Matrix product state algorithms for Gaussian fermionic states

Norbert Schuch

Max-Planck-Institute of Quantum Optics, Hans-Kopfermann-Str. 1, 85748 Garching, Germany
and Munich Center for Quantum Science and Technology, Schellingstraße 4, 80799 München, Germany

Bela Bauer

Station Q, Microsoft Corporation, Santa Barbara, California 93106, USA

While general quantum many-body systems require exponential resources to be simulated on a classical computer, systems of noninteracting fermions can be simulated exactly using polynomially scaling resources. Such systems may be of interest in their own right but also occur as effective models in numerical methods for interacting systems, such as Hartree-Fock, density functional theory, and many others. Often it is desirable to solve systems of many thousand constituent particles, rendering these simulations computationally costly despite their polynomial scaling. We demonstrate how this scaling can be improved by adapting methods based on matrix product states, which have been enormously successful for low-dimensional interacting quantum systems, to the case of free fermions. Compared to the case of interacting systems, our methods achieve an exponential speedup in the entanglement entropy of the state. We demonstrate their use to solve systems of up to one million sites with an effective matrix product state bond dimension of 10^{15}.

I. INTRODUCTION

Noninteracting or weakly interacting fermions play a crucial role in many aspects of condensed matter theory and quantum chemistry. For example, using density functional theory, many materials can be accurately approximated by weakly interacting fermions; Hartree-Fock theory is an excellent starting point for quantum chemical calculations; and transport in electronic nanodevices is often captured by noninteracting models. While the computational effort in all these applications scales polynomially with the size of the system that must be considered, the desire to obtain a realistic description—reached, for example, by including many orbitals in a complex material or a microscopic multiband model for a nanostructure—means that computing properties of a noninteracting fermion model can nevertheless become a computational bottleneck.

For interacting quantum systems, tensor networks have in many cases become the tool of choice for numerical simulations. They exploit the locality of entanglement in the low-energy states of local quantum Hamiltonians for a more compact classical description of these states. The prototypical examples are matrix product states [1,2], which form the basis of the density-matrix renormalization group (DMRG) [3]. Since its inception, DMRG has become the standard method to solve one-dimensional quantum systems [4,5]. Many extensions of matrix product states to higher-dimensional systems have been suggested [6–13]; for recent reviews, see Refs. [14–16].

In this paper, we will show that it is possible to apply the benefits of tensor networks to noninteracting fermions to obtain significantly more efficient numerical methods for the simulation of very large free-fermion systems, allowing us to simulate systems far beyond what is possible with existing methods. While tensor network states have previously been described for systems of free fermions [17–25], thus far they have not been used primarily as numerical tools. In this work, we will describe practical tensor network algorithms to compute the many-body ground state as well as the dynamics of many-body states for free-fermion systems. Our methods capture the full equal-time Green's function of the system, from which any physical observable can be computed. While exact methods for this property scale cubically with the system size, our approach scales linearly with the length of a quasi-one-dimensional system and thus outperforms exact methods significantly.

We work within the framework of Gaussian fermionic matrix product states (GFMPS) [19]. These states are constructed using Gaussian fermionic states, which are the most general class of states for which a Wick's theorem holds, i.e., whose equal-time correlation functions are characterized fully through the equal-time single-particle Green's function. We will review this formalism in Sec. II, where we also introduce the required building blocks for Gaussian tensor network algorithms, such as contraction and decomposition of tensors. We introduce the specific case of GFMPS in Sec. III, where we also discuss the key technical tool of our work: a *canonical form* for GFMPS. This crucial technical step will allow us to generalize many well-established numerical algorithms for matrix product states, such as single- and two-site DMRG as well as time evolution approaches such as TDVP [26,27] to the Gaussian setting; this will be discussed in Secs. IV and V. Finally, in Sec. VI we numerically demonstrate the performance of our algorithms.

A. Quadratic fermion systems

To set the stage, consider a lattice system of N sites, each comprised of some number n_s of fermionic modes. For our purposes, it will be convenient to work in a basis of self-adjoint, so-called Majorana fermions c_i, with $c_i^\dagger = c_i$ and $\{c_i, c_j\} = 2\delta_{ij}$. A fermionic system comprised of conventional complex fermions created by some operator $a_i^\dagger$ can always be rewritten in such a Majorana basis by taking $c_{2i-1} = a_i + a_i^\dagger$, $c_{2i} = -i(a_i - a_i^\dagger)$. We will be interested in quadratic Hamiltonian, i.e., Hamiltonians of the form

$$\mathcal{H} = -i \sum_{ij} H_{ij} c_i c_j + E, \qquad (1)$$

with E a constant offset. We denote the total number of Majorana modes, which must be even in a physical system, by $M = 2 \sum n_s$. Thus, H is a real matrix of size $M \times M$, which we can choose without loss of generality antisymmetric, $H^{\mathrm{T}} = -H$, and we will do so in the following.

To compute properties of such a system, one can fully diagonalize the matrix H to obtain all single-particle eigenvalues and eigenvectors, from which all other observables can be constructed. In the absence of any special structure of H, this will scale as $\mathcal{O}(N^3)$. If, on the other hand, only a few low-lying single-particle eigenstates (or a few states near a particular energy, such as the Fermi energy) are desired and the matrix A is sufficiently sparse (as is the case for short-range hopping models), one can use sparse Krylov-space diagonalization methods such as Lanczos or shift-and-invert methods. These are generally expected to scale linearly in N and the number of nonzero elements per row of A, as well as exhibit some dependence on spectral properties such as the energy gap. However, it should be emphasized that having computed only a few low-lying states, physical observables such as the density or, more generally, the equal-time Green's function cannot be computed.

There are many other methods that overcome this limitation. A very efficient method to compute the spectral density of some operator (as well as certain dynamical correlation functions) in a time that scales only linearly with the number of modes is the kernel polynomial method (KPM) [28]. There is also a large variety of Green's-function-based methods. To compute the frequency-dependent Green's function, the recursive-Green's-function (RGF) approach [29,30] can be used. For a one-dimensional system, this method scales linearly with the length; for two- or three-dimensional systems, the method can be applied by taking the system to be a one-dimensional collection of "slices." However, to obtain equal-time properties, a frequency integral must be taken. Finally, the nonequilibrium Green's function approach [31] incorporates systems driven out of equilibrium, for example, by an applied voltage in a nanodevice. For a recent review and a related wave-function-based approach, see Ref. [32]. Finally, in the context of density functional theory, several approximate methods to solve the Kohn-Sham equation in linear time have been developed (see, e.g., Ref. [33]). In contrast to these methods, our GFMPS method is based on the *many-body* wave function for a low-energy or time-dependent state of the system and thus allows the straightforward calculation of arbitrary expectation values without having to integrate over frequencies. Furthermore, our approach gives access to quantities such as the entanglement spectrum and entanglement entropy of the state.

B. Gaussian tensor networks

The power of matrix product states is most easily understood from the properties of the Schmidt decomposition. For a quantum state $|\psi\rangle$ on a tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$, it is given by

$$|\psi\rangle = \sum_{\alpha=1}^{M} s_\alpha |A_\alpha\rangle |B_\alpha\rangle, \qquad (2)$$

where $s_\alpha \geqslant 0$ are real, and the *Schmidt vectors* $|A_\alpha\rangle$, $|B_\alpha\rangle$ are orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. The von Neumann entanglement entropy between A and B is given by $S = -\operatorname{Tr}(\rho_A \log \rho_A) = -\sum s_\alpha^2 \log s_\alpha^2$ and is bounded by $S \leqslant \log(M)$. Thus, coarsely speaking, the number of terms that are relevant grows exponentially with the bipartite entanglement of the state.

