

Finding purifications with minimal entanglement

Johannes Hauschild,^{1,*} Eyal Leviatan,² Jens H. Bardarson,³ Ehud Altman,⁴ Michael P. Zaletel,⁵ and Frank Pollmann¹¹*Department of Physics, T42, Technische Universität München, James-Frank-Straße 1, D-85748 Garching, Germany*²*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 7610001, Israel*³*Department of Physics, KTH Royal Institute of Technology, Stockholm SE-10691, Sweden*⁴*Department of Physics, University of California, Berkeley, California 94720, USA*⁵*Department of Physics, Princeton University, Princeton, New Jersey 08540, USA*

(Received 24 November 2017; revised manuscript received 24 October 2018; published 28 December 2018)

Purification is a tool that allows to represent mixed quantum states as pure states on enlarged Hilbert spaces. A purification of a given state is not unique and its entanglement strongly depends on the particular choice made. Moreover, in one-dimensional systems, the amount of entanglement is linked to how efficiently the purified state can be represented using matrix-product states (MPS). We introduce an MPS based method that allows to find the minimally entangled representation by iteratively minimizing the second Rényi entropy. First, we consider the thermofield double purification and show that its entanglement can be strongly reduced especially at low temperatures. Second, we show that a slowdown of the entanglement growth following a quench of an infinite temperature state is possible.

DOI: [10.1103/PhysRevB.98.235163](https://doi.org/10.1103/PhysRevB.98.235163)

I. INTRODUCTION

Simulating quantum many-body systems faces a fundamental difficulty due to the complexity required to represent highly entangled states. Significant progress has been made through the observation that quantum ground states of interest often have only limited (area-law) entanglement, and thus can be represented efficiently using matrix-product states (MPS) [1–3] in one dimension (1D) and generalized tensor-product states [4] in higher dimensions. Such approaches have been particularly successful in the study of ground-state properties of 1D systems, where the density matrix renormalization group (DMRG) method [5] revolutionized the efficiency of numerical methods.

To extend the success of DMRG to transport and nonequilibrium phenomena, it is necessary to simulate real-time evolution [6–8]. The bipartite entanglement of pure states generically grows linearly with time, which leads to a rapid exponential blowup in computational cost, limiting pure-state time evolution to rather short times. But, while the entanglement growth limits the ability to compute the real-time evolution of pure quantum states, it need not impose the same restriction on the imaginary-time evolution of mixed states [9,10]. It is then natural to ask if the time evolution of mixed states can be represented efficiently using MPS and what sets the difficulty of such computations.

There are different techniques for simulating mixed states using MPS methods, including a direct representation of the density matrix as a matrix product operator (MPO) [11], using minimally entangled typical thermal states (METTS) [12–15], and purification [16,17]; in this paper we focus on the latter. In purification, a density matrix ρ acting on a *physical* Hilbert

space \mathcal{H}^P is represented as a pure state $|\psi\rangle$ in an enlarged space $\mathcal{H}^P \otimes \mathcal{H}^A$:

$$\rho = \text{Tr}_A |\psi\rangle \langle \psi|. \quad (1)$$

It is always sufficient to choose \mathcal{H}^A to be identical to \mathcal{H}^P , “doubling” each degree of freedom (DOF) as illustrated in Fig. 1(a). We note that the purification description can be a limitation for infinite systems [18,19]. Yet on finite systems, a purification can be found formally by diagonalizing the density matrix. In equilibrium this gives the thermofield double (TFD) purification $|\psi_\beta\rangle = \frac{1}{\sqrt{Z}} \sum_n e^{-\beta E_n/2} |n\rangle_P |n\rangle_A$, where $|n\rangle$ are the eigenvectors and E_n the eigenvalues of the Hamiltonian. It was recently argued that the TFD state can be efficiently represented with an MPS of bond dimension that grows at most polynomially with the inverse temperature [9]. The TFD is only one possible choice of purification since Eq. (1) is left invariant under an arbitrary unitary transformation U_{anc} which acts only on the ancilla space \mathcal{H}^A .

This gauge freedom may be used to reduce the entanglement in $|\psi\rangle$, rendering the MPS representation more efficient [20,21]. Here, we propose a way to find the minimally entangled purification. This minimum defines the entanglement of purification E_p [22] [defined below in Eq. (2)], which thus plays a role similar to the entanglement entropy in the pure case: it bounds the bond dimension $\chi \geq e^{E_p}$ [23]. However, this lower bound is irrelevant unless there is an efficient algorithm to *find* the minimally entangled purification at a cost comparable to DMRG [e.g., $\mathcal{O}(\chi^3)$], which, since it constitutes a global optimization problem over the many-body Hilbert space, is not *a priori* obvious.

