



Time-dependent variational principle for mixed matrix product states in the thermodynamic limitYantao Wu *Department of Physics, Princeton University, Princeton, New Jersey 08540, USA* (Received 29 July 2020; accepted 28 September 2020; published 12 October 2020)

We describe a time evolution algorithm for quantum spin chains whose Hamiltonians are composed of an infinite uniform left and right bulk part, and an arbitrary finite region in between. The left and right bulk parts are allowed to be different from each other. The algorithm is based on the time-dependent variational principle (TDVP) of matrix product states. It is inversion-free and very simple to adapt from an existing TDVP code for finite systems. The importance of working in the projective Hilbert space is highlighted. We study the quantum Ising model as a benchmark and an illustrative example. The spread of information after a local quench is studied in both the ballistic and the diffusive case. We also offer a derivation of TDVP directly from symplectic geometry.

DOI: [10.1103/PhysRevB.102.134306](https://doi.org/10.1103/PhysRevB.102.134306)**I. INTRODUCTION**

Over the last two decades, research in quantum dynamics has benefited greatly from numerical algorithms that can simulate accurately the real-time dynamics of many-body quantum systems. For one-dimensional systems, two time evolution algorithms, both based on matrix product states (MPSs), have proved reliable: the time-evolving block decimation (TEBD) method [1] and the time-dependent variational principle (TDVP) algorithm [2,3]. For translationally invariant systems, both methods can generalize to the thermodynamic limit: the iTEBD [4] and the iTDVP [5,6], eliminating the undesirable finite-size effects and reducing the complexity dependence of the system size from linear to constant. Based on locality [7], one expects that for systems composed of uniform left and right bulk parts and finite impurities in between, the time evolution algorithms should also have an efficient thermodynamic version. While it is not clear to us how this can be done for TEBD, a TDVP-based method to deal with such cases has been put forth in [8].

Since [8] was published, tangent-space methods of MPSs have developed significantly [3,5,6,9]. It is thus worthwhile to revisit the problem and apply these developments. In this paper, we greatly simplify the algorithm in [8] and improve it in many ways. While [8] only treats nearest-neighbor interactions, we will be able to treat any Hamiltonian that can be written as a matrix product operator (MPO). Reference [8] also uses inverses of matrices conditioned by the MPS Schmidt coefficients, which can be very small. The algorithm described below will be completely inversion-free. Reference [8] considers only the Hamiltonians whose left and right bulk parts are the same, and the quenches which only change the finite region of impurities. We will allow the left and right bulks to be different and the quenches to change the bulk parts.

The core idea of TDVP is very simple. The states representable by MPSs with a given bond dimension form a submanifold, \mathcal{H}_{MPS} , of the entire Hilbert space [10]. For a state, $|\Psi(t)\rangle$, at time t , the time evolution governed by its

Hamiltonian \hat{H} leads the state out of \mathcal{H}_{MPS} ; i.e., $\hat{H}|\Psi(t)\rangle$ is not in the tangent space of \mathcal{H}_{MPS} at $|\Psi(t)\rangle$. For the time evolution to stay in \mathcal{H}_{MPS} , the TDVP mandates to approximate $\hat{H}|\Psi(t)\rangle$ as its *orthogonal projection* onto the tangent space in the integration of the time evolution. One then chooses a small time step, and integrates the projected $\hat{H}|\Psi(t)\rangle$ to obtain a trajectory in \mathcal{H}_{MPS} . The technical difficulty in applying TDVP to MPSs comes from the gauge freedom in an MPS; i.e., the same quantum state can be represented by two MPSs with very different matrix elements. This means that the time evolution of the quantum state does not uniquely specify how the matrix elements of an MPS should evolve. One thus needs to specify a gauge choice for the MPS and its tangent vectors.

This paper is organized as follows. In Sec. II, we describe the system of interest and its MPS approximation. We will examine very carefully the gauge freedom of the MPS. In Sec. III, we review some facts about the tangent space of \mathcal{H}_{MPS} and provide a gauge choice for the tangent vectors. In Sec. IV, we present the orthogonal projection of $\hat{H}|\Psi\rangle$. The derivation of the results in this section is technical, and is given in Appendix A. In Sec. V, we give an integration scheme to obtain the TDVP dynamics. In Sec. VI, we study the quantum Ising model as an example. The speed of information spreading after a local quench is studied in both the ballistic and the diffusive case. In Sec. VII, we discuss and conclude. For completeness, we give a derivation of the TDVP principle directly from symplectic geometry in Appendix C.

II. THE SYSTEM OF INTEREST, ITS MPS APPROXIMATION, AND GAUGE FREEDOM

We consider an infinite quantum spin chain with a local Hilbert space of dimension d on each site. The system has an infinite left and right bulk part, and a finite region of impurities with length n_W in between. Let the Hamiltonian \hat{H} be written as an infinite MPO with four-index MPO elements $W_{ab}^{ss'}$ with $a, b = 1, \dots, d_W$ and $s, s' = 1, \dots, d$, where d_W is the bond

if the leading eigenvalue of $E_{A_L}^{[W]}$ equals one and is defective. In fact, for a typical MPO, the leading eigenvalue of $E_{A_L}^{[W]}$ is indeed one with algebraic multiplicity two and geometric multiplicity one [9]; i.e., $E_{A_L}^{[W]}$ has one eigenvector and one generalized eigenvector in the leading eigenspace. This behavior can be attributed to the Schur form (lower triangular form) of the W matrix of an MPO [9,12], on which we give a review in Appendix B. We denote the left generalized eigenvector of $E_{A_L}^{[W]}$ by $(L_A^{[W]})$, and the right generalized eigenvector of $E_{A_R}^{[W]}$ by $(R_A^{[W]})$. The $(L_A^{[W]})$ and $(R_A^{[W]})$ can be efficiently computed by an algorithm given in the Appendix of [9]. (They are known as quasifixed points there.) We analogously define $(L_Z^{[W]})$ and $(R_Z^{[W]})$.

