2 \rfloor$. In the DMT algorithm, on the other hand, the entanglement entropy of the left half of the chain is entanglement entropy not only with the right half but also with a notional bath, and it increases with every truncation on bonds within the left half.

Figure 7 shows the system's total energy over time as simulated by ordinary TEBD and our density-matrix TEBD. In the DMT simulation, the energy is constant. MPS time evolution, however, imposes an additional 'heating' whereas DMT is designed to conserve total energy. Because the MPS representation biases towards low-entanglement states, the system drifts towards the extrema of the energy spectrum over time. (The simulation begins with positive energy, and hence drifts towards
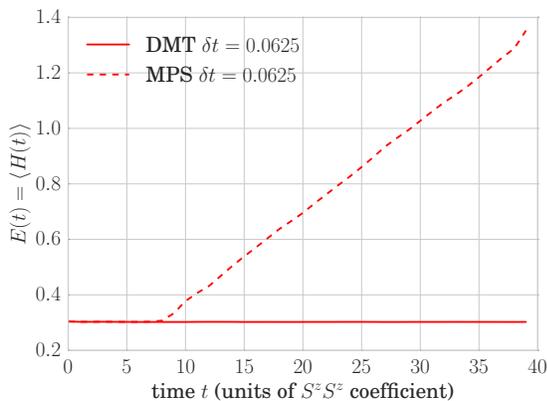
FIG. 7. Energy over time at fixed $\chi = 16$ for the initial pure state (32) evolving under the Hamiltonian (30) at system size $L = 64$. By design, the total energy remain constant under DMT.

the negative temperature state $T \to 0^-$.)

Onsite spins are easy to measure but hard to analyze: their expection values are noisy and they decay quickly with time. Instead, we measure a Fourier component of the energy density. The energy density ϵ_j, defined over a pair of sites, is

$$\epsilon_1 = \frac{h^z}{2}\big(S_1^z + \tfrac{1}{2}S_2^z\big) + \frac{h^x}{2}\big(S_1^x + \tfrac{1}{2}S_2^x\big) + S_1^z S_2^z,$$

$$\epsilon_{1<j<L-1} = \frac{h^z}{2}\big(\tfrac{1}{2}S_j^z + \tfrac{1}{2}S_{j+1}^z\big) + (x \leftrightarrow z) + S_j^z S_{j+1}^z,$$

$$\epsilon_{L-1} = \frac{h^z}{2}\big(\tfrac{1}{2}S_{L-1}^z + S_L^z\big) + (x \leftrightarrow z) + S_{L-1}^z S_L^z.$$
$$(34)$$

We measure a Fourier component of the energy density

$$\epsilon_{k=\pi/4} = -\frac{1}{L}\sum_{j=1}^{L-1} e^{ikj}\epsilon_j, \qquad k = \pi/4 \qquad (35)$$

with a wavelength of 8 sites. Fourier components are eigenmodes of the diffusion equation which should govern the system's long-time non-equilibrium behavior. We choose this particular component (and choose the initial state accordingly) because its wavelength is long enough that it should not immediately decay, but not so long as to be longer than accessible system sizes.

Figure 8(a) shows the the Fourier component of energy density (35) in as simulated by MPS (dashed) and DMT for $L = 16, 20, 24$. At fixed maximum bond dimension $\chi \ll 2^{\lfloor L/2 \rfloor}$, DMT is more accurate than MPS TEBD, which illustrates the power of DMT in both short- and long-time dynamics. Moreover, where MPS TEBD error increases with system size, DMT error decreases. This is due to finite-size deviations from thermalizing behavior: oscillations about the local equilibrium values for local operators result from long-range coherences that we do not expect to be able to capture.

Any pure state on a system of length L can be represented exactly by matrix product states with bond dimension $\chi_{\max} = 2^{\lfloor L/2 \rfloor}$. At this bond dimension the evolution by MPS TEBD becomes semiexact for all times; no truncation occurs for the simulation. We can therefore simulate pure-state evolution of an MPS using exactly the same Hamiltonian (30) and boustrophedon Trotter decomposition (A6), and compare the results to those of DMT, shown in Fig. 8(b). The data is a measure of the error introduced by the truncation; these are small $\approx 10^{-3}$ for a wide range of bond dimensions in the DMT simulations.

C. Application: mixed-state evolution (near equilibrium)

To probe the behavior of our algorithm near equilibrium, we take as our initial state a Gibbs state with a spatially-varying temperature

$$\rho_0 \propto \exp\left[\sum_j \beta_j \epsilon_j\right] \qquad (36)$$

with ϵ_j the energy density of Eq. (34), and

$$\beta_j = \beta_0\big(1 + g_j'\big) \qquad (37)$$

where

$$g_j' = 0.1 \times \begin{cases} 0 & j \bmod 8 = 1, 2, 7, \text{ or } 0, \\ 1 & j \bmod 8 = 3, 4, 5, \text{ or } 6. \end{cases} \qquad (38)$$

This temperature profile is broadly similar to the S^z profile we impose on the pure initial state (see Eq. (32) and Fig. 4).

See Appendix C 1 for details of the construction of the Gibbs state.

In Figs. 9 and 10 we compare DMT to the purification method of Karrasch, Bardarson and Moore [44], which we label "purification". This method takes advantage of the freedom to apply unitaries to the ancillae by time-evolving the ancillae backwards even as it evolves the physical system forwards. The time-evolution framework is therefore very similar to ours; the chief differences are in the interpretation of the vector space $(\mathbf{C}^{d^2})^L$ in which one works and in the truncation algorithm. The similarity is magnified by our choice to use the boustrophedon Trotter decomposition (A6) not only for DMT but for purification.