Following, we introduce a method to find an approximately optimal purification by sequentially applying local disentangling operations to the ancilla DOF. The cost of the disentangling procedure is comparable to DMRG, and the

*johannes.hauschild@tum.de

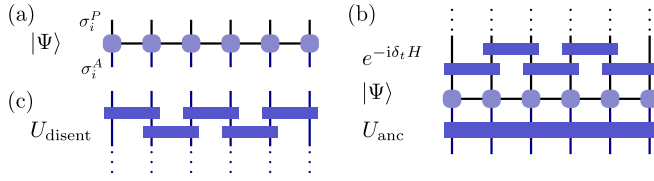


FIG. 1. (a) Schematic representation of purified states using MPS. (b) A purified state is evolved in real or imaginary time by acting on the physical degrees of freedom (e.g., using a Trotter decomposition of the time-evolution operator). The auxiliary degrees of freedom are only defined up to a global unitary U_{anc} which can be chosen to minimize the entanglement on the bonds. (c) The global U_{anc} is decomposed into a network of two-site gates to produce a disentangler U_{disent} .

resulting entanglement \tilde{E}_p reproduces the known properties of E_p in certain limits. We use the method to optimize both the equilibrium purification and that of a time-dependent state. We find that the method can significantly slow the entropy growth during real-time evolution down, as we demonstrate for both the transverse field Ising model and a disordered Heisenberg chain. For the latter, we find a slow spreading of \tilde{E}_p already for intermediate disorder strengths. In equilibrium, \tilde{E}_p approaches half of the entropy in the TFD state at low temperature.

Before proceeding we comment on the difference between the method presented here and two other proposals to compute long-time dynamics efficiently using MPS. Some of us have shown recently that the dynamics of local quantities in thermalizing systems can be captured accurately using the time-dependent variational principle (TDVP) [24], allowing to extract transport coefficients and even characteristics of chaos. Another one of us proposed a new truncation method to approximate the time evolution of a density matrix, represented as an MPO, to long times [25]. Both of these methods rely on the assumption that the increase of the nonlocal information encoded by the ever-growing entanglement entropy is irrelevant to the evolution of observable properties in thermalizing systems. These methods attempt to simulate the correct macrostate rather than the nearly exact microstate. Thus, the “truncation error” as usually defined in DMRG studies can be large as it is measured with respect to the exact state. In contrast, the approach presented here attempts, by optimizing the purification, to minimize the truncation error in order to compute the exact microstate.

II. PURIFICATIONS WITHIN THE MPS FORMALISM

To represent a purification as an MPS, we take $\mathcal{H}^P \sim \mathcal{H}^A$ so that each “site” contains a doubled DOF. The purification then takes the standard MPS form with an enlarged local Hilbert space as shown in Fig. 1(a). At infinite temperature, the TFD purification is obtained by maximally entangling the physical and ancilla DOF on every site i , e.g., $|\psi_0\rangle = \prod_i (\frac{1}{\sqrt{d}} \sum_{\sigma_i} |\sigma_i\rangle_P |\sigma_i\rangle_A)$, where σ_i runs over the local Hilbert space, resulting in a $\chi = 1$ MPS. In the standard purification approach, the finite-temperature TFD is obtained from $|\psi_0\rangle$ using imaginary-time evolution $|\psi_\beta\rangle \propto e^{-\frac{\beta}{2}H} |\psi_0\rangle$, from which thermal expectation

values are evaluated as $\langle Y \rangle_\beta = \langle \psi_\beta | Y | \psi_\beta \rangle$. Here, H acts only on \mathcal{H}^P . Similarly, to compute dynamical properties, for instance, $C(t, \beta) = \langle B^\dagger Y(t) B \rangle_\beta$, we define $|B(t, \beta)\rangle = e^{-itH} B |\psi_\beta\rangle$, so that $C(t, \beta) = \langle B(t, \beta) | Y | B(t, \beta) \rangle$. By taking $B = e^{i\epsilon X}$, this form is sufficient to find quantities of interest such as the spectral function $-i\partial_\epsilon C(t, \beta) = A_{YX}(t, \beta) = \langle [Y(t), X(0)] \rangle_\beta$. The requisite time evolution (both imaginary and real) can be simulated using standard methods [6–8, 26, 27].

The computational complexity of such simulations is generically linked to the bipartite von Neumann entanglement entropy $S_{LL'} = -\text{Tr}(\rho_{LL'} \log(\rho_{LL'}))$, where $\rho_{LL'} = \text{Tr}_{RR'}(|\psi\rangle\langle\psi|)$ is the reduced density matrix defined by a bipartition $\mathcal{H}^P = L \otimes R$ and $\mathcal{H}^A = L' \otimes R'$ at any of the bonds of the MPS; the bond dimension χ is bounded by $\chi \geq e^{S_{LL'}}$. Since other purifications can be obtained by acting with U_{anc} on the ancilla space [see Fig. 1(b)], it is desirable to exploit this choice to reduce $S_{LL'}$. Karrasch *et al.* [20] noticed that a natural choice is the “backward time evolution” $U_{\text{anc}} = e^{itH}$ because if B is local, this choice leaves $|B(t, \beta)\rangle$ invariant outside the growing “light cone” of the perturbation. Barthel [21] improved this approach by evolving both X and Y in the spectral function $A_{YX}(t, \beta) = \langle [Y(t), X(0)] \rangle_\beta = \langle [Y(t/2), X(-t/2)] \rangle_\beta$ as Heisenberg operators, which allows reaching times twice as long with comparable numerical effort [21, 28]. However, these prescriptions need not be optimal; ideally, we would minimize $S_{LL'}$ over all possible purifications, which would result in the entanglement of purification E_p [22]:

$$E_p[\rho_{LR}] \equiv \min_{|\psi\rangle} S_{LL'}[|\psi\rangle] = \min_{U_{\text{anc}}} S_{LL'}[U_{\text{anc}} |\psi\rangle]. \quad (2)$$

Equivalently, given an ansatz purification $|\psi\rangle$, we search for U_{anc} such that $U_{\text{anc}} |\psi\rangle$ has minimal entanglement; from this perspective, U_{anc} is a “disentangling” operation.