We now give the effective parameters, X_A , X_{B^i} , X_Z , and b_L^i , of $|\Phi\rangle_H$:

$$X_A = \text{contraction of } L_A^{[W]}, W_A, V_{A_L}, R_A^{[W]} \quad (17)$$

$$X_Z = \text{contraction of } L_Z^{[W]}, W_Z, V_{Z_n}, R_Z^{[W]} \quad (18)$$

$$X_{B^i} = \text{contraction of } L_A^{[W]}, \prod_{j=1}^{i-1} E_{B_L^j}^{[W]}, W_i, \prod_{k=i+1}^n E_{B_R^k}^{[W]}, R_Z^{[W]} \quad (19)$$

for $i = 1, \dots, n-1$.

$$b_L^i = \text{contraction of } L_A^{[W]}, \prod_{j=1}^{n-1} E_{B_L^j}^{[W]}, W_n, R_Z^{[W]} \quad (20)$$

We can now put Eqs. (17)–(20) back into Eq. (12) to obtain $|\Phi\rangle_H = \text{Proj}_{\mathcal{T}\mathcal{P}\mathcal{H}_{\text{MPS}}} \hat{H} |\Psi\rangle$.

X_A contains no information about B^i and Z , and in fact, is exactly the same effective parameter as in a translationally invariant system composed of only A and W_A [5,6]. Thus, the bulk tensors A and Z should evolve as if they are in an entirely uniform MPS, by the iTDVP algorithm in [5,6]. The effect of the left and the right bulks on the B tensors only comes through the boundary tensors $(L_A^{[W]})$ and $(R_Z^{[W]})$. In fact, in a finite system parametrized only by the B tensors, the tensors at the left (right) boundary have no left (right) indices, and the effective parameters are given by the terms in Eqs. (19) and (20) without the $(L_A^{[W]})$ and $(R_Z^{[W]})$ tensors [3]. Thus, the B matrices can be evolved by the same TDVP algorithm in [3] of a finite system, except under the additional influence of $(L_A^{[W]})$ and $(R_Z^{[W]})$. The only thing unclear is how to patch the

time evolutions of A , B^i , and Z together, which we explain in the next section.

V. INTEGRATION SCHEME

Here we explain how to evolve $|\Psi\rangle$ to $e^{\delta t \hat{H}} |\Psi\rangle$ using $|\Phi\rangle_H$. In iTDVP, one first puts the center site A_C at left infinity. Then one exponentiates the terms in $|\Phi\rangle_H$, one by one from left to right, to sequentially act on the current state. As the algorithm sweeps from left infinity to site 0, the effect of the left boundary tensor decays away and the A_C and C_A tensors converge to their respective limits. The iTDVP algorithm in [5] finds these limits without doing the actual sweep, and is thus very efficient. However, there is something very peculiar about the sweeping process: in obtaining $\{A_C(t + \delta t), C_A(t + \delta t)\}$ from $\{A_C(t), C_A(t)\}$, when the action of one term in $|\Phi\rangle_H$ is completed, one ends up with $C_A(t)$ instead of $C_A(t + \delta t)$ as the bond matrix. [One step of the sweep consists of two half steps, and $C_A(t + \delta t)$ is obtained after the first half step.] See page 35 of [6] or Table I of [5] for the details. This peculiar fact is the key to patch the iTDVP and the finite-TDVP algorithms.

Suppose that at time t , we have a mixed iMPS centered at $B_C^1(t)$:

To make the MPS centered at $A_C(t)$ at left infinity, one needs to borrow a $C_A(t)$ from $B_C^1(t)$, so that one has

One then performs iTDVP on A for δt to arrive at

Thus, the bond matrix $C_A(t)$ cancels, and one next carries out the right sweep of the finite-TDVP algorithm on B for $\delta t/2$ with boundary tensors $(L_{A(t+\delta t)}^{[W]})$ and $(R_{Z(t)}^{[W]})$. Then one does iTDVP on Z for δt and sweeps on B leftward for $\delta t/2$ with boundary tensors $(L_{A(t+\delta t)}^{[W]})$ and $(R_{Z(t+\delta t)}^{[W]})$. This completes the mixed-iTDVP for one step of δt . For a pseudocode, see Table I. We call this algorithm *mixed-iTDVP*. Globally, mixed-iTDVP is second order in δt if A and Z are eigenstates of the bulk Hamiltonian on the left and right, which is the same as the finite-TDVP algorithm. It is first order in δt if A and Z evolve nontrivially, which results from the iTDVP algorithm.

The algorithm can also be used to find the ground state when δt is real and negative. When the time step is infinite, the algorithm reduces to the conventional one-site density matrix renormalization group [13]. When the time step approaches 0, however, the time-evolution algorithm has the benefit of ensuring finding the global energy minimum, as long as the initial state has nonzero overlap with the ground state.