Given that low-energy states of local Hamiltonians are weakly entangled in the sense that they exhibit area-law scaling of the entanglement entropy [34–37], one can exploit the properties of the Schmidt decomposition to construct more compact representations of such states. In particular, it is natural to assume that the Schmidt values s_α decay very quickly with α in such weakly entangled states [35–41]. This directly leads to the construction of matrix product states, which can be obtained by recursively performing a Schmidt decomposition between all sites of a one-dimensional system and truncating the sum on each bond to the D largest Schmidt values s_α. The computational cost of such a state will scale exponentially with the bipartite entanglement entropy.

In systems of noninteracting fermions, the Schmidt decomposition and thus also the bipartite entanglement take a simpler form. This is due to the fact that the reduced density matrix for a subsystem is itself a Gibbs state of a quadratic Hamiltonian [42–45]. Thus, even its many-body spectrum can be obtained through single-particle modes and their eigenvalues, and the entanglement can be fully characterized through *modewise* entanglement [42,45] between single-particle modes in each part of the bipartite system. In other words, the Schmidt decomposition of a Gaussian state can itself be expressed entirely in terms of Gaussian states. Crucially, each mode can contribute a fixed amount to the bipartite entanglement, and the maximum entanglement is thus linear in the number of modes rather than logarithmic in the number of terms in the Schmidt decomposition. This constitutes an exponential compression of the Schmidt decomposition in the Gaussian case.

To exploit this computationally, one can construct a Gaussian fermionic matrix product state (MPS) in an analogous fashion to the interacting case by recursively applying the Schmidt decomposition on each bond of a given Gaussian state. Owing to the structure of Gaussian states, such a GFMPS will scale only polynomially rather than exponentially with the amount of bipartite entanglement.

This exponential speedup also benefits more general tensor networks. In the context of a conventional tensor network, a

tensor is nothing but a high-dimensional array of numbers, which are enumerated by an array of indices. The number of indices is commonly referred to as "tensor rank" R (not to be confused with the matrix rank), and each index runs over integers $1, \ldots, D$, where D is referred to as the bond dimension. A quantum state is expressed through such a tensor network by identifying some indices of the tensors with physical degrees of freedom and contracting (summing) over the remaining ones in some predetermined pattern.

To understand the case of a Gaussian tensor network, it is convenient to think of a rank-R tensor of bond dimension D as a quantum state on the tensor product of R D-dimensional Hilbert spaces. One can now consider the case where this state is Gaussian, i.e., it satisfies a Wick theorem and can thus be described completely through the expectation values of quadratic operators. Such a representation is exponentially more compact, i.e., requires only $\mathcal{O}(R \log D)$ rather than $\mathcal{O}(D^R)$ numbers. Using techniques that we will discuss in more detail below, one can generalize the contraction of tensors and other important operations on tensor networks to this framework.

II. COVARIANCE MATRIX FORMALISM FOR GAUSSIAN TENSOR NETWORKS

We will now introduce Gaussian tensor networks on a more technical level. This will establish the essential tools that we will use for the construction of Gaussian MPS in Sec. III as well as specific optimization algorithms in Sec. IV. We first review the fermionic covariance matrix formalism. We follow the conventions of Ref. [46]; for further details, see also Ref. [47]. Throughout this section, we will assume a system of $2n$ Majorana fermions described by a Hamiltonian of the form (1).

A. Covariance matrix formalism

Gaussian states are states that are fully characterized through their equal-time two-point correlation functions. A convenient formalism for describing a (possibly mixed) Gaussian state described by a density matrix ρ is the covariance matrix (CM), which is given by

$$\gamma_{ij} = \frac{\mathrm{i}}{2}\mathrm{tr}(\rho[c_i, c_j]). \tag{3}$$

The covariance matrix γ is real and antisymmetric, and satisfies $\gamma^2 \geqslant -\mathbb{1}$; for pure states,

$$\gamma^2 = -\mathbb{1}. \tag{4}$$

In the following, we discuss some further properties of quadratic Hamiltonians and Gaussian states.

a. Energy of a state and normal form. The energy of a state ρ under the Hamiltonian $\mathcal{H}$ [Eq. (1)] is

$$
\begin{aligned}
\mathrm{tr}[\mathcal{H}\rho] &= -\sum H_{ij} \,\mathrm{i}\,\mathrm{tr}(\rho c_i c_j) + E \\
&= -\sum H_{ij}\gamma_{ij} + E \\
&= \mathrm{tr}(H\gamma) + E.
\end{aligned}
\tag{5}
$$

Any real antisymmetric matrix A (such as H or γ) can be brought into a canonical form by an orthogonal

transformation,

$$OAO^\mathsf{T} = \bigoplus_{k=1}^{n}\begin{pmatrix} & \lambda_k \\ -\lambda_k & \end{pmatrix}, \tag{6}$$

with $\lambda_k \geqslant 0$. If A is of odd size, additionally a single 0 block appears. Note that $\pm\lambda_k$ are exactly the eigenvalues of the Hermitian matrix iA.

We can use this to bring a Hamiltonian of the form Eq. (1) into a normal form

$$\mathcal{H} = -\frac{\mathrm{i}}{2}\sum_{k=1}^{n}\epsilon_k[\eta_{2k-1}, \eta_{2k}], \tag{7}$$

where η_k are new canonical Majorana operators given by $\eta_k = \sum_i O_{ki}c_i$, with $OHO^\mathsf{T} = \bigoplus_{k=1}^{n}(_{-\epsilon_k/2} {}^{\epsilon_k/2})$. The ϵ_k correspond exactly to the non-negative half of the single-particle eigenvalues of the equivalent Bogoliubov–de Gennes Hamiltonian. The ground state of $\mathcal{H}$, which we will denote as $|0\rangle$, is characterized by having $i\eta_{2k-1}\eta_{2k}|0\rangle = |0\rangle$ for all $k = 1, \ldots, n$, and thus has total energy $E_0 = -\sum\epsilon_k$. Its covariance matrix in the basis of canonical operators η_k is thus given by

$$\tilde{\gamma}_0 = \frac{i}{2}\langle 0|[\eta_i, \eta_j]|0\rangle = \bigoplus\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}. \tag{8}$$

Using the orthogonal matrix O that brings A into the normal form, we can obtain the covariance matrix in the basis of local, physical Majorana operators c_i as $\gamma_0 = O^\mathsf{T}\tilde{\gamma}_0 O$. Using Eq. (5), we can verify that the ground-state energy is indeed $E_0 = -\sum\epsilon_k$. Numerically, γ_0 can be determined by diagonalizing the antisymmetric matrix H (giving imaginary eigenvalues) and then replacing them by the sign of the imaginary part. For other approaches, see Ref. [48].

b. Expectation values and state overlaps. The expectation value of arbitrary operators in a Gaussian state can be computed using Wick's theorem, which takes a particularly simple form in the representation of Gaussian covariance matrices. Consider a product of Majorana operators $C = \prod_{i\in\mathcal{I}}c_i$. The expectation value of C in a state $|\phi\rangle$ with CM γ is given by [46]

$$\langle\phi|C|\phi\rangle = \mathrm{Pf}(i\gamma_\mathcal{I}), \tag{9}$$

where $\gamma_\mathcal{I}$ is the submatrix of M with row and column indices in $\mathcal{I}$, and $\mathrm{Pf}(\cdot)$ denotes the Pfaffian.