III. DISENTANGLING ALGORITHM

We propose an algorithm to approximately identify the optimal U_{anc} via a sequence of local disentangling operations, producing a circuit $U_{\text{anc}} = U_{\text{disent}}$ of the form shown in Fig. 1(c): the time evolution is applied to the purified state using the time-evolving block-decimation algorithm (TEBD) [6]. The TEBD algorithm [6] is based on a Trotter decomposition of e^{-itH} into two-site unitaries $e^{-i\delta_t H_{i,i+1}}$ as illustrated in Fig. 1(b). These unitaries are applied to the physical indices of the effective two-site wave function

$$|\Theta\rangle = \sum_{\substack{\sigma_i^P, \sigma_{i+1}^A, l \\ \sigma_{i+1}^P, \sigma_i^A, r}} \Theta_{l,r}^{\sigma_i^P \sigma_{i+1}^A, \sigma_{i+1}^P \sigma_i^A} |l\rangle |\sigma_i^P \sigma_{i+1}^A\rangle |\sigma_{i+1}^P \sigma_i^A\rangle |r\rangle,$$

where $|l\rangle$ (and $|r\rangle$) labels a basis consisting of Schmidt states to the left of site i (and right of site $i+1$, respectively). During a real-time evolution, we disentangle the two-site wave function right after each Trotter step using a unitary acting on the auxiliary space. These two-site disentanglers can be found using an iterative scheme based on minimizing the second Rényi entropy as a cost function as explained below, similar to the optimizations of a multiscale entanglement renormalization ansatz (MERA) [29]. As the time

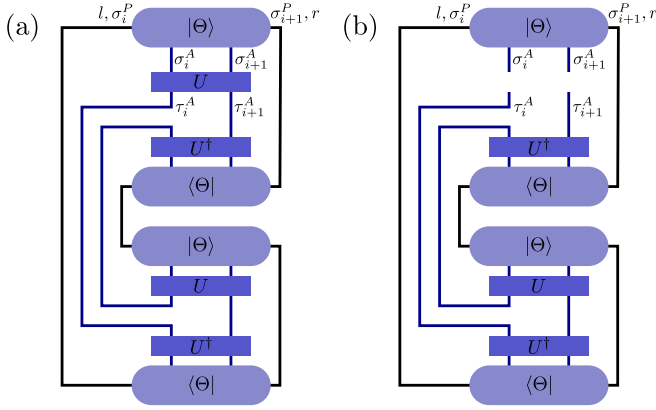


FIG. 2. (a) Tensor network for $Z_2(U, \Theta)$. (b) Effective environment $E_2(U, \Theta)$ such that $Z_2(U, \Theta) = \text{Tr}(U E_2(U, \Theta))$.

evolution proceeds, the disentangling unitary circuit U_{dissent} is then gradually built up by two-site unitaries, as depicted in Fig. 1(c). During an imaginary-time evolution we use a different scheme outlined in Sec. III C.

While the algorithm can suffer from numerical instabilities, we find empirically that it converges to a purification with significantly less entanglement compared to both backward time evolution and no disentangling at all, as shown in the benchmark section. The method described above is particularly suitable for correlation functions which involve only a single purification, e.g., $C(t, \beta)$, as there is no need to keep track of U_{dissent} . When two distinct purifications $|B(t)\rangle$ and $|A(t)\rangle$ are required, one would have to compress $U_{\text{dissent}}^\dagger U_{\text{dissent}}$ as a separate MPO.

A. Two-site disentangler minimizing the entropy

We explain now how to find a two-site unitary $U = U_{\sigma_i^A, \sigma_{i+1}^A}^{\tau_i^A, \tau_{i+1}^A}$ (i.e., acting in \mathcal{H}^A) which minimizes the entanglement of an effective two-site wave function $U|\Theta\rangle$, similar as during the optimization of MERA [29]. We chose to minimize the second Rényi entropy $S_2(U|\Theta) = -\log \text{Tr}(\rho_{LL'}^2)$, where $\rho_{LL'}$ is the reduced density matrix $\rho_{LL'} = \text{Tr}_{\sigma_{i+1}^P, \tau_{i+1}^A, r}(U|\Theta)\langle\Theta|U)$ [30]. In contrast to the von Neumann entropy, the second Rényi entropy is readily expressed as $S_2(U|\Theta) = -\log(Z_2)$ with the tensor network Z_2 depicted in Fig. 2(a); Z_2 is to be maximized. We solve this nonlinear optimization problem iteratively: in the n th iteration, we consider one U_{n+1} formally as independent of the other U_n and write $Z_2(U_{n+1}, U_n, \Theta) = \text{Tr}(U_{n+1} E_2(U_n, \Theta))$, where the network for the “environment” $E_2(U_n, \Theta)$ is shown in Fig. 2(b). It is easy to see that the unitary U_{n+1} maximizing this expression is given by a polar decomposition of $E_2(U_n, \Theta)$, in other words, we set $U_{n+1} := YX^\dagger$ where X and Y are obtained from a singular value decomposition of $E_2(U_n, \Theta) = X\Lambda Y^\dagger$. The unitary minimizing $Z_2(U, \Theta)$ is then a fixed point U_* of this iteration procedure. As a starting point of the iteration, one can choose the identity $U_1 := \mathbb{1}$. At later times, one can also use the result of U_n from previous iterations (for the same time step and at the same bond) as initial guess for the next disentangler, which reduces the number of necessary iterations in many cases.