To dynamically expand n , simply upgrade some number of A and Z matrices to be part of B . The procedure used in Sec. VI is that, during the time-evolution process, when the half-chain entanglement entropy at $B^{i=5}$ differs from that at A by more than 10^{-5} , we add five more B tensors equal to A to the left end of the inhomogeneous region. The same is done to the right, too. The fact that n can be expanded dynamically

TABLE I. Pseudocode of mixed-iTDVP for step δt .

algorithm 1 Mixed-iTDVP: Evolving $ \Psi\rangle$ to $e^{\delta t \hat{H}} \Psi\rangle$
Input: MPO tensor $W_A, W_1, \dots, W_n, W_Z$; MPS tensor $\{A_L, A_R, C_A, A_C\}, \{Z_L, Z_R, C_Z, Z_C\}, B_C^1, B_R^2, \dots, B_R^n, L_A^{[W]}, R_Z^{[W]}$; time step δt
Output: MPS tensor $\{A_L, A_R, C_A, A_C\}, \{Z_L, Z_R, C_Z, Z_C\}, B_C^1, B_R^2, \dots, B_R^n, L_A^{[W]}, R_Z^{[W]}$
1: $\{A_L, A_R, C_A, A_C\} \leftarrow \text{iTDVP}(W_A, A_L, A_R, C_A, A_C, \delta t)$
2: Compute $L_A^{[W]}$ with A_L and W_A
3: $\{B_L^1, \dots, B_L^{n-1}, B_C^n\} \leftarrow \text{right sweep of finite-size TDVP}(B_C^1, B_R^2, \dots, B_R^n, L_A^{[W]}, R_Z^{[W]}, \delta t/2)$
4: $\{Z_L, Z_R, C_Z, Z_C\} \leftarrow \text{iTDVP}(W_Z, Z_L, Z_R, C_Z, Z_C, \delta t)$
5: Compute $R_Z^{[W]}$ with Z_R and W_Z
6: $\{B_C^1, B_R^2, \dots, B_R^n\} \leftarrow \text{left sweep of finite-size TDVP}(W_1, \dots, W_n, B_L^1, \dots, B_L^{n-1}, B_C^n, L_A^{[W]}, R_Z^{[W]}, \delta t/2)$

means that one can start with a very small inhomogeneous region at the early times of the time evolution and expand it gradually as time increases. This is an advantage compared to a finite-size algorithm.

VI. EXAMPLE: QUANTUM ISING MODEL

As an illustrative example, we study the quantum dynamics of the quantum Ising chain:

$$\hat{H}_{\text{Ising}} = J \sum_{i=-\infty}^{\infty} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z + \sum_{i=-\infty}^{\infty} (h_x \hat{\sigma}_i^x + h_z \hat{\sigma}_i^z), \quad (21)$$

where $\hat{\sigma}^{x,z}$ are the Pauli matrices. It is integrable when $h_z = 0$ or $h_x = 0$, and is critical when $h_z = 0$ and $h_x/J = \pm 1$ [14]. At criticality, the dispersion relation becomes linear: $E(\mathbf{k}) = v_s |\mathbf{k}|$, giving a characteristic sound velocity $v_s = 2$ [14]. We denote the pre-quenched Hamiltonian by \hat{H}_0 and the post-quenched Hamiltonian by \hat{H}_1 . In the following, $\hat{H}_1 = \hat{H}_0 + \delta \hat{H}$, where $\delta \hat{H}$ is a local field on site i_0 at the middle of region B . When the quench is local, we observe that the entanglement entropy saturates at long time. This means that one can study the quantum dynamics for long times with a relatively small bond dimension, well into the stationary limit.

A. Benchmark

We benchmark our algorithm with \hat{H}_0 with $J = -1$, $h_x = 1.5$, and $h_z = 0$, and $\delta \hat{H} = \hat{\sigma}_{i_0}^x$. This local quench does not break the integrability of the transverse-field Ising chain, and thus the quench dynamics can be computed exactly on a finite chain. We follow [15] to compute the quench dynamics. In Fig. 1, we show the transverse magnetization at site i_0 , $\langle \hat{\sigma}_{i_0}^x(t) \rangle$, as a function of time, obtained both with mixed-iTDVP and the Ising exact solution. As seen, the mixed-iTDVP works correctly well into the stationary regime.

B. Effect of finite size

The defining feature of mixed-iTDVP is that it works directly in the thermodynamic limit. We demonstrate the lack of the finite-size effect by computing the ground state of an inhomogeneous Hamiltonian: $\hat{H}_{\text{Ising}} + \hat{\sigma}_{i_0}^z$ with $J = -1$, $h_x = 1.05$, and $h_z = 0$, where i_0 is in the middle of the chain. The transverse magnetization of the ground state is shown in Fig. 2, in comparison with a finite-size calculation with 500 sites.

C. Speed of information spreading

Here we consider the spread of information after a local quench in the Ising chain both in the ballistic and the diffusive case. In the ballistic case, the system is integrable and admits an extensive number of noninteracting quasiparticles in its spectrum, which transports energy ballistically. When both h_x and h_z are nonzero, however, the Ising chain is no longer integrable, and the only locally conserved quantity is the energy. In this case, there are no ballistically propagating quasiparticles so that, in an extended quantum quench, the energy is transported in a way similar to a random walk, at a speed which is proportional to \sqrt{t} [16]. This is called a diffusive system.