The square modulus of the overlap between two states is easily determined from the covariance formalism [46,49]. Consider two many-body states $|A\rangle$ and $|B\rangle$ with corresponding covariance matrices γ_A and γ_B and the same fermion parity, i.e., $\mathrm{Pf}(i\gamma_A) = \mathrm{Pf}(i\gamma_B) = p$. Then,

$$|\langle A|B\rangle|^2 = p2^{-n}\,\mathrm{Pf}(\gamma_A + \gamma_B). \tag{10}$$

When the phase of the overlap is also desired, a complication arises from the fact that the covariance matrix formalism does not capture the global phase of the state—clearly for some fermionic state $|\psi\rangle$, all states $e^{i\phi}|\psi\rangle$ will have the same expectation values of physical operators and thus the same covariance matrix. To overcome this problem, one can introduce a reference state $|C\rangle$ and observe that $\langle C|A\rangle\langle A|B\rangle\langle B|C\rangle$ is invariant under multiplying any of the three states by a phase.

Therefore, it can be computed from the covariance formalism; for explicit expressions, see Ref. [46].

c. Composite systems. Consider a Gaussian state described by a covariance matrix γ' on N modes and γ'' on M modes. The joint state γ is characterized by the direct sum of the covariance matrices,

$$\gamma = \gamma' \oplus \gamma'' = \begin{pmatrix} \gamma' & 0 \\ 0 & \gamma'' \end{pmatrix}. \tag{11}$$

To describe a general state of such a joint system, it is convenient to introduce collective labels for the Majorana modes in each subsystem. For example, we can denote the N modes of the state γ' as a, and the M modes of the other system as b. A general Gaussian state of the joint system thus takes the block form

$$\gamma = \begin{pmatrix} \gamma_{aa} & \gamma_{ab} \\ -\gamma_{ab}^{\mathsf{T}} & \gamma_{bb} \end{pmatrix}, \tag{12}$$

where γ_{aa} contains the correlations of the a modes, γ_{ab} the correlations of a modes with b modes, etc. We will occasionally write $|a|$ for the number of modes in a, i.e., $|a| = N$. More generally, for multipartite systems, such as a tripartite system with modes a, b, and c, we can express γ in different partitions, e.g.,

$$\gamma = \begin{pmatrix} \gamma_{aa} & \gamma_{ab} & \gamma_{ac} \\ \gamma_{ba} & \gamma_{bb} & \gamma_{bc} \\ \gamma_{ca} & \gamma_{cb} & \gamma_{cc} \end{pmatrix} = \begin{pmatrix} \gamma_{ab|ab} & \gamma_{ab|c} \\ \gamma_{c|ab} & \gamma_{c|c} \end{pmatrix}, \tag{13}$$

where we use a vertical bar to separate row and column indices when ambiguous, e.g., $\gamma_{ab|c}$ holds the correlations of the modes a and b vs c, i.e., $\gamma_{ab|c} = \begin{pmatrix} \gamma_{ac} \\ \gamma_{bc} \end{pmatrix}$. We will generally follow this notation throughout this paper. Note that due to the explicit labeling of the modes in this notation, the actual ordering of the blocks in the matrix does not matter.

d. Partial traces. The dual operation to composing systems is the partial trace, i.e., considering only the state of a subsystem; we will make frequent use of it in this paper. Given a system composed of $N + M$ modes a and b with covariance matrix

$$\gamma = \begin{pmatrix} \gamma_{aa} & \gamma_{ab} \\ -\gamma_{ab}^{\mathsf{T}} & \gamma_{bb} \end{pmatrix}, \tag{14}$$

the covariance matrix which describes the state of the a modes (i.e., the system after tracing out the b modes) is given by γ_{aa}; this follows immediately from the definition of the covariance matrix.

B. Tensor networks and contraction

a. Formalism. Tensor networks are a way to describe multipartite quantum states

$$|\Psi\rangle = \sum_{i_1,\dots,i_N} c_{i_1,\dots,i_N} |i_1,\dots,i_N\rangle \tag{15}$$

by rewriting $c_{i_1,\dots,i_N} = \sum_{\alpha_k} \prod C_{\{i_s,\alpha_p\}}$ using tensors $C_{\{i_s\},\{\alpha_p\}}$ which depend on only on a few (if any) physical indices $\{i_s\}$ and auxiliary indices (or *entanglement degrees of freedom*) $\{\alpha_p\}$ each, where each auxiliary index is contained in

exactly two tensors, and where the auxiliary indices are contracted (i.e., summed over). In the context of fermionic tensor networks, it is convenient to instead consider each $C_{\{i_s\},\{\alpha_p\}}$ as describing a pure state ("fiducial state") $|C_{\{i_s\},\{\alpha_p\}}\rangle = \sum C_{\{i_s\},\{\alpha_p\}} |\{i_s\}, \{\alpha_p\}\rangle$ of the physical *and* virtual indices jointly, where the contraction is achieved by projecting pairs of virtual indices onto a maximally entangled state $|\omega_{\alpha,\alpha'}\rangle$:

$$|\Psi\rangle = \left(\bigotimes \langle\omega_{\alpha,\alpha'}| \right) \left(\bigotimes |C_{\{i_s\},\{\alpha_p\}}\rangle \right). \tag{16}$$

This framework is particularly suited for the case of fermionic tensor network states, as it removes the requirement to specify an ordering of the fermionic tensors (and bonds), since both $\langle\omega_{\alpha,\alpha'}|$ and $|C_{\{i_s\},\{\alpha_p\}}\rangle$ have fixed fermion parity, and all tensors C mutually commute (as they don't share fermionic operators), and similarly all bonds ω.

b. Contraction. In the case of Gaussian fermionic tensor networks, the fiducial states are given by covariance matrices γ which are indexed by the corresponding physical and auxiliary modes. The maximally entangled state $|\omega\rangle$ (which can either be a product of maximally entangled states between complex fermions or a product of pairs of Majorana fermions in their joint vacuum) is itself Gaussian, such that each contraction gives rise to a new Gaussian state. The basic step is thus the contraction of two Gaussian tensors (i.e., states). This corresponds to integrating out the auxiliary degrees of freedom in a Gaussian integral and thus gives rise to a Schur complement [47,50–52].