Since this iteration is based on a descent, it tends to go into local minima within the optimization space. To find the global optimum, we can perform multiple iterations in parallel: one starting from the identity, and others starting from initially random unitaries (chosen according to the Haar measure, i.e., from the so-called circular unitary ensemble). From the unitaries obtained by the parallel iterations, we choose the one with the smallest final entropy.

The disentangler U_n obtained by the above procedure preserves the quantum numbers of symmetries in the Hamiltonian, at least if the initial guess U_0 preserves them. In the presence of such a symmetry, one should choose U_0 accordingly from the Haar measure on unitaries preserving the symmetry to avoid an artificial buildup of entanglement. In our case, we exploited the S^z conservation in the Heisenberg chain (4) to reduce the computational cost in the tensor contractions and singular value decompositions [31,32].

B. Two-site norm disentangler

In this section, we discuss an alternative way to obtain a two-site disentangler, which directly focuses on the required bond dimension. The procedure described below is equivalent to finding the “entanglement branching operator” introduced by Harada [33]. In order to reduce the bond dimension, we look for a two-site unitary U (acting on the ancilla DOF) for which the truncation of the effective two-site wave function $U|\Theta\rangle$ has the smallest truncation error. To find this U , we use a similar, iterative scheme as above: given U_n , we calculate the truncated $(U_n|\Theta)_{\text{trunc}}$ and find the U_{n+1} maximizing the overlap $|\langle\Theta|U_{n+1}^\dagger(U_n|\Theta)_{\text{trunc}}|$. Again, the new U_{n+1} can be found by a polar decomposition of the “environment” consisting of the corresponding tensor network for $|\langle\Theta|U_{n+1}^\dagger(U_n|\Theta)_{\text{trunc}}|$, but excluding the U_{n+1} . Since the optimal U depends on the final bond dimension χ after truncation, we need to gradually increase χ and repeat the iteration procedure until the truncation error for the given bond dimension is below a desired accuracy threshold. While we found that this gradual increase of χ also helps to find the optimal disentangler, it substantially increases the computational cost.

C. Global disentangling for imaginary-time evolution

In contrast to the real-time evolution, the Trotter gate $e^{-\delta\beta H_{i,i+1}}$ in imaginary-time evolution is nonunitary. Thus, it can change the Schmidt values and thus generate entanglement on sites it does not even act on, which creates the necessity for a more global scheme of disentangling than the one presented above for the real-time evolution. Instead, we perform the imaginary-time evolution as usual (with $U_{\text{anc}} = \mathbb{1}$) and disentangle only after each n th time step in a more global fashion: in this case, we find that generating the network of U_{dissent} by optimizing bonds with right and left sweeps similar as in DMRG is more effective than the Trotter-type scheme of even and odd bonds depicted in Fig. 1(c). Moreover, it is straightforward to generalize the two-site disentangling described above to multiple sites by grouping multiple sites. For example, we can disentangle the wave function of four sites $i, i+1, i+2, i+3$ by grouping each two sites as $(i, i+1)$ and $(i+2, i+3)$ and then using the above-described method.

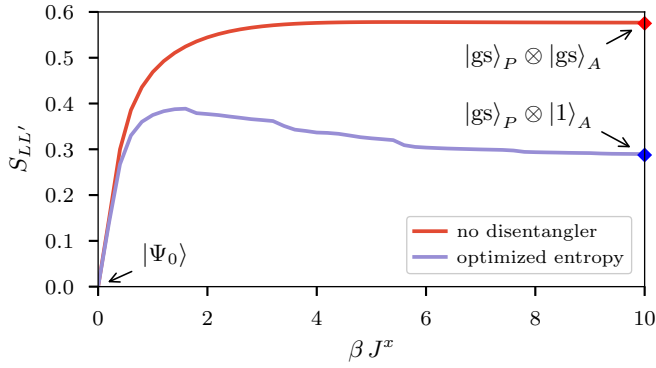


FIG. 3. Half-chain entanglement entropy of the finite-temperature purification $|\psi_\beta\rangle \propto e^{-\frac{\beta}{2}H} |\psi_0\rangle$ in the generalized Ising model (3) with $L = 50$ sites, for the TFD state ($U_{\text{anc}} = \mathbb{1}$, upper line) and when disentangling up to four sites at once (lower line). The parameters $J^x = h^z = 1$ and $J^z = 0.1$ are chosen to be in the vicinity of the quantum phase transition. The diamonds on the right axis indicate the half-chain entanglement S_{gs} (blue) and $2S_{\text{gs}}$ (red) of the ground state $|\text{gs}\rangle$ obtained from DMRG.

As the resulting disentangler can perform arbitrary “onsite” rotations within each group, it is necessary to disentangle the obtained wave function (recursively) within each group. While such a grouping provides additional freedom in the unitary to be found and is thus a systematic improvement for finding the optimal *global* disentangler, it comes at the cost of a scaling of required computational resources which is exponential in the number of included sites. In practice, we limited ourselves to optimizing at most four sites at once.