Both DMT and this purification time evolution converge very quickly, as one might expect: the results are essentially identical between the methods and between bond-dimension cutoffs, even to quite small bond dimensions. (Note that we subtract the thermal value in each case.)
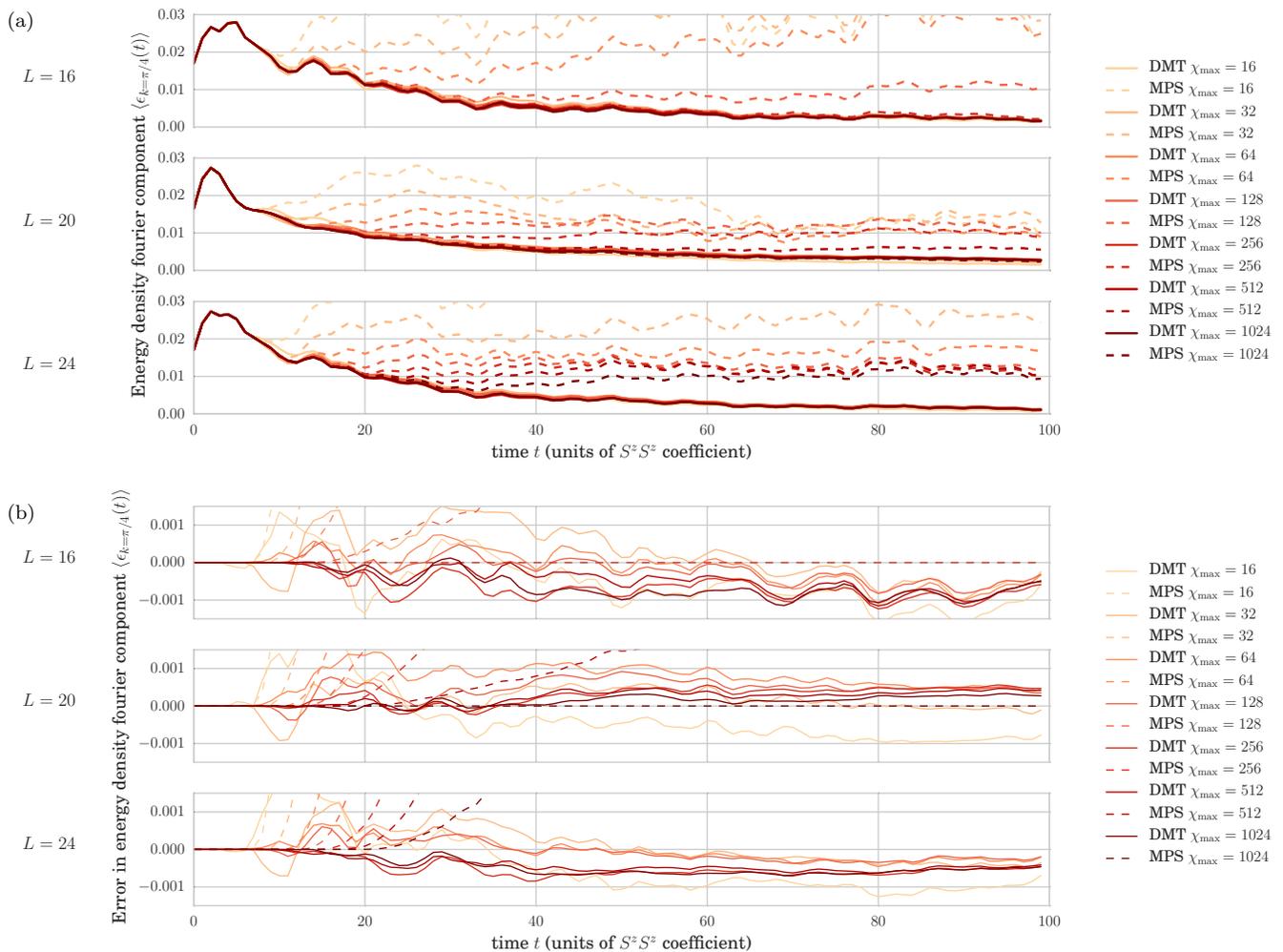
FIG. 8. (a) Fourier component of energy density for the initial pure state (32) evolving under the Hamiltonian (30) on chains of length $L = 16, 20, 24$. (b) The 'error' of the energy density, measured by comparing each data set with the semiexact result as simulated by MPS at $\chi = 2^{L/2}$. Note we do not show the MPS simulations for $\chi_{max} = 2048, 4096$.

V. CONCLUSION

We have presented an algorithm for approximating density operators by low-rank matrix product operators suitable for simulating long-time quantum dynamics. The method exactly preserves expectation values of operators on up to three contiguous sites, and it slots neatly into a standard Trotter-decomposition framework for time evolution of matrix product structures (TEBD), allowing time evolution by an ETH Hamiltonian of a variety of initial states.

Our algorithm, DMT, qualitatively outperforms its nearest competitor (ordinary MPS TEBD) for pure initial states. We use the fact that matrix product density operators with small bond dimension can represent states with high entropy to circumvent the area-law entanglement bound on matrix product states. Thus far the work is unoriginal: Zwolak and Vidal realized this was possible more than a decade ago. Our key insight is that we can

preserve the trace of the density matrix and the expectation values of conserved quantities by appropriately rotating the Schmidt spaces at the bond at which we truncate. Consequently, DMT can simulate time evolution by ETH Hamiltonians to arbitrary times using memory and computation time polynomial in system size.