This leads to the following explicit expressions for contraction: Given two tensors

$$G = \begin{pmatrix} G_{aa} & G_{ac} \\ -G_{ac}^{\mathsf{T}} & G_{cc} \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} H_{bb} & H_{bc'} \\ -H_{bc'}^{\mathsf{T}} & H_{c'c'} \end{pmatrix} \tag{17}$$

(possibly with blocked indices), the tensor resulting from contracting the c index of G with the c' index of H is given by

$$K = \begin{pmatrix} K_{aa} & K_{ab} \\ -K_{ab}^{\mathsf{T}} & K_{bb} \end{pmatrix} = \begin{pmatrix} G_{aa} \\ & H_{bb} \end{pmatrix}$$
$$+ \begin{pmatrix} G_{ac} \\ & H_{bc'} \end{pmatrix} \begin{pmatrix} G_{cc} & \mathbb{1} \\ -\mathbb{1} & H_{c'c'} \end{pmatrix}^{-1} \begin{pmatrix} G_{ac} \\ & H_{bc'} \end{pmatrix}^{\mathsf{T}}. \tag{18}$$

Note that the contraction is oriented (swapping c and c' changes the sign of the $\mathbb{1}$'s); we will from now on use the notation $c \triangleright c'$ to indicate the order. The matrix inverse in Eq. (18) can be carried out blockwise using Schur complements. Explicitly,

$$\begin{pmatrix} A & \mathbb{1} \\ -\mathbb{1} & D \end{pmatrix}^{-1} = \begin{pmatrix} (DA + \mathbb{1})^{-1} \\ & (AD + \mathbb{1})^{-1} \end{pmatrix} \begin{pmatrix} D & -\mathbb{1} \\ \mathbb{1} & A \end{pmatrix}, \tag{19}$$

and as a special case which we will use later,

$$\begin{pmatrix} A & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & A \end{pmatrix}. \tag{20}$$

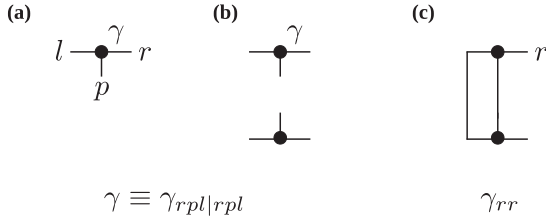For generalizations of this approach as well as the normalization of the state after contraction, see Appendix B.

(a) (b) (c)

$$\gamma \equiv \gamma_{rpl|rpl} \qquad\qquad \gamma_{rr}$$

FIG. 1. Graphical calculus for Gaussian fermionic tensor networks (see Sec. II Bc). (a) Tensors are described by their covariance matrix γ, which has blocks of rows/columns labeled by the different indices; the size of each block equals the number of Majorana modes in the block. (b) Since γ describes pure and mixed state on equal footing, we will sometimes draw even pure state tensors with both ket and bra indices. (c) This allows, e.g., to naturally capture tracing of an index in the graphical calculus.

c. Graphical calculus. Tensor networks can be conveniently expressed using a graphical calculus where each tensor is described by a box with one leg per index (where each index can contain any number of Majorana modes), Fig. 1(a). Contraction of indices is indicated by connecting the corresponding legs; the orientation of the contraction will be indicated by an arrow. This is shown in Fig. 2.

Note that the covariance matrix formalism (and thus Gaussian tensor networks) does not distinguish between pure and mixed states. Yet, we will generally use the tensor notation for *pure* states; when referring to the density operator, we will instead draw the tensor twice, once for the "ket" and once for the "bra" system [Fig. 1(b)]. In particular, an essential operation will be tracing out part of a tensor, which corresponds to contracting the "ket" and "bra" part of the corresponding index, and which will be drawn as in Fig. 1(c).

C. Entanglement in Gaussian states and Schmidt decomposition

A crucial step in the DMRG algorithm is the Schmidt decomposition, i.e., the decomposition of a bipartite entangled state $|\Psi\rangle$ in the form $|\Psi\rangle = \sum_k \lambda_k |\ell_k\rangle |r_k\rangle$ with orthonormal vectors $\{|\ell_k\rangle\}$ and $\{|r_k\rangle\}$ (see also Sec. II D 1).
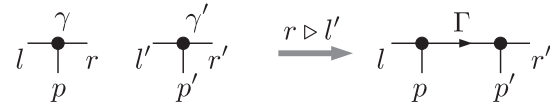
$$\Gamma_{lpp'r'|lpp'r'} = A + BD^{-1}B^T$$

$$A = \begin{pmatrix} \gamma_{lp|lp} & 0 \\ 0 & \gamma'_{p'r'|p'r'} \end{pmatrix}$$

$$B = \begin{pmatrix} \gamma_{lp|r} & 0 \\ 0 & \gamma'_{p'r'|l'} \end{pmatrix} \qquad D = \begin{pmatrix} \gamma_{rr} & \mathbb{1}_{rl'} \\ -\mathbb{1}_{l'r} & \gamma'_{l'l'} \end{pmatrix}$$

FIG. 2. Contraction of two indices r and l' (contraction is oriented, denoted as $r \rhd l'$, and indicated by the arrow on the contracted index). The CM Γ of the resulting state is obtained as a Schur complement, as given by the formulas in the figure or Eq. (18). (Here $\mathbb{1}_{xy}$ denotes the identity matrix between modes x and y.)

The procedure for fermionic Gaussian states is quite analogous. Consider a pure bipartite Gaussian state

$$\gamma = \begin{pmatrix} \gamma_{aa} & \gamma_{ab} \\ -\gamma_{ab}^\mathsf{T} & \gamma_{bb} \end{pmatrix} \tag{21}$$

and assume for the moment that γ_{ab} is square (i.e., the two systems have the same dimension) and has full rank; this corresponds to the case where all modes on either side are entangled. Since γ is pure, we have $\gamma^2 = -\mathbb{1}$, which implies

$$\gamma_{aa}^2 = \gamma_{ab}\gamma_{ab}^\mathsf{T} - \mathbb{1}, \qquad \gamma_{bb}^2 = \gamma_{ab}^\mathsf{T}\gamma_{ab} - \mathbb{1}, \tag{22}$$

as well as $\gamma_{aa}\gamma_{ab} = -\gamma_{ab}\gamma_{bb}$, or equivalently,

$$\gamma_{bb} = -\gamma_{ab}^{-1}\gamma_{aa}\gamma_{ab}. \tag{23}$$

We now perform a singular value decomposition (SVD) of γ_{ab}, $O\gamma_{ab}Q^\mathsf{T} = \Lambda_{ab}$, where Λ_{ab} is diagonal with positive entries which are ordered descendingly, and O and Q are real orthogonal, as they diagonalize the real symmetric matrices $\gamma_{ab}\gamma_{ab}^\mathsf{T}$ and $\gamma_{ab}^\mathsf{T}\gamma_{ab}$, respectively. Equation (22) then immediately implies that also $O\gamma_{aa}^2 O^\mathsf{T}$ is diagonal with descending entries. Following Eq. (6), we find that

$$O\gamma_{aa}O^\mathsf{T} = \bigoplus_k \begin{pmatrix} & \lambda_k \\ -\lambda_k & \end{pmatrix} \tag{24}$$