As an alternative for the global disentangling, we tried a method along the lines of Hyatt *et al.* [34]. Here, the idea is to identify pairs of sites with maximal mutual information as candidates for disentangling. Using swap gates (commonly used for TEBD with longer-range interactions [13]), we bring the two sites next to each other and disentangle them with a two-site disentangler as described above for the real-time evolution. Yet, we find that this approach is very limited by the fact that the purification can not be disentangled completely (except for $\beta \rightarrow \infty$), such that we fail at some point to identify the next candidate pair to be disentangled.

IV. BENCHMARKS

A. Finite temperatures

To benchmark our algorithm, we study a concrete example, the generalized transverse field Ising model

$$H = -J^x \sum_{i=1}^{L-1} \sigma_i^x \sigma_{i+1}^x - J^z \sum_{i=1}^{L-1} \sigma_i^z \sigma_{i+1}^z - h^z \sum_{i=1}^L \sigma_i^z. \quad (3)$$

For $J^z = 0$, the model maps onto free fermions and exhibits a quantum phase transition at $h_c^z = J^x$. The term proportional to J^z introduces interactions and breaks integrability.

Figure 3 compares the entanglement of the optimized purification with the entanglement of the TFD state obtained by imaginary-time evolution without disentangling, i.e., $U_{\text{anc}} = \mathbb{1}$. The infinite-temperature state $|\psi_0\rangle$ has maximal entanglement between the physical and auxiliary DOF on

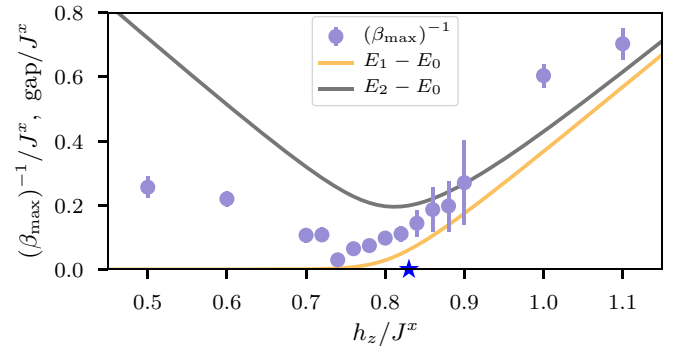


FIG. 4. Behavior of the position of the maximum β_{max} in Fig. 3 with the parameter h^z , for $L = 50$ sites and $J^z = 0.1 J^x$. Error bars indicate uncertainties in extracting β_{max} stemming from a limited resolution in β and numerical instabilities of the algorithm. For comparison, the energy gaps of the first and second states above the ground state (extracted with DMRG) are also shown. The critical h^z / J^x in the thermodynamic limit is indicated by the blue star on the x axis.

each site, but no correlations between different sites, hence, $S_{LL'} = 0$. For small β , the imaginary-time evolution starts to build up correlations between neighboring sites, but it is not immediately possible to disentangle the state with a rotation in \mathcal{H}^A . For example, a nontrivial unitary acting on σ_i^A and σ_{i+1}^A would lead to a strong correlation between σ_i^P and σ_{i+1}^A , and thus larger entanglement for a cut between sites i and $i+1$. However, due to the monogamy of entanglement, the buildup of quantum correlations between different sites ensures the reduction of the entanglement between the physical and auxiliary spaces. Consequently, the disentangler can reduce the entanglement at larger β . This is most evident in the limit of large β in which $e^{-\frac{\beta}{2}H}$ becomes a projector $|\text{gs}\rangle \langle \text{gs}|$ onto the ground state $|\text{gs}\rangle$. In this limit, the TFD purification ends up with two copies $|\text{gs}\rangle_P \otimes |\text{gs}\rangle_A$ of the ground state in \mathcal{H}^P and \mathcal{H}^A . In contrast, a perfect disentangling algorithm should be able to rotate $|\text{gs}\rangle_A$ into an unentangled product state $|1\rangle_A$, ending up with the state $|\text{gs}\rangle_P \otimes |1\rangle_A$ which has only half as much entanglement as the TFD. The fact that we find a purification with an entanglement close to that of the ground state shows that our algorithm can indeed find the minimum, i.e., it finds E_P .

Notably, we also find a maximum at intermediate β (although our algorithm suffers from numerical instabilities in this region). This can be understood from the fact that the entanglement of purification has contributions from both quantum fluctuations and thermal fluctuations, and the latter vanish for $\beta \rightarrow \infty$. A similar maximum is also present in the holographic prescription for the entanglement of purification [35,36]. Figure 4 shows that the maximum moves to larger β when tuning h^z towards the phase transition. We attribute this increase of β_{max} to the closing energy gap which induces thermal fluctuations at smaller temperatures (and thus additional entanglement entropy in the purified state on top of the ground-state entropy reached in the limit $\beta \rightarrow \infty$). In the symmetry-broken phase for $h^z \lesssim 0.75$, the ground state is (for the finite system almost) twofold degenerate, and $S_{LL'}(\beta \rightarrow \infty)$ is increased by $\log(2)$ on top of the