In addition, DMT matches the current state of the art (purification time evolution) for near-equilibrium mixed initial states and outperforms it for far-from-equilibrium initial states.

We did not compare our algorithm to the interesting recent work of Leviatan et al.[19]. They use intuition not unlike ours to argue that the time-dependent variational principle of Haegeman et al. [12, 15], which approximates quantum dynamics as a classical Hamiltonian flow on the manifold of low-bond-dimension matrix product states and therefore exactly conserves energy, should give good results for ETH systems. The goals of Ref. 19—and, consequently, their benchmarking protocols—differ from
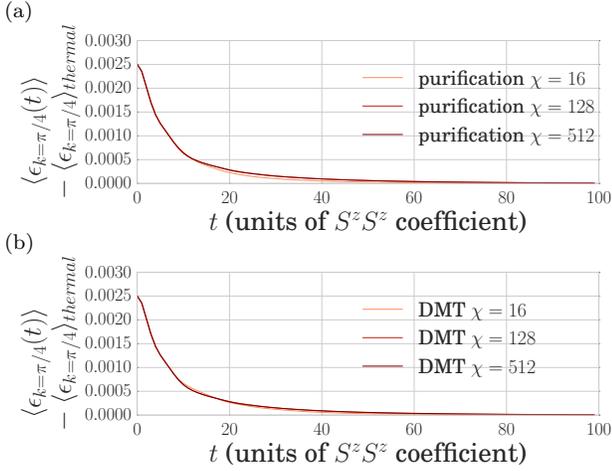
(a)

(b)

FIG. 9. Fourier component of energy density for (a) purification time evolution and (b) DMT for the near-equilibrium mixed state (36) evolving under the Hamiltonian (30) on a 128-site chain. The thermal value is $\langle \epsilon_{k=\pi/4} \rangle_{\text{thermal}} = -0.00038$
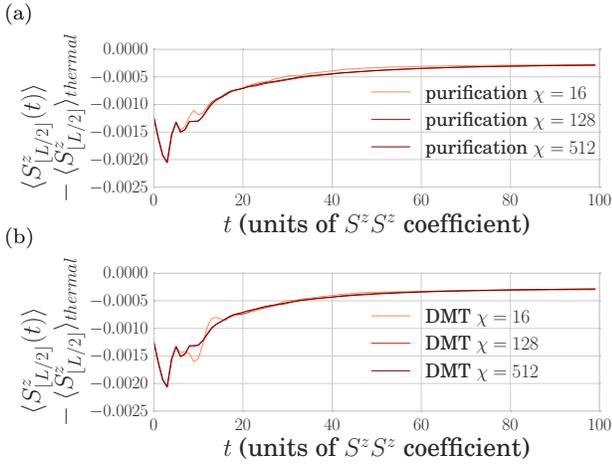
(a)

(b)

FIG. 10. Expectation value of S^z at the midpoint of the chain for (a) purification time evolution and (b) DMT for the near-equilibrium initial state (36) evolving under the Hamiltonian (30) on a 128-site chain. Both methods converge very quickly, so they give nearly identical results (cf. Fig. 14). This expectation value fails to approach the thermal value due to the large Trotter step we use ($dt = 1.0$). The thermal value is $\langle S^z_{\lfloor L/2 \rfloor} \rangle_{\text{thermal}} = -0.0622$.

ours, but a direct comparison could be an interesting topic for future work.

The reader would be right to worry that our method does not converge: as we increase the bond dimension above a certain value (perhaps 2^5–2^6), the accuracy of our method does not improve. We suspect that—once again—this is a result of the operators $O^y_{l,t}$, whose large expectation values result from the fact that we start near an S^y_l eigenstate. When we reduce the rank of the matrix $\tilde{M}$ in (24), we still do so in a way that minimizes error with respect to the Frobenius norm (even though

we have arranged to exactly preserve very-short-range operators). This means that the operators $O^y_{l,t}$ again dominate the error, and the matrix resulting from the truncation is pulled toward those operators. The obvious next step is to reduce rank in such a way that we minimize error with respect to a different norm, one that takes into account the spatial structure of the operator space: if we truncate at bond j, we should weight errors along $\sigma^z_{j-1}\sigma^z_{j+2}$ more heavily than errors along $\sigma^z_{j-7}\sigma^z_{j+6}$. Such controlled-metric truncation is a natural extension of this work.

One natural question to attack using our algorithm is the characterization of the ergodic side of the MBL transition. The random field Heisenberg model with small disorder appears to satisfy the ETH [57], but the nature of its dynamics is unclear (see the review of Luitz and Bar Lev [58]). Quantities like the spin-spin correlation $\langle S^z_{i+r}(t)S^z_i(t) \rangle$, from which one can compute a number of diagnostics for subdiffusion, should be straightforward to calculate using our method.

More interesting still are questions about interfaces between ETH and MBL systems. Besides being of inherent interest (how large must a bath be to thermalize an MBL system of a given size? How quickly does it thermalize?), answers to these questions will shed light on the phenomenological RG schemes of Potter, Vasseur, and Parameswaran [59] and Vosk, Huse, and Altman [60] for which ETH-MBL interfaces are fundamental building blocks. Because MBL systems display low entanglement in a wide variety of situations, we expect our algorithm to be able to simulate both bath and system out to large system sizes.