(possibly after rearranging rows of O within degenerate blocks; by performing the same permutation on Q, we can ensure that $O\gamma_{ab}Q^\mathsf{T}$ remains diagonal). Therefore,

$$Q\gamma_{bb}Q^\mathsf{T} = Q\left(-\gamma_{ab}^{-1}\gamma_{aa}\gamma_{ab}\right)Q^\mathsf{T} \tag{25}$$

$$= -(O\gamma_{ab}Q^\mathsf{T})^{-1}(O\gamma_{aa}O^\mathsf{T})(O\gamma_{ab}Q^\mathsf{T}) \tag{26}$$

must be again of the form

$$Q\gamma_{bb}Q^\mathsf{T} = \bigoplus_k \begin{pmatrix} & -\lambda_k \\ \lambda_k & \end{pmatrix}, \tag{27}$$

where the fact that the entries λ_k equal those in Eq. (24) follow from (22), as well as that

$$O\gamma_{ab}Q^\mathsf{T} = \bigoplus_k \begin{pmatrix} \mu_k & \\ & \mu_k \end{pmatrix} \tag{28}$$

with $\mu_k = \sqrt{1 - \lambda_k^2}$. Together, we find that the local rotations $O \oplus Q$ transform γ to

$$(O \oplus Q)\gamma(O \oplus Q)^\mathsf{T} = \bigoplus_k W(\lambda_k), \tag{29}$$

with

$$W(\lambda_k) = \begin{pmatrix} & \lambda_k & \mu_k & \\ -\lambda_k & & & \mu_k \\ -\mu_k & & & -\lambda_k \\ & -\mu_k & \lambda_k & \end{pmatrix}, \qquad \mu_k = \sqrt{1 - \lambda_k^2}, \tag{30}$$

where each block $W(\lambda_k)$ captures the correlations between two a and two b Majorana modes (in this order). This is the fermionic Gaussian equivalent of the Schmidt decomposition (and of the Williamson normal form for bosonic Gaussian states), a result first derived by Botero and Reznik [45].

From this form, the von Neumann entanglement entropy between the modes a and b is now easily found to be

$$S_{\text{vN}} = \sum_k H\left(\frac{1-\lambda_k}{2}\right), \qquad (31)$$

with $H(p) = -p \log p - (1-p)\log(1-p)$.

In the more general case where γ_{ab} does not have full rank (or is not square), we can still perform an SVD of γ_{ab}, which in that case will give rise to blocks where $\gamma_{ab}\gamma_{ab}^{\mathsf{T}}$ (or $\gamma_{ab}^{\mathsf{T}}\gamma_{ab}$) is zero. Equation (22) implies that the corresponding blocks of γ_{aa} (γ_{bb}) decouple and are equal to $\bigoplus(_1 \quad ^{-1})$, i.e., they describe a mode in a pure (vacuum) state, which is thus not entangled to any other mode. In the general case, we thus obtain the form

$$(O \oplus Q)\gamma(O \oplus Q)^{\mathsf{T}} = \left[\bigoplus\begin{pmatrix}_1 & ^{-1}\end{pmatrix}\right] \oplus \left[\bigoplus_k W(\lambda_k)\right]$$

$$\times \oplus \left[\bigoplus\begin{pmatrix}_1 & ^{-1}\end{pmatrix}\right], \qquad (32)$$

where the first (last) term in the sum lives only on the a (b) modes, and the middle term captures their correlations; as before, it is obtained from the SVD of γ_{ab}, together with a suitable rearrangement of degenerate modes.

Note that instead of performing the SVD of γ_{ab}, we could just as well have diagonalized γ_{aa} and γ_{bb}. The approach chosen here has two advantages: First, it does not require the eigenvalues to be matched, and second, the singular values of γ_{ab} relate to the square roots of the eigenvalues of γ_{aa}, thus yielding a higher accuracy in finite-precision arithmetic (also in the basis transformations O and Q) for weakly entangled modes.

D. Tensor decompositions

Tensor decompositions, such as the SVD and the closely related QR decomposition, are core ingredients in DMRG and other tensor network algorithms. Their main role is that they allow the state to decompose in terms of an orthonormal basis; furthermore, they are an efficient way of obtaining the Schmidt decomposition and thus information about the entanglement, including the full entanglement spectrum. We will start by introducing the Gaussian version of the SVD and then discuss how to build different related decompositions which are potentially easier or more efficient to implement.

1. Singular value decomposition

The SVD decomposes a matrix $M = UDV^\dagger$, where U and V are isometries ($U^\dagger U = \mathbb{1}$, $V^\dagger V = \mathbb{1}$) and D is diagonal and full rank. It is closely related to the Schmidt decomposition, which for a state $|\Psi\rangle = \sum_{ij} M_{ij}|i\rangle|j\rangle$ is obtained from the SVD of M as

$$|\Psi\rangle = \sum d_k |a_k\rangle |b_k\rangle \qquad (33)$$

$$d_k = D_{kk}, \quad |a_k\rangle = \sum U_{ik}|i\rangle, \quad |b_k\rangle = \sum V_{jk}^*|j\rangle. \qquad (34)$$

This can be viewed as an elementary tensor network: the state $|\Psi\rangle$ is given as the contraction of tensors U, D, and V^*, where
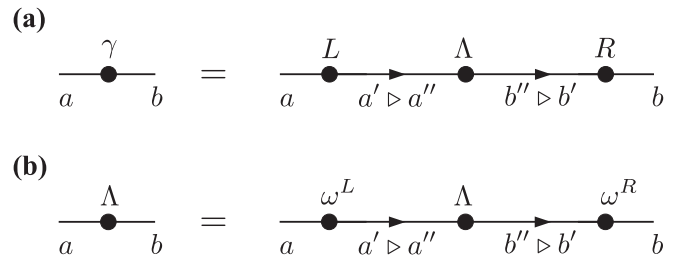

FIG. 3. Gaussian fermionic SVD. (a) The SVD expresses a CM γ as the contraction of three CMs L, Λ, and R, where L and R are isometries when contracting a and b, respectively (i.e., $L_{a'a'} = 0$, $R_{b'b'} = 0$), and $\Lambda = \bigoplus W(\lambda_k)$, Eq. (30). (b) The derivation proceeds by using the Schmidt decomposition of $\gamma = (O \oplus Q)^{\mathsf{T}}\Lambda(O \oplus Q)$, rewriting Λ as the contraction of itself with two maximally entangled states ω^L and ω^R, and applying O and Q to a and b, respectively.

U and V^* have one auxiliary and one physical index each, and D has only auxiliary indices.

We have discussed the Gaussian version of the Schmidt decomposition in the previous section, Sec. II C. Building on this, we will now describe how to view the Schmidt decomposition as a tensor decomposition, i.e., how to view the state as the contraction of a minimal tensor network of three tensors. Specifically, the Gaussian SVD corresponds to the decomposition of the covariance matrix γ of a bipartite system ab into the contraction of three covariance matrices L, Λ, and R, as indicated in Fig. 3(a) (including the labeling of the modes). Here, the isometry condition $U^\dagger U = \mathbb{1}$ translates naturally to $L_{a'a'} = 0$, which is the CM of the maximally mixed state $\mathbb{1}$, cf. Fig. 1(c), and correspondingly, $R_{b'b'} = 0$; furthermore, we require $\Lambda = \bigoplus W(\lambda_k)$ [cf. Eq. (29)].

As a starting point, consider the covariance matrix γ written as

$$\gamma = (O \oplus Q)^{\mathsf{T}}\Lambda(O \oplus Q) \qquad (35)$$

with orthogonal O and Q; note that at this point, we do not need to make any assumption about Λ, i.e., the decomposition can be more general than Eq. (29). We will generalize it to the case with additional local degrees of freedom as in Eq. (32) in a moment. The first step is to rewrite Λ as a contraction of three tensors ω^L, Λ, and ω^R, as depicted in Fig. 3(b) (note the relabeling of the indices of $\Lambda \equiv \Lambda_{a''b''|a''b''}$)—this is the equivalent of contracting with the identity tensor, i.e., writing $D = \mathbb{1} \cdot D \cdot \mathbb{1}$. The tensors $\omega^{L,R}$ are given by

$$\omega^L \equiv \omega^L_{aa'|aa'} = \begin{pmatrix} 0 & \mathbb{1}_{aa'} \\ -\mathbb{1}_{a'a} & 0 \end{pmatrix} \qquad (36)$$

and

$$\omega^R \equiv \hat{\omega}_{b'b|b'b} = \begin{pmatrix} 0 & \mathbb{1}_{b'b} \\ -\mathbb{1}_{bb'} & 0 \end{pmatrix}. \qquad (37)$$

They describe maximally entangled states between a and a' and b' and b, respectively (corresponding to the identity tensor $\mathbb{1}$). It is easy to check using Eq. (18) that contracting $a' \triangleright a''$ returns Λ, and likewise, for $b'' \triangleright b'$ [53].