For the ballistic case, we take \hat{H}_0 to be the \hat{H}_{Ising} with $J = -1$, $h_x = 1.5$, and $h_z = 0$. For the diffusive case, we take \hat{H}_0 to be the \hat{H}_{Ising} with $J = 1$, $h_x = 0.9045$, and $h_z = 0.8090$, which is shown to be robustly nonintegrable in [16]. In both cases, the local quench is done through $\delta \hat{H} = \hat{\sigma}_{i_0}^z$, where we place i_0 in the middle of the inhomogeneous region B . To monitor the spread of information, we measure the time dependence of $\langle \hat{\sigma}_i^x \rangle$ on the whole chain, shown in Figs. 3 and 4. The time dependence of other local observables is similar to $\langle \hat{\sigma}_i^x \rangle$.

A very sharp wave front is observed in both cases as the information of the local quench spreads. While this is expected for the ballistic system, it is surprising for the diffusive system, because the energy transports only diffusively in an extended quench. The slope of the wave front can be computed to give the speed of information spreading, v_w . More specifically, we do a linear fit of the function $i_{\text{ridge}}(t)$, which

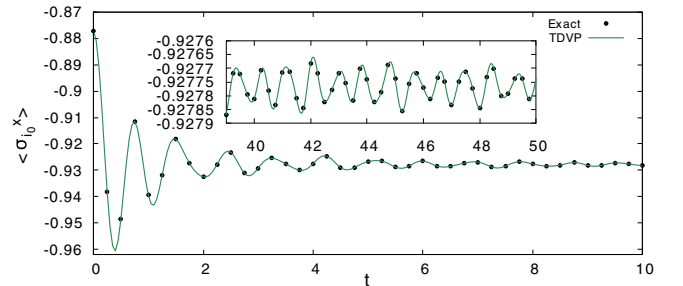


FIG. 1. $\langle \hat{\sigma}_{i_0}^x(t) \rangle$ as a function of time, \hat{H}_0 with $J = -1$, $h_x = 1.5$, and $h_z = 0$, and $\delta \hat{H} = \hat{\sigma}_{i_0}^x$. The mixed-iTDVP computation is done with $\delta t = 0.005$ and $D = 20$. The exact Ising solution is computed for an open chain with 512 sites. The inset is $\langle \hat{\sigma}_{i_0}^x(t) \rangle$ from $t = 40$ to 50.

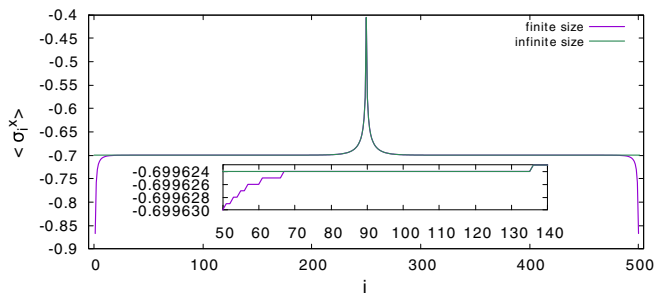


FIG. 2. $\langle \hat{\sigma}_i^x \rangle$ in the ground state of $\hat{H}_{\text{Ising}} + \hat{\sigma}_{i_0}^z$ with $J = -1$, $h_x = 1.05$, and $h_z = 0$. The calculation is done with $D = 20$. The inset is a zoomed-in version of the main plot. The curves for the finite system and the infinite system are overlapping for most times.

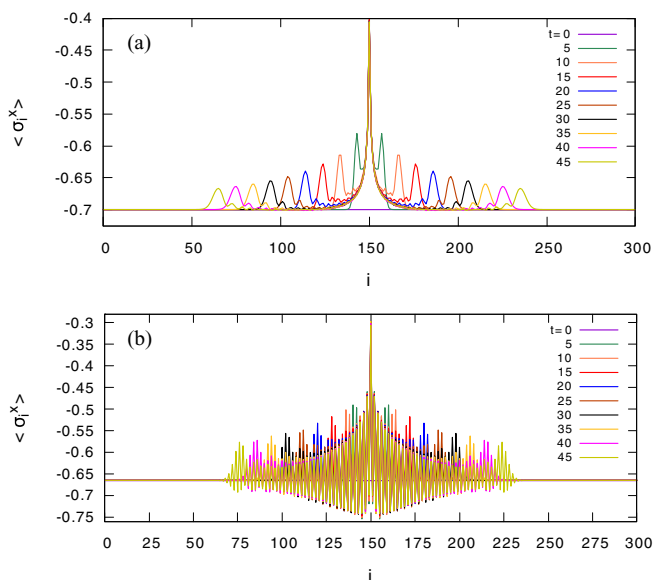


FIG. 3. $\langle \hat{\sigma}_i^x \rangle$ as a function of time, represented in a curve plot for both the ballistic system (top): $\hat{H}_0 = \hat{H}_{\text{Ising}}$ with $J = -1$, $h_x = 1.5$, and $h_z = 0$; and the diffusive system (bottom): $\hat{H}_0 = \hat{H}_{\text{Ising}}$ with $J = 1$, $h_x = 0.9045$, and $h_z = 0.8090$. The quenching Hamiltonian is $\delta \hat{H} = \hat{\sigma}_{i_0}^z$ in both cases. The computation is done with $\delta t = 0.005$ and $D = 20$. Computations with $D = 30$ are also done, and the results are well converged with the bond dimension.