ACKNOWLEDGMENTS

Appendix A: Time-evolution framework

Consider a Hamiltonian

$$H = \sum H_j \qquad (A1)$$

(for instance (30)) where H_j is supported on sites $j, j+1$. Call the onsite Hilbert space $\mathcal{H}_j$ and its dual $\mathcal{H}_j^*$. Pure states then live in a Hilbert space $\mathcal{H} = \mathcal{H}_j^{\otimes L}$, and density operators ρ in a Hilbert space $\mathcal{H} \otimes \mathcal{H}^* = [\mathcal{H}_j \otimes \mathcal{H}_j^*]^{\otimes L}$. Closed-system Hamiltonian evolution, then, is

$$\frac{d}{dt}\rho = -i[H, \rho] \equiv -iH^\sharp \rho \qquad (A2)$$

with a linear superoperator Hamiltonian defined

$$H^\sharp : \mathcal{H} \otimes \mathcal{H}^* \to \mathcal{H} \otimes \mathcal{H}^*,$$
$$H^\sharp = H \otimes 1 - 1 \otimes H. \qquad (A3)$$

For a spin-1/2 chain, we can write $S_j^{x,y,z}$ for operators on the ordinary ("ket") space $\mathcal{H}_j$ and $T_j^{x,y,z}$ for operators on the dual ("bra") space $\mathcal{H}_j^*$; in this notation, the superoperator corresonding to our fruit-fly Ising Hamiltonian (30) is

$$\begin{aligned} H^\sharp = \sum H_j^\sharp \\ \equiv \sum \Big[(S_j^z S_{j+1}^z - T_j^z T_{j+1}^z) \\ + \frac{1}{2}h^x(S_j^x - T_j^x) + \frac{1}{2}h^z(S_j^z - T_j^z) \Big]. \quad (A4) \end{aligned}$$

In order to time-evolve a density matrix by a time t, one applies the superoperator unitary $e^{-iH^\sharp t}$. We discretize this operator by a timestep δt, as usual:

$$e^{-iH^\sharp t} = \left[e^{-iH^\sharp \delta t} \right]^{t/\delta t}. \qquad (A5)$$

We then perform a second-order Trotter decomposition into $2L-1$ two-site unitaries ("gates")

$$e^{-iH^\sharp \delta t} \simeq \underbrace{\prod_{j=1}^{L-1} \left[e^{-iH_j^\sharp \delta t/2} \right]}_{\text{leftward sweep after}} \underbrace{\prod_{j=L-1}^{1} \left[e^{-iH_j^\sharp \delta t/2} \right]}_{\text{rightward sweep}}; \quad (A6)$$

the error is of order $L\|h_j\|^3 \delta t^3$, where $\|h_j\|$ is an estimate of the typical magnitude of the terms h_j. (Note that this is a boustrophedon, DMRG-like "sweep", not the usual even-odd Trotter decomposition of Vidal's TEBD. We choose this Trotter decomposition for reasons of numerical stability.)

We apply this series of gates to a so-called matrix product density operators (MPDO): a representation of a density matrix ρ of the form

$$\rho = \sum \left[A_1^{\mu_1} \cdots A_j^{\mu_{j-1}} \right] s \left[B_j^{\mu_j} \cdots B_L^{\mu_L} \right] \hat\sigma_1^{\mu_1} \cdots \hat\sigma_L^{\mu_L}. \quad (A7)$$

The common dimension of the matrices $A_n^{\mu_n}, A_{n+1}^{\mu_n}$ (for any n) is called the *bond dimension*; the bond dimension controls the time and memory requirements for storing and operating on the MPDO. (This MPDO is in multicanonical form with orthogonality center at site j. All MPDOs with which we work will be in multicanonical form.) It is convenient to think of applying a gate (say, $e^{-ih_{j,j+1}^\sharp \delta t/2}$) as moving the orthogonality center from a bond adjacent to the bond on which the gate is applied (in this case bond $j+1$) to the bond on which it was applied; the boustrophedon Trotter decomposition (A6) thus drags the orthogonality center from right end to left and back, over and over again.

When applied to an MPDO, each two-site gate increases the bond dimension at bond by up to a factor of $(\dim \mathcal{H})^2$. Consequently, we must *truncate*: approximate the MPDO by another with a smaller bond dimension (and hence less demanding memory and time requirements); section III describes our truncation algorithm.

Appendix B: Relation between truncation of correlation matrix and bond dimension of MPDO

In Sec. III C, we truncated

$$\tilde{M} \mapsto \tilde{M}' \qquad (B1)$$

such that the block $\tilde{M}'_{\alpha\beta}, \alpha, \beta \geq 4$ has rank χ', then claimed that

$$\text{rank}\left[M'_{\alpha\beta} = \tilde{M}'_{\alpha\beta} + \frac{M_{\alpha 0} M_{0\beta}}{M_{00}} \right] \leq \chi' + 8. \qquad (B2)$$

To see that this is true, first decompose $\tilde{M}'$ as

$$\tilde{M}' = \tilde{M}'^A + \tilde{M}'^B + \tilde{M}'^C \qquad (B3)$$

with $\tilde{M}'^{A,B,C}$ left, upper, and lower right blocks of M' respectively:

$$\begin{aligned} \tilde{M}'^A_{\alpha\beta} &= \tilde{M}'_{\alpha\beta}, 0 \leq \alpha \leq 3 \\ \tilde{M}'^B_{\alpha\beta} &= \tilde{M}'_{\alpha\beta}, 3 < \alpha, 0 \leq \beta \leq 3 \qquad (B4) \\ \tilde{M}'^C_{\alpha\beta} &= \tilde{M}'_{\alpha\beta}, 3 < \alpha, \beta \end{aligned}$$

(other elements zero). These have ranks

$$\begin{aligned} \text{rank } \tilde{M}'^A &\leq 4, \\ \text{rank } \tilde{M}'^B &\leq 4, \qquad (B5) \\ \text{rank } \tilde{M}'^C &\leq \chi', \end{aligned}$$

so

$$\begin{aligned} \text{rank}(\tilde{M}') &\leq \text{rank } \tilde{M}'^A + \text{rank } \tilde{M}'^B + \text{rank } \tilde{M}'^C \\ &\leq 8 + \chi'. \end{aligned} \qquad (B6)$$

Since

$$\text{range } \frac{M_{\alpha 0} M_{0\beta}}{M_{00}} \subseteq \text{range } \tilde{M}'^A \qquad (B7)$$

(the range of an operator is also known as its column space) we have

$$\mathrm{rank}\, M' = \mathrm{rank}\, \tilde{M}' \leq \chi' + 8, \qquad (B8)$$

as desired.