In a second step, we now take Fig. 3(b) and apply O and Q to the a and b system, respectively. On the left-hand side, this yields $(O \oplus Q)\Lambda(O \oplus Q)^{\mathsf{T}} = \gamma$, while on the right-hand side,

ω^L and ω^R are transformed to

$$L = (O^\mathsf{T} \oplus \mathbb{1})\omega^L(O \oplus \mathbb{1}) = \begin{pmatrix} 0 & O^\mathsf{T} \\ -O & 0 \end{pmatrix} \quad (38)$$

and

$$R = (\mathbb{1} \oplus Q^\mathsf{T})\omega^R(\mathbb{1} \oplus Q) = \begin{pmatrix} 0 & Q \\ -Q^\mathsf{T} & 0 \end{pmatrix}, \quad (39)$$

respectively. Furthermore, the contraction and the application of O and Q commute, as they act on different indices (and as can be also checked explicitly). We thus find that γ can be written as the contraction of L, Λ, and R, as in Fig. 3(a), with L and R isometries. In the case where Λ is the Schmidt decomposition, this yields the Gaussian analog of the SVD.

In the general case where the Schmidt decomposition contains additional unentangled local degrees of freedom, as in Eq. (32), these modes need to be attached to the a and b system of ω^L and ω^R, respectively, before applying the transformations in Eqs. (38) and (39); the resulting L and R are still isometries, and the decomposition in Fig. 3(a) still holds with $\Lambda = \bigoplus W(\lambda_k)$. Note that we can always regard a pair of decoupled modes as $W(\lambda_k \equiv 1)$ if we do not want to minimize the bond dimension (i.e., the dimension of Λ).

2. Truncation of bond dimension

An important scenario, in particular in two-site DMRG, is to truncate the bond dimension. This corresponds to carrying out an SVD $M = UDV^\dagger$ of a matrix M and replacing it by a matrix $M' = (UW)(W^\dagger DV^\dagger)$ with an isometry W such that $WW^\dagger$ projects onto the χ largest singular values of D.

In the Gaussian scenario, given a CM γ this amounts to carrying out the Schmidt decomposition Eq. (32) and keeping only the blocks $W(\lambda_k)$ with the κ smallest λ_k, while replacing the additional blocks with

$$W(\lambda = 1) \equiv \begin{pmatrix} & 1 & & \\ -1 & & & \\ & & & -1 \\ & & 1 & \end{pmatrix}, \quad (40)$$

i.e., decoupled modes. Those modes are henceforth counted towards the decoupled modes in Eq. (32), and the truncated SVD is obtained by applying the SVD construction, Fig. 3(a), with Λ from the truncated Schmidt decomposition.

3. Variants: QR and related decompositions

In most applications, a full SVD is not needed. For example, to obtain the canonical form of an MPS (see below for details) one only needs to be able to rewrite a covariance matrix $\gamma \equiv \gamma_{ab|ab}$, with $|a| \geqslant |b|$, as the contraction of an isometry and a tensor with equal-sized systems on both sides. (In DMRG, a corresponds to the left virtual + physical modes while b corresponds to the right virtual modes.) To this end, we need to isolate modes in a that are not entangled with any modes in b; we will denote the modes that are unentangled as a_0 and the others as a_1, i.e., $|a| = |a_0| + |a_1|$ and $|a_1| = |b|$. (We assume that all modes in b are entangled with modes in a, as happens in practice; the generalization is straightforward.) To achieve this, we find an $|a| \times |a|$ orthogonal transformation

O such that

$$(O \oplus \mathbb{1})\gamma(O \oplus \mathbb{1})^\mathsf{T} = \left(\begin{array}{c|cc} \gamma_{a_0 a_0} & 0 & 0 \\ \hline 0 & \gamma_{a_1 a_1} & X \\ 0 & -X^\mathsf{T} & \gamma_{bb} \end{array} \right) \quad (41)$$

where $\gamma_{a_0 a_0} = \bigoplus \left(\begin{smallmatrix} & 1 \\ -1 & \end{smallmatrix} \right)$, and X is $|b| \times |b|$. Note that this is a weaker version of the decomposition of Eq. (29), where no transformation is applied on the b modes. To obtain this form, it is sufficient to choose O such that

$$O\gamma_{ab} = \begin{pmatrix} 0 \\ X \end{pmatrix}, \quad (42)$$

with X a square matrix; this can be achieved, for example, by determining the kernel of γ_{ab}^T (or the image of γ_{ab}). The fact that this implies $\gamma_{a_0 a_1} = 0$ can be verified by invoking purity of the state. Once such an O is found, following the SVD construction of Fig. 3(a) with $Q = \mathbb{1}$ and $\Lambda \equiv \begin{pmatrix} \gamma_{a_1 a_1} & X \\ -X^\mathsf{T} & \gamma_{bb} \end{pmatrix}$ gives the desired result. For the case where $|a| \leqslant |b|$, the equivalent transformation is obtained by acting with O on the b modes. Note that for $|a| = |b|$, O can be chosen trivial.

Similarly, in order to obtain a decomposition which allows the bond dimension to be truncated (such as in two-site DMRG), it is sufficient to determine O such that γ_{aa} is diagonalized; this allows one to determine which modes decouple and yields the λ_k of $W(\lambda_k)$, which can be used to identify and disentangle the least entangled modes, without the need to determine Q.

III. GAUSSIAN FERMIONIC MPS

A. Construction

For the remainder of this manuscript we will focus on the Gaussian fermionic version of matrix product states. A matrix product state is a particular type of tensor network state in one dimension, where the sites of the physical system are arranged on a chain and one tensor is associated with each physical site. The tensor is connected to its left and right neighbors. In the following discussion, we will focus on the case of open boundary conditions. For a chain of N sites, an MPS consists of N rank-3 tensors. In the case of a Gaussian fermionic MPS (GFMPS), it is thus fully specified by a set of covariance matrices

$$\gamma^s_{l_s p_s r_s | l_s p_s r_s}, \quad (43)$$

where $s = 1, \ldots, N$ denotes the sites of the chain (from left to right), p_s corresponds to the physical modes at each site, and the dimensions $|r_s| = |l_{s+1}|$ of the virtual modes r_s, l_{s+1} are the *Majorana bond mode number* of the bond $(s, s+1)$. As in conventional MPS, the bond mode number, analogous to the bond dimension, can be increased to systematically increase the class of variational states, and an exact description is recovered for $|l_s| \sim \mathcal{O}(N)$. As noted in the Introduction, the bond dimension M in a conventional MPS is related to the Majorana bond number χ used here by $M = \sqrt{2}^\chi$. For open boundary conditions, there are no left modes on the leftmost tensor and no right modes on the rightmost tensor, i.e., $|l_1| = 0$ and $|r_N| = 0$. It is therefore often convenient to omit these and set $\gamma^1 = \gamma^1_{p_1 r_1 | p_1 r_1}$ and $\gamma^N = \gamma^N_{l_N p_N | l_N p_N}$. When clear from the
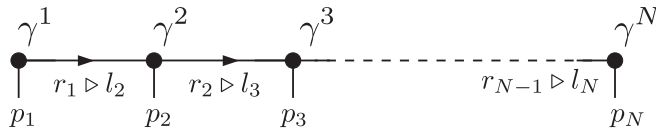
FIG. 4. Construction of a Gaussian fermionic MPS. To each site we associate a tensor $\gamma^s_{l_s p_s r_s | l_s p_s r_s}$ (with l_1 and r_N trivial); the state of the "physical modes" $p_1, \ldots, p_N$ is obtained by contracting the "virtual modes" $r_s \triangleright l_{s+1}$.