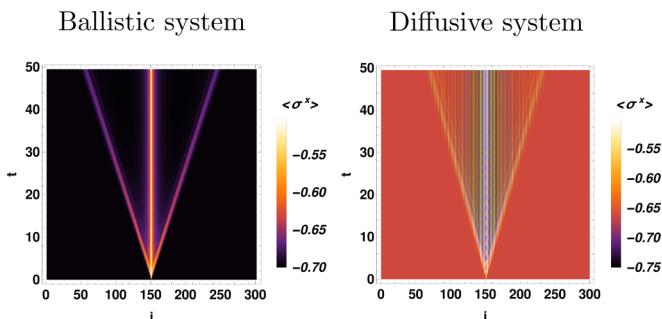


FIG. 4. $\langle \hat{\sigma}_i^x \rangle$ as a function of time, represented in a contour plot, for the same two quenching described in Fig. 3.

TABLE II. Velocity of the wave front in local quenches. The number in parentheses is the uncertainty of the fit on the last digit. R^2 is the R squared of the linear fit.

System	D	v_w	R^2
Ballistic	20	1.94(2)	0.99979
Diffusive	20	1.71(3)	0.99949

for the ballistic system equals the site of the leftmost local maximum of $\langle \hat{\sigma}_i^x \rangle$ at time t , and take the slope of the linear fit as the slope of the wave front. For the diffusive system, we note that there exists a secondary peak in the magnetization profile, for example at around $i = 75$ at $t = 45$. We take the $i_{\text{ridge}}(t)$ to be the site on which $\langle \hat{\sigma}_i^x \rangle$ is the largest in this secondary peak.

The fitted v_w 's are shown in Table II. Because of the discrete nature of i , $i_{\text{ridge}}(t)$ can be ambiguous up to ± 1 . This contributes to the slight nonlinearity of $r_{\text{ridge}}(t)$, indicated by $R^2 < 1$. For the ballistic system, there are well-defined quasi-particles whose velocities are given by the dispersion relation: $E(k) = 2\sqrt{1 - 2h_x \cos(k) + h_x^2}$ [14]. One thus expects that the speed of information spreading should be

$$v_w = \max_{k \in [-\pi, \pi]} \frac{dE(k)}{dk}, \quad (22)$$

which equals 2 for all h_x for the transverse-field Ising model. This is very close to the velocity actually measured in the local quench. The presence of the light cone in the diffusive system, however, suggests that the ballistic spread of information is generic in a local quench, and happens not only in integrable systems.

VII. DISCUSSION

In this paper, we gave a detailed derivation of the TDVP equation for mixed infinite MPSs. The result is a simple combination of the finite-TDVP and infinite-TDVP algorithms, both of which are inversion-free. The method was applied to local quenches of the quantum Ising model, and interesting phenomena were found, which calls for future work. We also expect future work on the algorithmic side. For example, we note that the mixed infinite MPS is very similar to the variational ansatz of the elementary excitations [6] of a translationally invariant system:

$$|\Psi_k\rangle = \sum_x e^{ikx} \dots \textcircled{A} \textcircled{A} \textcircled{B^*} \textcircled{Z} \textcircled{Z} \dots, \quad (23)$$

where x labels the position of spin sites. We thus hope that the current method can help develop a time-evolution algorithm for the elementary excitations.

The code is based on ITensor [17] (version 3, C++) and is available upon request.

ACKNOWLEDGMENTS

The author is grateful for the help received on the ITensor Support Q&A. He is grateful for mentorship from his advisor

Roberto Car at Princeton. He acknowledges support from DOE Award No. DE-SC0017865.

APPENDIX A: DERIVATION OF EQS. (17)–(20)

Before we start, we need some facts about the MPS transfer matrices E_{A_L} and E_{A_R} , defined as

$$E_{A_L} = \begin{array}{c} \text{---} (A_L) \text{---} \\ | \\ \text{---} (A_L) \text{---} \end{array} \quad E_{A_R} = \begin{array}{c} \text{---} (A_R) \text{---} \\ | \\ \text{---} (A_R) \text{---} \end{array}. \quad (\text{A1})$$

We note that the canonical condition, Eq. (10), is the eigenrelation for the nondegenerate leading eigenvalue of the transfer operators, which is 1 for a normalized uniform MPS [6]. This is very important, because it means that if one propagates an arbitrary boundary tensor from left through infinitely many E_{A_L} , only the leading left eigenvector of E_{A_L} survives, which is a two-index delta tensor. The analogous fact is true for E_{A_R} , too.

We now determine the $|\Phi(X_A; X_{B^i}; X_Z; b_L^n)\rangle$ that is the orthogonal projection of $\hat{H}|\Psi\rangle$ on the tangent space at $|\Psi\rangle$. To do this, we need to first compute the inner product $\langle\Phi|\Phi\rangle$, also known as the Gram matrix. Using Eqs. (10) and (12)–(15), we have

$$\begin{aligned} & \langle\Phi(\bar{X}_A; \bar{X}_{B^i}; \bar{X}_Z; \bar{b}_L^n)|\Phi(X_A; X_{B^i}; X_Z; b_L^n)\rangle \\ &= \sum_{m=0}^{\infty} \begin{array}{c} (X_A) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_A) \end{array} \begin{array}{c} (B_R^1) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{B}_R^1) \end{array} + \sum_{m=0}^{\infty} \begin{array}{c} (B_L^n) \\ | \\ (E_{Z_L})^m \\ | \\ (\bar{X}_Z) \end{array} \\ &+ \sum_{i=1}^{n-1} \begin{array}{c} (X_{B^i}) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_{B^i}) \end{array} + \begin{array}{c} (b_L^n) \\ | \\ (E_{Z_L})^m \\ | \\ (\bar{b}_L^n) \end{array}. \end{aligned}$$