Appendix C: Matrix product density operator representations of Gibbs states

1. Construction

In sections IV C and D we require a Gibbs MPDO as our initial state. The Gibbs state is

$$\rho \propto e^{-\beta H} = e^{-\beta H/2} I e^{-\beta H/2}; \qquad (C1)$$

which is precisely the imaginary-time evolution of the product MPDO I by the Hamiltonian superoperator

$$H^{\sharp}_{\mathrm{therm}} = H \otimes 1 + 1 \otimes H \qquad (C2)$$

out to time $\beta/2$. We approximate this imaginary-time evolution not by tDMRG with the boustrophedon Trotter decomposition (A6), but by ordinary TEBD using the trick of Hastings for numerical stability (as described in 7.3.2 of Schollwöck's magisterial review[52], q.v.).

2. Estimating thermal expectation values

In analyzing the time evolution of ETH states, one naturally requires Gibbs state expectation values as a function of total energy (or, equivalently, energy density): the long time limit of an expectation value is given by its expectation value in a Gibbs state whose energy density matches that of the initial state. We tabulate energy densities and observables of interests for Gibbs states at a variety of temperatures using MPDOs as described above; To find the long-time limit of an expectation value, we measure the energy density of the initial state and linearly interpolate between the two nearest Gibbs energy densities. Note that this does not account for Trotter heating (that is, the fact that—because the system actually simulated is a Floquet system with period given by the Trotter step δt, its energy as measured by the Hamiltonian simpliciter gradually increases).

Appendix D: Application: mixed-state evolution (far from equilibrium)

One might worry that the two initial states (32) from Sec. IV B and (36) from Sec. IV C are each special cases in their own ways: the first is a pure state, and the second is very near equilibrium.
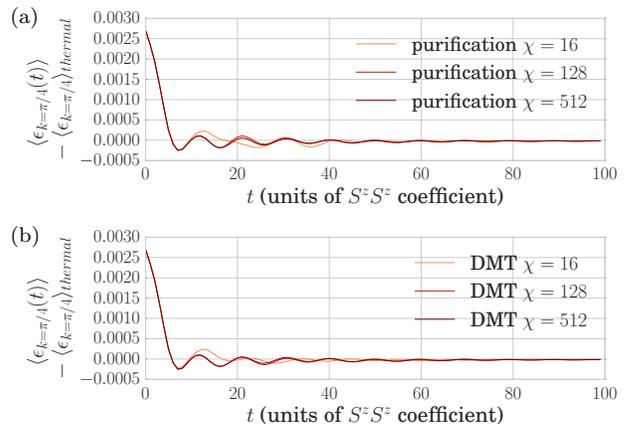

FIG. 11. Fourier component of energy density for purification time evolution and DMT starting from a far-from-equilibrium initial state on a 128-site chain. The thermal value is $\langle \epsilon_{k=\pi/4} \rangle_{\mathrm{thermal}} = -0.00031$.

In order to probe the performance of DMT for in more generic situations, we quench from a Gibbs state of the TFIM (30) with

$$\begin{aligned} h_0^x &= 0.5 \\ h_0^z &= 0.5 \end{aligned} \qquad (D1)$$

and

$$\beta_j = \beta_0 \left(1 + g_j'\right) \qquad (D2)$$

where

$$g_j' = 0.1 \times \begin{cases} 0 & j \bmod 8 = 1,\, 2,\, 7,\, \text{or } 0, \\ 1 & j \bmod 8 = 3,\, 4,\, 5,\, \text{or } 6. \end{cases} \qquad (D3)$$

(as in section IV C) to a TFIM (30) with

$$\begin{aligned} h_1^x &= 2.0 \\ h_1^z &= 0.5. \end{aligned} \qquad (D4)$$

We again compare to the purification method [44], and find that our method and that purification time evolution both converge quickly (see Figs. 11 and 12). Even very small bond dimensions (e.g. $\chi = 16$) can accurately treat long-time, hydrodynamic behavior; accurately treating short-time behavior requires somewhat higher bond dimension.