context, we will often omit the site subscripts to the modes and write, for example, $\gamma^s_{lpr|lpr}$.

A description of the state on N sites is now obtained by arranging the γ^s on a line and contracting the adjacent virtual indices $r_s \triangleright l_{s+1}$, as depicted in Fig. 4, which yields a CM describing the physical modes $p_1, \ldots, p_N$. Note that on physical grounds, it is clear that the state obtained through the above construction is independent of the order of the contractions, even though this is not evident from the description of the contraction in the CM formalism through Schur complements.

Owing to the properties of the Gaussian formalism, a GFMPS always describes a normalized state up to an overall global phase. This fact has the potential to complicate practical MPS algorithms somewhat: in the conventional formalism, physical observables such as the energy are generally bilinear in the individual tensor elements. Thus, optimization at least over individual tensors manifestly is a quadratic and thus well-behaved problem. Since GFMPS are a special case of MPS, this must, in principle, also be true for GFMPS, but it is not apparent in the Gaussian formulation. However, with some extra steps the usual structure can be exposed and the conventional and well-tested algorithms, such as single- and two-site DMRG optimization as well as related time-evolution algorithms, can be formulated. A central role is played by the canonical form of the MPS. While this form is in practice very useful for the stability and performance of conventional MPS algorithms, it is not strictly required; in the Gaussian case, on the other hand, we will find it to be essential. In the following, we will first introduce the canonical form of GFMPS and in the following sections will proceed to discuss some important algorithms, such as efficient computation of the total energy, single- and two-site optimization, and finally, the time-dependent variational principle.

B. Canonical form

a. Definition. The first step in defining the canonical form of an MPS is to choose a center "working" site s_0 with respect to which the canonical form is defined. Then, all tensors to the left of s_0 are brought into left-canonical form and the ones to its right into right-canonical form. The definitions of these are given by $\gamma^s_{rr} = 0$ (left-canonical form, $s < s_0$) and $\gamma^s_{ll} = 0$ (right-canonical form, $s > s_0$), as illustrated in Fig. 5. This definition precisely corresponds to that in the conventional MPS formalism. There, an MPS tensor is said to be left-canonical if the contraction with its own adjoint over the left and physical index yields the identity matrix [see tensor diagram of Fig. 5(a)], or equivalently, the tensor—seen as a
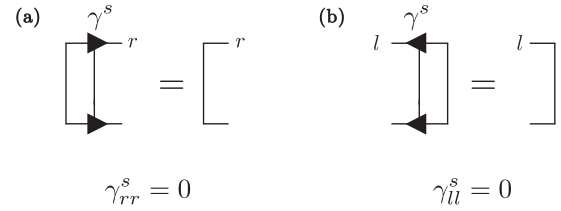

FIG. 5. (a) Left- and (b) right-canonical form of a tensor γ^s and the corresponding condition on γ^s.

map from right to physical and left index—is an isometry. The contraction corresponds to tracing out the left and physical system, and thus for GFMPS, the left-canonical form amounts to requiring that the reduced density matrix of the r modes be $\gamma^s_{rr} = 0$. Correspondingly, for the right-canonical one has $\gamma^s_{ll} = 0$. Note that this is precisely the isometry condition we used in the derivation of the Gaussian SVD.

b. Elementary move. The elementary move in establishing and keeping the canonical form is the following: Given a site s such that all tensors to the left of it are in left-canonical form, transform γ^s into left-canonical form by changing γ^s and γ^{s+1}, without changing the state described by the GFMPS. We focus on the left-canonical form, but the corresponding procedure for the right-canonical form is completely analogous.

We proceed in two steps, as shown in Fig. 6: In step (I), γ^s is rewritten with two tensors, namely, a new tensor $\hat{\gamma}^s$ in left-canonical form, contracted with a new tensor Λ on its right virtual bond. In step (II), Λ is contracted with γ^{s+1}, giving a new tensor $\hat{\gamma}^{s+1}$.

Step (I) is implemented by blocking $(lp) \equiv a$ and performing the SVD of $\gamma^s_{ar|ar}$ as described in Sec. II D 1, or one of the simplified versions in Sec. II D 3, which yields a decomposition of γ^s into an isometry $L \equiv \hat{\gamma}^s$ and Λ, as desired. (If a decomposition with nontrivial R [cf. Fig. 3(a)], such as the SVD, is used, Λ and R need to be contracted.) Specifically, it is sufficient to choose an O such that
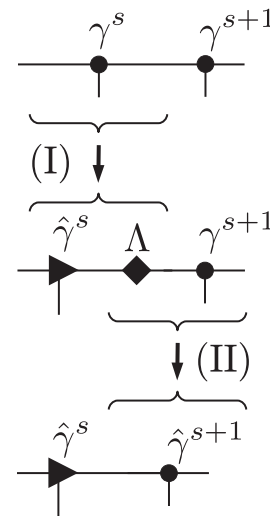

FIG. 6. Bringing a tensor into left-canonical form. In step (I), γ^s is split into a left-canonical tensor $\hat{\gamma}^s$ and a tensor Λ on the bond. In step (II), Λ is absorbed in the tensor γ^{s+1} to the right, yielding $\hat{\gamma}^{s+1}$.

$O\gamma_{ar}^s = \binom{0}{Y}$ with Y square, define

$$
\begin{pmatrix} X & 0 & 0 \\ \hline 0 & & \\ 0 & & \Lambda \end{pmatrix} \equiv (O \oplus \mathbb{1})\gamma^s(O \oplus \mathbb{1})^{\mathsf{T}} \qquad (44)
$$

(on the left-hand side, the first two blocks correspond to a and the third block to p), and let

$$
\hat{\gamma}^s \equiv (O \oplus \mathbb{1})^{\mathsf{T}} \begin{pmatrix} X & 0 & 0 \\ \hline 0 & 0 & \mathbb{1} \\ 0 & -\mathbb{1} & 0 \end{pmatrix}(O \oplus \mathbb{1}). \qquad (45)
$$

For the sizes of X and Λ and more details on the procedure, see Sec. II D 3. In the second step, Λ and γ^{s+1} are contracted as indicated in Fig. 6, using Eq. (18).