To simplify $\langle\Phi|\Phi\rangle$ further, we explicitly split out the contribution of E_{A_R} from its leading eigenspace:

$$E_{A_R} = \begin{array}{c} \text{---} (A_R) \text{---} \\ | \\ \text{---} (A_R) \text{---} \end{array} + \tilde{E}_{A_R} \quad (\text{A2})$$

where l_{A_R} is the leading left eigenvector of E_{A_R} , and \tilde{E}_{A_R} is the contribution from the subleading eigenspace of E_{A_R} . Then,

$$\sum_{m=0}^{\infty} (E_{A_R})^m = \sum_{m=0}^{\infty} \begin{array}{c} \text{---} (A_R) \text{---} \\ | \\ \text{---} (A_R) \text{---} \end{array} + \sum_{m=0}^{\infty} (\tilde{E}_{A_R})^m. \quad (\text{A3})$$

This splitting is useful because \tilde{E}_{A_R} has a spectral radius less than one, and the second term on the right-hand side of Eq. (A3) converges. We now have

$$\sum_{m=0}^{\infty} \begin{array}{c} (X_A) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_A) \end{array} \begin{array}{c} (B_R^1) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{B}_R^1) \end{array} = \sum_{m=0}^{\infty} \begin{array}{c} (X_A) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_A) \end{array} + F_A \quad (\text{A4})$$

where F_A is a finite number. Here we have used the normalization of the state:

$$\langle\Psi|\Psi\rangle = \begin{array}{c} (B_R^1) \\ | \\ (A_R) \\ | \\ (B_R^1) \end{array} = 1. \quad (\text{A5})$$

A relation analogous to Eq. (A4) holds for the Z tensors, too. This gives the final form of the Gram matrix:

$$\begin{aligned} & \langle\Phi(\bar{X}_A; \bar{X}_{B^i}; \bar{X}_Z; \bar{b}_L^n)|\Phi(X_A; X_{B^i}; X_Z; b_L^n)\rangle \\ &= \sum_{m=0}^{\infty} \begin{array}{c} (X_A) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_A) \end{array} \begin{array}{c} (B_R^1) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{B}_R^1) \end{array} + \sum_{m=0}^{\infty} \begin{array}{c} (X_Z) \\ | \\ (E_{Z_L})^m \\ | \\ (\bar{X}_Z) \end{array} + \begin{array}{c} (b_L^n) \\ | \\ (E_{Z_L})^m \\ | \\ (\bar{b}_L^n) \end{array} \\ &+ \sum_{i=1}^{n-1} \begin{array}{c} (X_{B^i}) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_{B^i}) \end{array} + F_A(X_A) + F_Z(X_Z). \end{aligned} \quad (\text{A6})$$

The Gram matrix is thus essentially diagonal in the effective parameters of $|\Phi\rangle$.

We are ready to compute the orthogonal projection of $\hat{H}|\Psi\rangle$, which is given by the solution to the minimization problem

$$\min_{X_A, X_{B^i}, X_Z, b_L^n} \|\hat{H}|\Psi\rangle - |\Phi(X_A; X_{B^i}; X_Z; b_L^n)\rangle\|_2^2.$$

X_A is determined by

$$\frac{\partial\langle\Phi|\Phi\rangle}{\partial X_A} = \frac{\partial F_A}{\partial X_A} + \sum_{m=0}^{\infty} \begin{array}{c} (X_A) \\ | \\ (E_{A_R})^m \\ | \\ (\bar{X}_A) \end{array} = \frac{\partial\langle\Phi|\hat{H}|\Psi\rangle}{\partial X_A}. \quad (\text{A7})$$

Here,

$$\begin{aligned} & \frac{\partial\langle\Phi|\hat{H}|\Psi\rangle}{\partial X_A} \\ &= \sum_{m=0}^{\infty} \dots \begin{array}{c} (A_C) \\ | \\ (E_{A_L}^{[W]})^m \\ | \\ (\bar{V}_{A_L}) \end{array} \begin{array}{c} (W_A) \\ | \\ (E_{A_R}^{[W]})^m \\ | \\ (\bar{V}_{A_L}) \end{array} \prod_{i=1}^n \begin{array}{c} (E_{B^i}^{[W]}) \\ | \\ (E_{B^i}^{[W]}) \\ | \\ (\bar{V}_{A_L}) \end{array} \begin{array}{c} (E_{Z_R}^{[W]}) \\ | \\ (E_{Z_R}^{[W]}) \\ | \\ (\bar{V}_{A_L}) \end{array} \dots \end{aligned} \quad (\text{A8})$$

where the MPO transfer matrices $E_{A_L}^{[W]}$, etc., are defined in Eq. (16). In addition to their generalized eigenvectors, we denote the left eigenvectors of $E_{A_L}^{[W]}$ and right eigenvectors of $E_{A_R}^{[W]}$ respectively as $\langle I_1|$ and $|I_{d_W}\rangle$. In fact, these eigenvectors do not depend on the values of the MPO, and thus are the same for E_{Z_L} and E_{Z_R} (see Appendix B). As the left boundary tensor at left infinity propagates through infinitely many $E_{A_L}^{[W]}$ to meet the center site A_C in Eq. (A8), only the leading eigenspace survives. The same applies to the right side. Thus,

$$\frac{\partial\langle\Phi|\hat{H}|\Psi\rangle}{\partial X_A} = [(\langle I_1^{[W]}| + \alpha\langle I_1|)E_C[|R_Z^{[W]}] + \beta|I_{d_W}\rangle], \quad (\text{A9})$$

where

$$E_C \equiv \begin{array}{c} \text{---} A_C \text{---} \\ | \\ \text{---} W_A \text{---} \\ | \\ \text{---} \bar{V}_{A_L} \text{---} \end{array} \left(E_{A_R}^{[W]} \right)^m \prod_{i=1}^n E_{B_R}^{[W]} \quad (\text{A10})$$