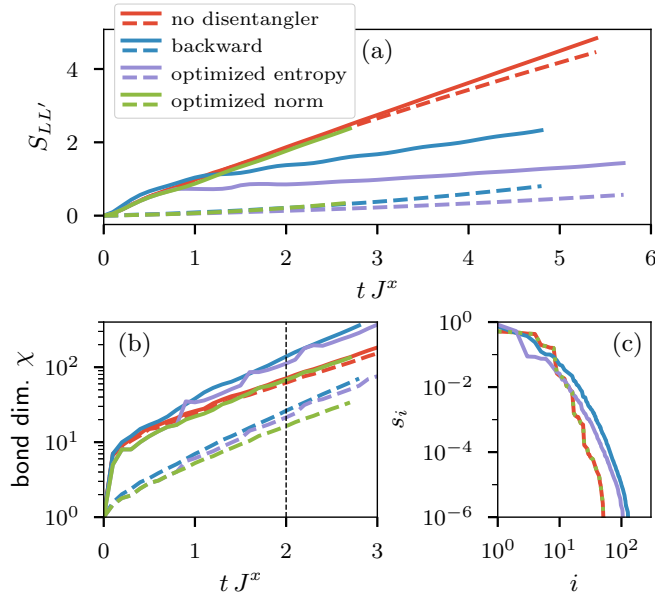


FIG. 5. (a) Comparison of the entanglement in the purification state $|S_{L/2}^+(t, \beta = 0)\rangle = e^{-itH} S_{L/2}^+ |\psi_0\rangle$ for the Ising chain (3) with $L = 40$ sites, $J^x = h^z = 1$, $J^z = 0.1$. (b) MPS bond dimension when the truncation error is kept below 10^{-6} in each step. (c) Decay of the Schmidt values s_i on the central bond at time $tJ^x = 2$. In (a) and (b), solid lines (dashed lines) show the maximum (mean) over different bonds. In all panels, different colors compare different disentanglers U_{anc} .

ground-state entanglement entropy. We still observe a maximum of $S_{LL'}$ at finite β in this phase, yet less pronounced than in the paramagnetic phase.

B. Real-time evolution at infinite temperature

Next, we consider the time evolution of a local operator applied to the infinite-temperature purification $|S_{L/2}^+(t, \beta = 0)\rangle = e^{-itH} S_{L/2}^+ |\psi_0\rangle$, where $S_i^+ = S_i^x + iS_i^y$. Figure 5(a) compares the resulting entanglement for no disentangling ($U_{\text{anc}} = \mathbb{1}$), backward time evolution ($U_{\text{anc}} = e^{itH}$), and the optimized disentangler ($U_{\text{anc}} = U_{\text{disent}}$) using the two-site disentanglers described in Secs. III A and III B. Note that for $\beta = 0$ backward time evolution is equivalent to the Heisenberg evolution of $S_{L/2}^+$. The maximum of the entropy over different bonds (solid lines) grows roughly linear in all three cases, yet with very different prefactors. While the growth is spatially almost uniform in the case $U_{\text{anc}} = \mathbb{1}$, both the backward time evolution and our optimized algorithm develop entropy only within a causal “light-cone,” which leads to a significant reduction when the mean over different bonds is taken (dashed lines). Figure 5(b) compares the growth of the required MPS bond dimension when the truncation error is kept fixed. Both backward time evolution and the optimized disentangler minimizing the entropy require a slightly higher maximal bond dimension close to where $S_{L/2}^+$ was applied. This apparent contradiction of a larger bond dimension despite a lower entropy can be understood from the fact that the entropy has large weight on the largest Schmidt values, but the required bond dimension

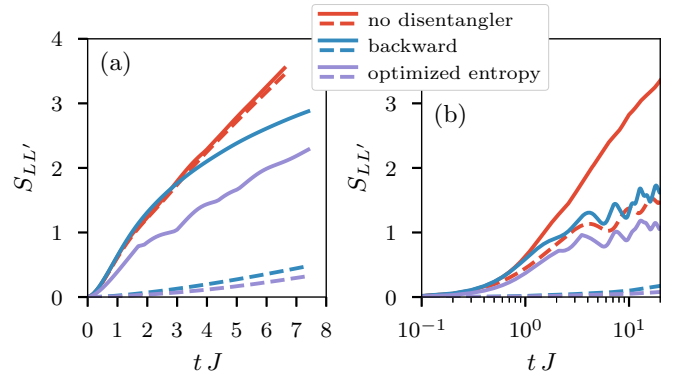


FIG. 6. Comparison of the entanglement in the purification state $|S_{L/2}^+(t, \beta = 0)\rangle = e^{-itH} S_{L/2}^+ |\psi_0\rangle$ for the Heisenberg chain (4) with $L = 80$ sites without disorder ($W = 0$) (a) and for a single disorder realization with $W = 5J$ (b). In both panels, different colors compare different disentanglers U_{anc} , and solid lines (dashed lines) show the maximum (mean) over different bonds.

is determined by the decay of Schmidt values in the tail. Indeed, we show in Fig. 5(c) that the optimization of the entropy leads to a reduction in the first few Schmidt values accompanied by a slightly longer tail of small Schmidt values compared to $U_{\text{anc}} = \mathbb{1}$. Nevertheless, the tail decays faster than with backward time evolution. In contrast, when the two-site disentangler described in Sec. III B is used, we can indeed slightly reduce the required maximal bond dimension as a proof of principle, even though in practice performing the optimization itself is computationally more expensive than the speedup gained by the reduced bond dimension. In this case, the disentangler acts almost trivially in the region where $S_{L/2}^+$ was applied, such that a larger tail of the singular values is avoided. While this optimization reduced the bond dimension during the real-time evolution, in the case of imaginary-time evolution we were not able to reduce the bond dimension with the same method.