Appendix E: Convergence of mixed-state evolution

It is difficult to judge convergence of any of the three algorithms from plots like Figure 9 or 11. In Figure 13 we take a near-equilibrium initial state and plot the deviation in $\epsilon_{k=\pi/4}$, as measured for a series of bond dimensions χ, from the last (largest) χ in the series. In Figure 15 we do the same for a far-from-equilibrium mixed state,
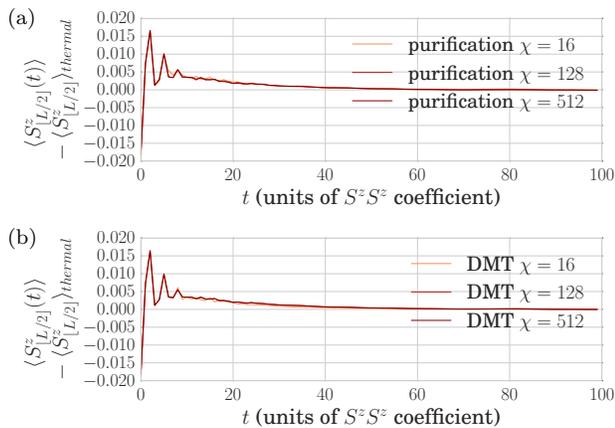
FIG. 12. Expectation value of S^z at the midpoint of the chain for purification time evolution and DMT starting from a far-from-equilibrium initial state on a 128-site chain. The thermal value is $\langle S^z_{\lfloor L/2 \rfloor} \rangle_{\text{thermal}} = -0.021$.

and in Figures 14, 16 for $S^z_{L/2}$ for near-equlibrium and far-from-equilibrium mixed states.

Our method converges with approximately the same bond dimension vs. accuracy tradeoff as purification time evolution for both the near-equilibrium initial state (Fig. 13) and the far-from-equilibrium initial state (Fig. 15). In both cases, Frobenius time evolution converges more slowly than either method.

[1] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).
[2] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).
[3] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).
[4] J. J. Garca-Ripoll, New Journal of Physics **8**, 305 (2006).
[5] M. C. Bañuls, M. B. Hastings, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **102**, 240603 (2009).
[6] A. Holzner, A. Weichselbaum, I. P. McCulloch, U. Schollwck, and J. von Delft, Physical Review B **83**, 195115 (2011).
[7] A. Mller-Hermes, J. I. Cirac, and M. C. Bauls, New Journal of Physics **14**, 075003 (2012).
[8] M. L. Wall and L. D. Carr, New Journal of Physics **14**, 125015 (2012).
[9] T. Enss and J. Sirker, New Journal of Physics **14**, 023008 (2012).
[10] D. Tamascelli, R. Rosenbach, and M. B. Plenio, Physical Review E **91**, 063306 (2015).
[11] F. A. Wolf, J. A. Justiniano, I. P. McCulloch, and U. Schollwöck, Phys. Rev. B **91**, 115144 (2015).
[12] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Physical Review Letters **107**, 070601 (2011).
[13] J. Haegeman, T. J. Osborne, and F. Verstraete, Physical Review B **88**, 075133 (2013).
[14] J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, Journal of Mathematical Physics **55**, 021902 (2014).
[15] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Physical Review B **94**, 165116 (2016).
[16] J. H. Bardarson, F. Pollmann, and J. E. Moore, Phys. Rev. Lett. **109**, 017202 (2012).
[17] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. Lett. **110**, 260601 (2013).
[18] T. Grover, ArXiv e-prints (2014), arXiv:1405.1471 [cond-mat.dis-nn].
[19] E. Leviatan, F. Pollmann, J. H. Bardarson, and E. Altman, ArXiv e-prints (2017), arXiv:1702.08894 [cond-mat.stat-mech].
[20] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[21] M. B. Hastings, Phys. Rev. B **73**, 085115 (2006).
[22] A. Molnar, N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. B **91**, 045138 (2015).
[23] D. Pekker and B. K. Clark, ArXiv e-prints (2014), arXiv:1410.2224 [cond-mat.str-el].
[24] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[25] T. Prosen and M. Žnidarič, Journal of Statistical Mechanics: Theory and Experiment **2009**, P02035 (2009).
[26] M. Žnidarič, T. c. v. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).
[27] T. Prosen and I. Pižorn, Physical Review A **76**, 032316 (2007).
[28] T. Prosen and I. Pižorn, Physical Review Letters **101**, 105701 (2008).
[29] M. Žnidarič, T. Prosen, and I. Pižorn, Physical Review A **78**, 022103 (2008).
[30] I. Pižorn and T. Prosen, Physical Review B **79**, 184416 (2009).
[31] G. Benenti, G. Casati, T. Prosen, D. Rossini, and M. Žnidarič, Physical Review B **80**, 035110 (2009).
[32] S. Jesenko and M. Žnidarič, Physical Review B **84**, 174438 (2011).
[33] M. Žnidarič, Physical Review E **83**, 011108 (2011).
[34] T. Prosen and M. Žnidarič, Physical Review B **86**, 125118 (2012).
[35] I. Pižorn, V. Eisler, S. Andergassen, and M. Troyer, New Journal of Physics **16**, 073007 (2014).
[36] J. Cui, J. I. Cirac, and M. C. Bauls, Physical Review Letters **114**, 220601 (2015).
[37] E. Mascarenhas, H. Flayac, and V. Savona, Phys. Rev. A **92**, 022116 (2015).
[38] U. Marzolino and T. Prosen, Physical Review A **93**, 032306 (2016).
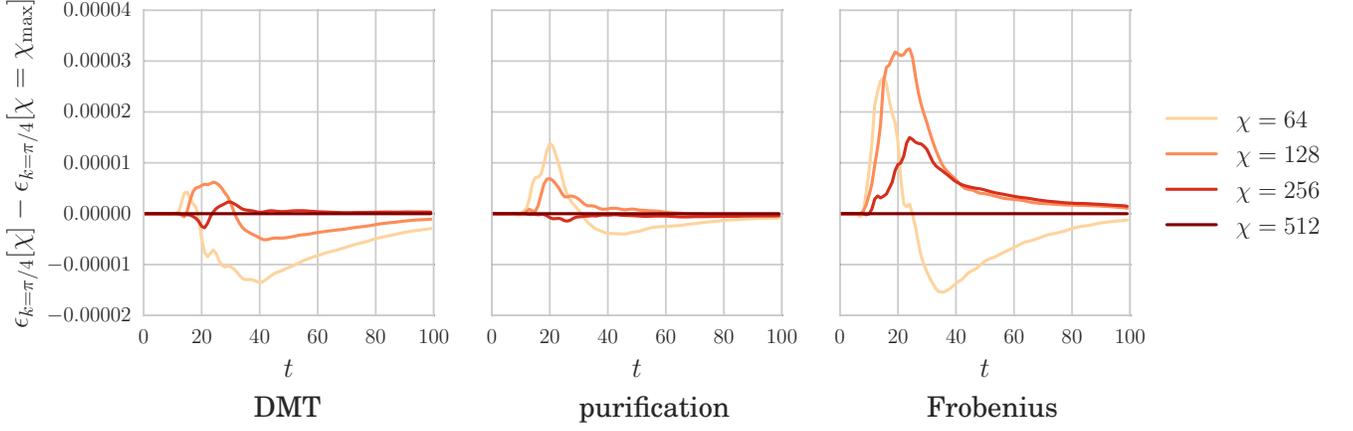[39] M. Kliesch, D. Gross, and J. Eisert, Phys. Rev. Lett.