Here, α and β are two complex numbers. They occur because every time $(L_A^{[W]}|$ passes through $E_{A_L}^{[W]}$, there arises a new term of $(I_1|: (L_A^{[W]}|E_{A_L}^{[W]} = (L_A^{[W]}| + e(I_1|$, where e is the energy density of the chain [9]. Their values, however, do not matter because of the following lemmas.

Lemma 1: $(I_1|E_C = 0$. (This lemma, and others below, are based on the Schur form of the MPO. See Appendix B for a discussion of their proofs.)

Lemma 2: $(L_A^{[W]}|E_C|I_{d_w}) = 0$.

Thus,

$$\frac{\partial \langle \Phi | \hat{H} | \Psi \rangle}{\partial \bar{X}_A} = \sum_{m=0}^{\infty} \begin{array}{c} \text{---} A_C \text{---} \\ | \\ \text{---} L_A^{[W]} \text{---} \\ | \\ \text{---} W_A \text{---} \\ | \\ \text{---} \bar{V}_{A_L} \text{---} \end{array} \left(E_{A_R}^{[W]} \right)^m \prod_{i=1}^n E_{B_R}^{[W]} R_Z^{[W]} \quad (\text{A11})$$

As with E_{A_R} , we split out of $E_{A_R}^{[W]}$ the term associated with the leading eigenspace. To do this, we need the following lemma in linear algebra.

Lemma 3: Let E be a matrix with leading eigenvalue one, according to which there is one eigenvector and one generalized eigenvector. Let $(v_1|$ be the left generalized eigenvector, $(v_2|$ the left eigenvector, $|u_1\rangle$ the right eigenvector, and $|u_2\rangle$ the right generalized eigenvector. Then, for an integer $m > 0$,

$$E^m = |u_1\rangle(v_1| + m|u_1\rangle(v_2| + |u_2\rangle(v_2| + \tilde{E}^m, \quad (\text{A11})$$

where \tilde{E} is the contribution to E from the subleading eigenspace.

When applying Lemma 3 to $E_{A_R}^{[W]}$, the contribution associated with the $|u_1\rangle = |I_{d_w}\rangle$ drops because of the following lemma.

Lemma 4:

$$\begin{array}{c} \text{---} A_C \text{---} \\ | \\ \text{---} L_A^{[W]} \text{---} \\ | \\ \text{---} W_A \text{---} \\ | \\ \text{---} \bar{V}_{A_L} \text{---} \end{array} I_{d_w} = 0. \quad (\text{A12})$$

Thus, we have

$$\begin{aligned} \frac{\partial \langle \Phi | \hat{H} | \Psi \rangle}{\partial \bar{X}_A} &= \sum_{m=0}^{\infty} \begin{array}{c} \text{---} A_C \text{---} \\ | \\ \text{---} L_A^{[W]} \text{---} \\ | \\ \text{---} W_A \text{---} \\ | \\ \text{---} \bar{V}_{A_L} \text{---} \end{array} \left(E_{A_R}^{[W]} \right)^m \prod_{i=1}^n E_{B_R}^{[W]} R_Z^{[W]} \\ &+ \begin{array}{c} \text{---} A_C \text{---} \\ | \\ \text{---} L_A^{[W]} \text{---} \\ | \\ \text{---} W_A \text{---} \\ | \\ \text{---} \bar{V}_{A_L} \text{---} \end{array} \sum_{m=0}^{\infty} \left(E_{A_R}^{[W]} \right)^m \prod_{i=1}^n E_{B_R}^{[W]} R_Z^{[W]}, \end{aligned} \quad (\text{A13})$$

where we have made use of the following lemma.

Lemma 5:

$$l_{A_R}^{[W]} \prod_{i=1}^n E_{B_R}^{[W]} R_Z^{[W]} = 1, \quad (\text{A14})$$

where $l_{A_R}^{[W]}$ is the left eigenvector of $E_{A_R}^{[W]}$.

Note that the second term of Eq. (A13) converges. Now substitute Eq. (A13) into Eq. (A7), and divide the equation by $\sum_{m=0}^{\infty} 1$. The finite terms drop, and we obtain Eq. (17). Analogously, we obtain Eqs. (18) and (19).

We now determine b_L^n , which is given by

$$\begin{aligned} \frac{\partial \langle \Phi | \Phi \rangle}{\partial \bar{b}_L^n} &= -\langle b_L^n | = \frac{\partial \langle \Phi | \hat{H} | \Psi \rangle}{\partial \bar{b}_L^n} \\ &= \left[(L_A^{[W]}| + \alpha(I_1|) \right] E_D \left[|R_Z^{[W]}\rangle + \beta|I_{d_w}\rangle \right], \end{aligned}$$

where

$$E_D \equiv \prod_{i=1}^{n-1} E_{B_L}^{[W]} \begin{array}{c} \text{---} B_C^n \text{---} \\ | \\ \text{---} W_n \text{---} \end{array} \quad (\text{A15})$$

Here the α and β are the same as in Eq. (A9). Two lemmas are now in order.