As a second example, we consider the $S = \frac{1}{2}$ Heisenberg chain with disordered z -directed field

$$H = J \sum_{i=1}^{L-1} \vec{S}_i \cdot \vec{S}_{i+1} - \sum_{i=1}^L h_i^z S_i^z, \quad (4)$$

where h_i^z is chosen uniformly in the interval $[-W, W]$. This model has been established as a standard model in the study of many-body localization (MBL) [37–39] in one dimension. Numerically, a localization transition was found to occur at $W_c \approx 3.5J$ [40,41]. Figure 6(a) again compares the entanglement growth of $|S_{L/2}^+(t, \beta = 0)\rangle$ for the three choices of U_{anc} in the clean Heisenberg chain $W = 0$. While the entropy grows linearly when no disentangler is used, the integrability of the Heisenberg chain and the presence of S^z conservation restricts the entanglement of time-evolved local operators in the Heisenberg picture (here the “backward” evolution) to $S(t) \propto \log(t)$ [42]. Our results are compatible with the same $S(t) \propto \log(t)$ scaling when optimized, again with a smaller prefactor. In the MBL phase [Fig. 6(b)], even $U_{\text{anc}} = \mathbb{1}$ displays only a logarithmic entanglement growth, which is a characteristic feature of the MBL phase [43–46].

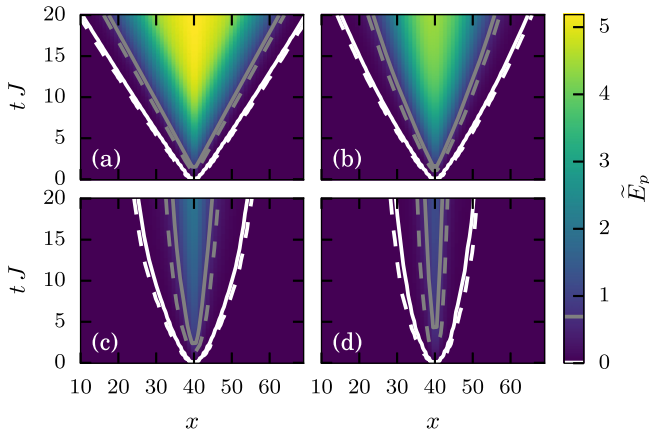


FIG. 7. Optimized entanglement entropy \tilde{E}_p in the state $e^{-itH} S_{L/2}^+ |\psi_0\rangle$ for the Heisenberg model (4) with disorder strength (a) $W = 0$, (b) $W = J$, (c) $W = 3J$, and (d) $W = 5J$, each averaged over 30 disorder realizations. The white and gray solid lines show the contour for the onset of finite values at a threshold of 0.01 and $\log(2)$. The dashed lines shows contours for the same threshold when backward time evolution is used.

Next, we focus on the spatial spread of the entanglement in $S_{L/2}^+ |\psi_0\rangle$ when using the optimized disentangler, tracking $S_{LL'}$ as a function of time t , and bipartition bond x , shown in Fig. 7. In the thermalizing regime, at small W [Figs. 7(a) and 7(b)], we observe the expected linear light-cone spreading [47]. Deep in the MBL phase [Fig. 7(d)], we find a qualitatively different spreading which is compatible with a logarithmic light cone. This is as expected from a generalized Lieb-Robinson bound $\mathbb{E} \| [A_i(t), B_j] \| \leq c t e^{\frac{|i-j|}{2\xi}}$, where ξ is the localization length and $c > 0$ some constant [48,49]. At intermediate disorder, near the MBL transition, we observe a sublinear spreading of the entanglement. Although there are extended eigenstates in this region, the system is expected to be subdiffusive and exhibits only slow transport on very long

time scales (inaccessible to our numerics) [50–55]. Since the backward time evolution already reduces $S_{LL'}$ to a zero (up to exponentially small corrections) outside of the light cone, it is not surprising that the contours of the onset are nearly unchanged compared to our optimized case.

V. CONCLUSIONS

We introduced an MPS based method to find a unitary U_{anc} acting on the ancilla DOF of a purification state, which reduces the entanglement both in equilibrium and during real-time evolution, at a similar cost to the TEBD algorithm. At low temperatures, the optimized entanglement entropy \tilde{E}_p is half as large as in the TFD state, providing evidence that the algorithm actually finds the entanglement of purification E_p . We find a maximum of \tilde{E}_p at intermediate β , the location of which diverges to $\beta_{\text{max}} \rightarrow \infty$ as the gap closes. During real-time evolution, the entanglement is significantly reduced both compared to $U_{\text{anc}} = \mathbb{1}$ and backward time evolution. In the clean Heisenberg chain, \tilde{E}_p shows a linear light-cone structure, which turns to a logarithmic spreading in the MBL phase (at large disorder). The minimization of the entanglement is, however, not directly accompanied by a reduction of the required bond dimensions, as it leads to a larger tail of small Schmidt values. This limitation might be overcome by another choice of local disentanglers.

ACKNOWLEDGMENTS

We acknowledge useful discussions with L. Schoonderwoerd. We have used the TeNPy package, which contains an implementation of our algorithm [32]. This work was partially supported by the ERC Starting Grant No. 679722. F.P. acknowledges the support of the DFG Research Unit FOR 1807 through Grants No. PO 1370/2-1 and No. TRR80, the Nanosystems Initiative Munich (NIM) by the German Excellence Initiative, and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No. 771537).