FIG. 13. Convergence of $\epsilon_{k=\pi/4}$ for three algorithms. Initial state is a near-equilibrium mixed state (cf Section IV C and Figure 9) on a 128-site chain. For each algorithm, we plot $\epsilon_{k=\pi/4}[\chi] - \epsilon_{k=\pi/4}[\chi = \chi_{\max}]$—that is, how far the measurement during a run with a certain bond dimension χ deviates from measurement during a run with some high bond dimension.
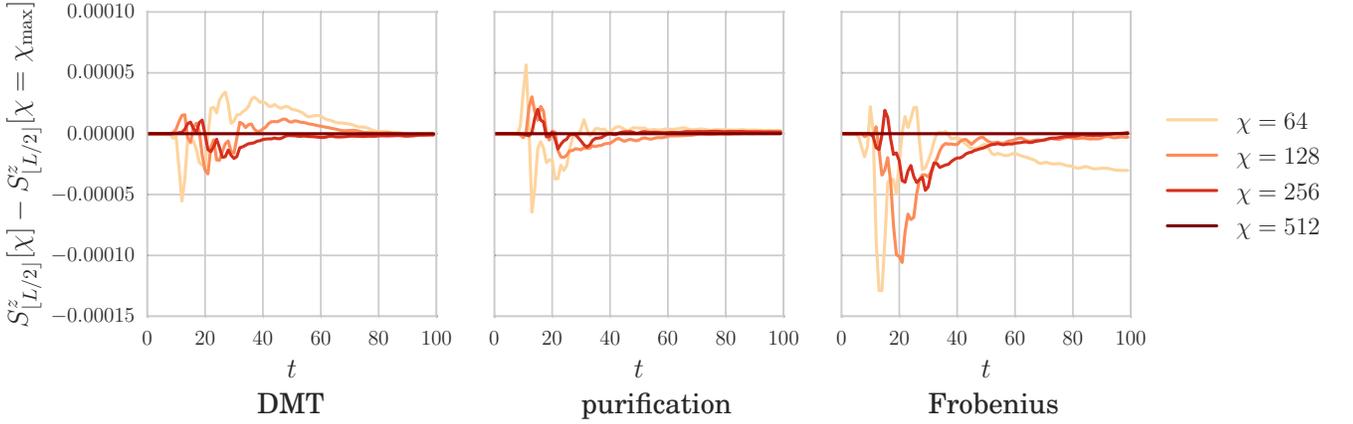

FIG. 14. Convergence of S^z at site $L/2$ for three algorithms. Initial state is a near-equilibrium mixed state (cf Section IV C and Figure 10) on a 128-site chain. For each algorithm, we plot $S^z_{L/2}[\chi] - S^z_{L/2}[\chi = \chi_{\max}]$—that is, how far the measurement during a run with a certain bond dimension χ deviates from measurement during a run with some high bond dimension.
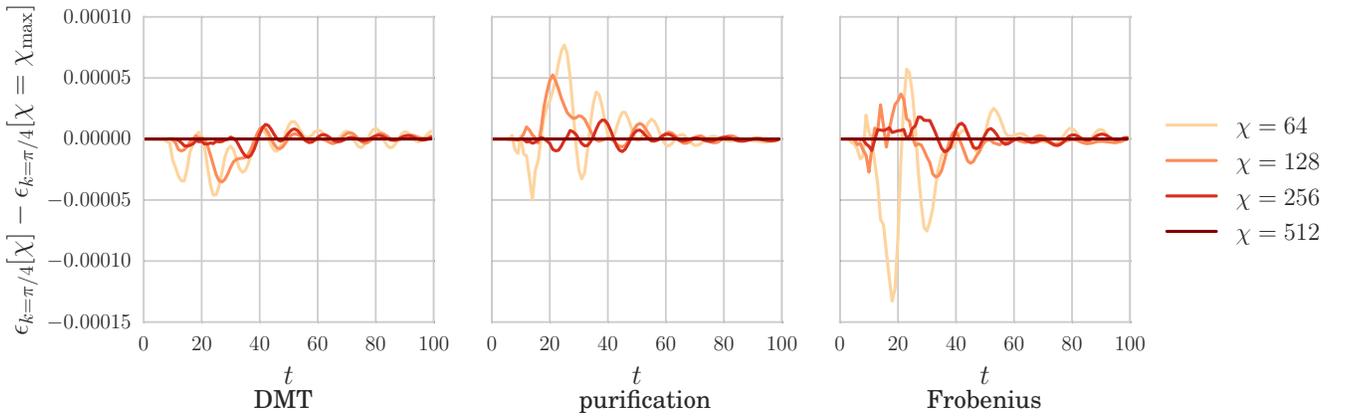