Lemma 6: $(I_1|E_D|I_{d_w}) = 0$.

Lemma 7: $(I_1|E_D|R_Z^{[W]}) = (L_A^{[W]}|E_D|I_{d_w}) = B_C^n$.

Thus,

$$b_L^n = (L_A^{[W]}|E_D|R_Z^{[W]}) + (\alpha + \beta)B_C^n. \quad (\text{A16})$$

But note that $b_L^n = (\alpha + \beta)B_C^n$ gives a contribution of $(\alpha + \beta)|\Psi\rangle$ to $|\Phi\rangle$, which can be dropped in the projective space. Also recall that we still have one last gauge symmetry to spare, which we now use to demand $\alpha + \beta = 0$ so that $b_L^n = (L_A^{[W]}|E_D|R_Z^{[W]})$ in Eq. (20).

APPENDIX B: SCHUR FORM OF MPO

As discussed in the main text, the W matrix of an MPO is lower triangular, known as the Schur form. For example, in

Both the Hamiltonian function and the symplectic form have a restriction on M :

$$H_M = H \circ \text{inc}, \quad \Omega_M = \text{inc}^* \Omega. \quad (\text{C8})$$

Because the exterior differentiation d and the pullback inc^* commute, $d\Omega_M = 0$, and thus M is also symplectic. We now look for the Hamiltonian flow X_{H_M} associated with H_M on M . For all $\xi, \eta \in (T_\Psi M)^+$, we look for $X_{H_M} \in (T_\Psi M)^{\mathbb{C}}$ such that $\Omega_M(X_{H_M}, \xi + \bar{\eta}) = dH_M(\xi + \bar{\eta})|_\Psi$.

$$\begin{aligned} dH_M(\xi + \bar{\eta})|_\Psi &= dH(\text{inc}^*(\xi + \bar{\eta}))|_\Psi \\ &= dH(\xi + \bar{\eta})|_\Psi \\ &= I(\eta, \hat{H}\Psi) + I(\hat{H}\Psi, \xi). \end{aligned} \quad (\text{C9})$$

Now here is the key: Because ξ, η are both only in $(T_\Psi M)^+$, $\hat{H}\Psi$ can be replaced with its orthogonal projection on $(T_\Psi M)^{\mathbb{C}}$, $\text{Proj}\hat{H}\Psi$:

$$\begin{aligned} dH_M(\xi + \bar{\eta}) &= I(\eta, \text{Proj}\hat{H}\Psi) + I(\text{Proj}\hat{H}\Psi, \xi) \\ &= \Omega(X_{H_M}, \xi + \bar{\eta}), \end{aligned} \quad (\text{C10})$$

where X_{H_M} is the Hamiltonian flow of H_M on M :

$$X_{H_M} = -i\text{Proj}\hat{H}\Psi + \overline{-i\text{Proj}\hat{H}\Psi}. \quad (\text{C11})$$

This gives the TDVP dynamics on M and the dynamics is symplectic.

-
- [1] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [2] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-Dependent Variational Principle for Quantum Lattices, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [3] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Unifying time evolution and optimization with matrix product states, *Phys. Rev. B* **94**, 165116 (2016).
- [4] G. Vidal, Classical Simulation of Infinite-Size Quantum Lattice Systems in One Spatial Dimension, *Phys. Rev. Lett.* **98**, 070201 (2007).
- [5] J. C. Halimeh and V. Zauner-Stauber, Dynamical phase diagram of quantum spin chains with long-range interactions, *Phys. Rev. B* **96**, 134427 (2017).
- [6] L. Vanderstraeten, J. Haegeman, and F. Verstraete, Tangent-space methods for uniform matrix product states, *SciPost Phys. Lect. Notes* **7** (2019).
- [7] M. B. Hastings, Locality in quantum systems, [arXiv:1008.5137](https://arxiv.org/abs/1008.5137).
- [8] A. Milsted, J. Haegeman, T. J. Osborne, and F. Verstraete, Variational matrix product ansatz for nonuniform dynamics in the thermodynamic limit, *Phys. Rev. B* **88**, 155116 (2013).
- [9] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, Variational optimization algorithms for uniform matrix product states, *Phys. Rev. B* **97**, 045145 (2018).
- [10] J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, Geometry of matrix product states: Metric, parallel transport, and curvature, *J. Math. Phys.* **55**, 021902 (2014).
- [11] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, *Commun. Math. Phys.* **28**, 251 (1972).
- [12] L. Michel and I. P. McCulloch, Schur forms of matrix product operators in the infinite limit, [arXiv:1008.4667](https://arxiv.org/abs/1008.4667).
- [13] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011).
- [14] J. B. Kogut, An introduction to lattice gauge theory and spin systems, *Rev. Mod. Phys.* **51**, 659 (1979).
- [15] A. P. Young, Finite-temperature and dynamical properties of the random transverse-field Ising spin chain, *Phys. Rev. B* **56**, 11691 (1997).
- [16] H. Kim and D. A. Huse, Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [17] ITensor Library (version 3.1.3), <http://itensor.org>.
- [18] M. Nakahara, *Geometry, Topology, and Physics* (Taylor and Francis, New York, 2003).