- [1] M. Fannes, B. Nachtergaele, and R. F. Werner, *Commun. Math. Phys.* **144**, 443 (1992).
- [2] M. B. Hastings, *Phys. Rev. B* **73**, 085115 (2006).
- [3] U. Schollwöck, *Ann. Phys. (NY)* **326**, 96 (2011).
- [4] F. Verstraete and J. I. Cirac, *arXiv:cond-mat/0407066*.
- [5] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [6] G. Vidal, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [7] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, *J. Stat. Mech.: Theory Exp.* (2004) P04005.
- [8] S. R. White and A. E. Feiguin, *Phys. Rev. Lett.* **93**, 076401 (2004).
- [9] T. Barthel, *arXiv:1708.09349*.
- [10] J. Dubail, *J. Phys. A Math. Theor.* **50**, 234001 (2017).
- [11] M. Zwolak and G. Vidal, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [12] S. R. White, *Phys. Rev. Lett.* **102**, 190601 (2009).
- [13] E. M. Stoudenmire and S. R. White, *New J. Phys.* **12**, 055026 (2010).
- [14] M. Binder and T. Barthel, *Phys. Rev. B* **95**, 195148 (2017).
- [15] M. Berta, F. G. S. L. Brandao, J. Haegeman, V. B. Scholz, and F. Verstraete, *arXiv:1709.07423*.
- [16] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [17] T. Barthel, U. Schollwöck, and S. R. White, *Phys. Rev. B* **79**, 245101 (2009).
- [18] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. Ignacio Cirac, *New J. Phys.* **15**, 123021 (2013).
- [19] G. De las Cuevas, T. S. Cubitt, J. I. Cirac, M. M. Wolf, and D. Pérez-García, *J. Math. Phys.* **57**, 071902 (2016).
- [20] C. Karrasch, J. H. Bardarson, and J. E. Moore, *Phys. Rev. Lett.* **108**, 227206 (2012).
- [21] T. Barthel, *New J. Phys.* **15**, 073010 (2013).
- [22] B. M. Terhal, M. Horodecki, D. W. Leung, and D. P. DiVincenzo, *J. Math. Phys.* **43**, 4286 (2002).
- [23] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, *Phys. Rev. Lett.* **100**, 030504 (2008).
- [24] E. Leviatan, F. Pollmann, J. H. Bardarson, and E. Altman, *arXiv:1702.08894*.

- [25] C. D. White, M. Zaletel, R. S. K. Mong, and G. Refael, *Phys. Rev. B* **97**, 035127 (2018).
- [26] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [27] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, *Phys. Rev. B* **91**, 165112 (2015).
- [28] C. Karrasch, J. H. Bardarson, and J. E. Moore, *New J. Phys.* **15**, 083031 (2013).
- [29] G. Evenbly and G. Vidal, *J. Stat. Phys.* **157**, 931 (2014).
- [30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
- [31] S. Singh, R. N. C. Pfeifer, and G. Vidal, *Phys. Rev. A* **82**, 050301 (2010).
- [32] J. Hauschild and F. Pollmann, *SciPost Phys. Lect. Notes* **5** (2018), the code is available online at <https://github.com/tenpy/tenpy/>.
- [33] K. Harada, *Phys. Rev. B* **97**, 045124 (2018).
- [34] K. Hyatt, J. R. Garrison, and B. Bauer, *Phys. Rev. Lett.* **119**, 140502 (2017).
- [35] K. Umemoto and T. Takayanagi, *Nat. Phys.* **14**, 573 (2018).
- [36] P. Nguyen, T. Devakul, M. G. Halbasch, M. P. Zaletel, and B. Swingle, *J. High Energ. Phys.* **01** (2018) 098.
- [37] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Ann. Phys. (N. Y.)* **321**, 1126 (2006).
- [38] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [39] D. A. Abanin and Z. Papić, *Ann. Phys. (N. Y.)* **529**, 1700169 (2017).
- [40] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [41] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103 (2015).
- [42] D. Muth, R. G. Unanyan, and M. Fleischhauer, *Phys. Rev. Lett.* **106**, 077202 (2011).
- [43] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
- [44] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [45] R. Vosk and E. Altman, *Phys. Rev. Lett.* **110**, 067204 (2013).
- [46] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [47] H. Kim and D. A. Huse, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [48] I. H. Kim, A. Chandran, and D. A. Abanin, *arXiv:1412.3073*.
- [49] D.-L. Deng, X. Li, J. H. Pixley, Y.-L. Wu, and S. Das Sarma, *Phys. Rev. B* **95**, 024202 (2017).
- [50] Y. Bar Lev, G. Cohen, and D. R. Reichman, *Phys. Rev. Lett.* **114**, 100601 (2015).
- [51] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, *Phys. Rev. Lett.* **114**, 160401 (2015).
- [52] A. Chandran, A. Pal, C. R. Laumann, and A. Scardicchio, *Phys. Rev. B* **94**, 144203 (2016).
- [53] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **93**, 060201 (2016).
- [54] D. J. Luitz and Y. B. Lev, *Ann. Phys. (N. Y.)* **529**, 1600350 (2017).
- [55] D. J. Luitz and Y. Bar Lev, *Phys. Rev. B* **96**, 020406 (2017).