FIG. 15. Convergence of $\epsilon_{k=\pi/4}$ for three algorithms. Initial state is a far-from-equilibrium mixed state (cf Section D and Figure 11) on a 128-site chain. For each algorithm, we plot $\epsilon_{k=\pi/4}[\chi] - \epsilon_{k=\pi/4}[\chi = \chi_{\max}]$—that is, how far the measurement during a run with a certain bond dimension χ deviates from measurement during a run with some high bond dimension.
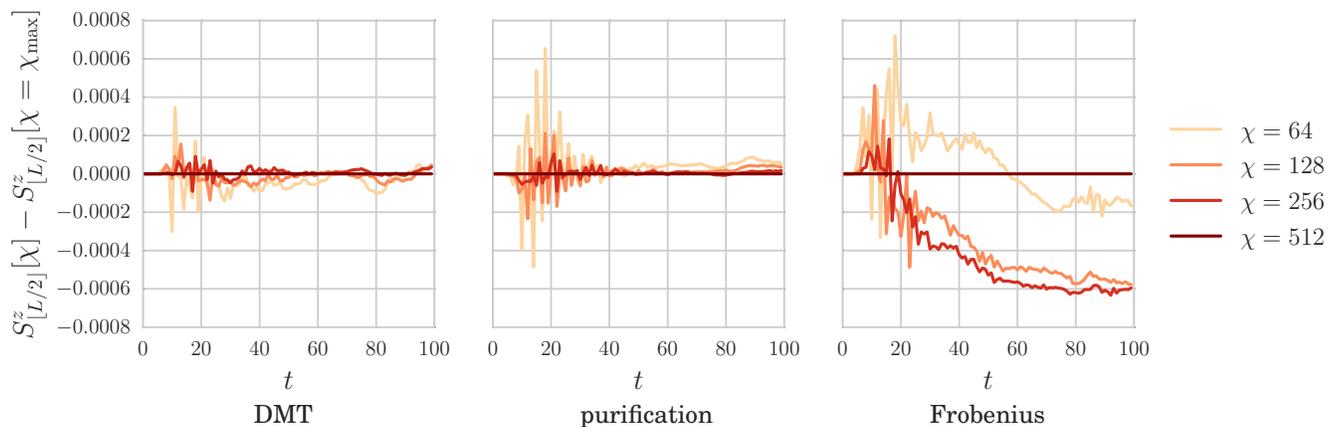
FIG. 16. Convergence of S^z at site $L/2$ for three algorithms. Initial state is a far-from-equilibrium mixed state (cf Section D and Figure 12) on a 128-site chain. For each algorithm, we plot $S^z_{L/2}[\chi] - S^z_{L/2}[\chi = \chi_{\max}]$—that is, how far the measurement during a run with a certain bond dimension χ deviates from measurement during a run with some high bond dimension.

113, 160503 (2014).

[40] T. Barthel, U. Schollwöck, and S. R. White, Phys. Rev. B **79**, 245101 (2009).

[41] A. E. Feiguin and G. A. Fiete, Physical Review B **81**, 075108 (2010).

[42] C. Karrasch, J. H. Bardarson, and J. E. Moore, Phys. Rev. Lett. **108**, 227206 (2012).

[43] T. Barthel, New Journal of Physics **15**, 073010 (2013).

[44] C. Karrasch, J. H. Bardarson, and J. E. Moore, New Journal of Physics **15**, 083031 (2013).

[45] A. C. Tiegel, S. R. Manmana, T. Pruschke, and A. Honecker, Phys. Rev. B **90**, 060406 (2014).

[46] M. Binder and T. Barthel, Phys. Rev. B **92**, 125119 (2015).

[47] C. Karrasch, D. M. Kennes, and F. Heidrich-Meisner, Physical Review B **91**, 115130 (2015).

[48] D. Kennes and C. Karrasch, Computer Physics Communications **200**, 37 (2016).

[49] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Phys. Rev. Lett. **116**, 237201 (2016).

[50] Y.-K. Liu, eprint arXiv:quant-ph/0604166 (2006),

quant-ph/0604166.

[51] We hat all operators except for the density matrix ρ, and identity I.

[52] U. Schollwöck, Annals of Physics **326**, 96 (2011), January 2011 Special Issue.

[53] H. Kim, T. N. Ikeda, and D. A. Huse, Phys. Rev. E **90**, 052105 (2014).

[54] M. C. Bañuls, J. I. Cirac, and M. B. Hastings, Phys. Rev. Lett. **106**, 050405 (2011).

[55] H. Kim, M. C. Bañuls, J. I. Cirac, M. B. Hastings, and D. A. Huse, Phys. Rev. E **92**, 012128 (2015).

[56] C.-J. Lin and O. I. Motrunich, ArXiv e-prints (2016), arXiv:1610.04287 [cond-mat.stat-mech].

[57] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).

[58] D. J. Luitz and Y. Bar Lev, ArXiv e-prints (2016), arXiv:1610.08993 [cond-mat.dis-nn].

[59] A. C. Potter, R. Vasseur, and S. A. Parameswaran, Phys. Rev. X **5**, 031033 (2015).

[60] R. Vosk, D. A. Huse, and E. Altman, Phys. Rev. X **5**, 031032 (2015).