In order to obtain the right-canonical form, the same steps have to be followed; the only core difference is the inverse contraction order, which in particular implies that the signs of the $\pm\mathbb{1}$ in the SVD and the contraction have to be reversed.

c. Obtaining and updating the canonical form. Given this elementary step, the GFMPS can be kept in canonical form at all times using the same procedures that are used for conventional MPS and described in detail in, e.g., Ref. [5]. In particular, given an initial GFMPS on N sites that is not in canonical form, it can be brought into canonical form by sweeping from one end to the other, i.e., in N elementary operations. All practical optimization or time evolution algorithms for MPS perform local operations in each step and sweep over the system from one end to the other (see also Sec. V). Therefore, the state can be kept in canonical form by moving the center s, i.e., the site such that all tensors left (right) of it are left-canonical (right-canonical), by only a single site, requiring only local operations.

IV. ENERGY COMPUTATION FOR GFMPS

We now explain how to compute and optimize the energy of a GFMPS for a given Hamiltonian. We start with a simple elementary step which takes a two-site Hamiltonian and a two-site GFMPS in canonical form and expresses the energy of the GFMPS as a linear function of one of the GFMPS tensors through an effective single-site Hamiltonian. This will allow us to derive an iterative formula for the energy which is valid for any Hamiltonian, and which can moreover be used to optimize the energy of a given working site. Subsequently, we show how this formula can be specialized to different cases, such as local Hamiltonians, to obtain more efficient ways to compute the energy.

For sake of completeness, we recall from the Introduction [Eqs. (1) and (5)] that a general Gaussian Hamiltonian is of the form

$$
\mathcal{H} = -i \sum H_{ij} c_i c_j + E, \qquad (46)
$$

(with H an antisymmetric matrix and E a constant offset), and the energy of a state ρ with CM γ is given by

$$
\mathrm{tr}[\mathcal{H}\rho] = \mathrm{tr}(H\gamma) + E. \qquad (47)
$$

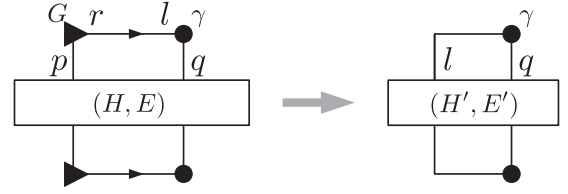We will henceforth denote this Hamiltonian by (H, E).

FIG. 7. Elementary step for evaluating energies in MPS. Dependence of the energy on one tensor γ (see text).

A. Energy in GFMPS: Elementary step

Let us first consider the scenario depicted in Fig. 7. We are given two tensors $G_{pr|pr}$ and $\gamma_{lq|lq}$, where $G_{rr} = 0$, i.e., it is in left-canonical form, and $r \triangleright l$ is contracted, and consider a Hamiltonian (H, E) acting on pq. Following Eq. (18) for the contraction, using Eq. (20) for the inverse, and rearranging some terms, we find that the contraction yields a CM:

$$
\Theta = \begin{pmatrix} G_{pp} & 0 \\ 0 & 0+ \end{pmatrix}\begin{pmatrix} G_{pr} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} \gamma_{ll} & -\gamma_{ql}^{\mathsf{T}} \\ \gamma_{ql} & \gamma_{qq} \end{pmatrix}\begin{pmatrix} G_{pr}^{\mathsf{T}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \qquad (48)
$$

We can thus rewrite

$$
\mathrm{tr}[H\Theta]
$$
$$
= \mathrm{tr}\left[H\begin{pmatrix} G_{pp} & 0 \\ 0 & 0 \end{pmatrix}\right] + \mathrm{tr}\left[H\begin{pmatrix} G_{pr} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\gamma\begin{pmatrix} G_{pr}^{\mathsf{T}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\right] \qquad (49)
$$
$$
= \mathrm{tr}[H_{pp}G_{pp}] + \mathrm{tr}[H'\gamma], \qquad (50)
$$

with

$$
H' = \begin{pmatrix} G_{pr}^{\mathsf{T}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}H\begin{pmatrix} G_{pr} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \qquad (51)
$$

where the size of the identities is $|q| \times |q|$, and, of course, $|r| = |l|$, and therefore H' has dimensions $(|l| + |q|) \times (|l| + |q|)$. The effective Hamiltonian acting on the γ system—comprised of the modes in l and q—is thus obtained from the old Hamiltonian through

$$
(H, E) \mapsto (H', E' = E + \mathrm{tr}[H_{pp}G_{pp}]), \qquad (52)
$$

with H' from Eq (51).

Note that if the right tensor were canonical and we instead contracted $l \triangleright r$, the signs of the $\mathbb{1}$'s in Eq. (51) had to be reversed.

B. Energy in GFMPS: Iteration formula

We are now ready to derive how the energy of a given GFMPS for some Hamiltonian $(H^{1,N}, E^{1,N})$, acting on $p_1, \ldots, p_N$, depends on the GFMPS tensor at a specific site $\hat{s}$. It is based on iterated application of the elementary step above and is illustrated in Fig. 8.

Consider an MPS in canonical form around $\hat{s}$, with H acting on all physical sites $p_1, \ldots, p_N$. For the description, we assume the generic case $1 < \hat{s} < N$, but the procedure is easily applied to any $\hat{s}$. By applying the procedure derived in the last section to γ^1 (corresponding to G above) vs the remaining tensors (seen as one blocked tensor γ), we obtain an effective
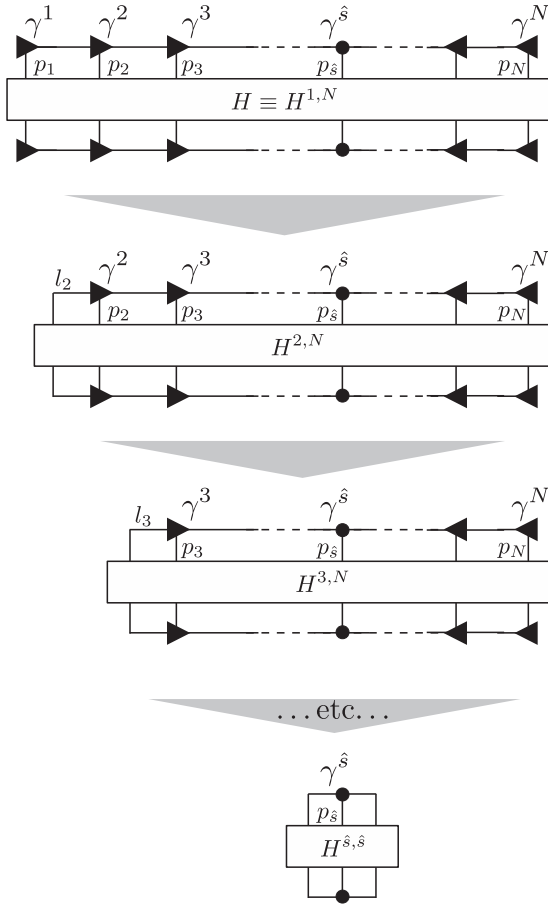
FIG. 8. Scheme for the contraction of the Hamiltonian. We proceed stepwise from both ends (here shown from the left end), which gives an effective Hamiltonian for successively smaller patches which also acts on the dangling virtual bonds at the boundary, finally yielding an effective Hamiltonian for $\gamma^{\hat{s}}$.

Hamiltonian

$$H^{2,N} \overset{l_2 \leftarrow r_1}{=} \left(\gamma^1_{p_1 r_1} \oplus \mathbb{1}\right)^\mathsf{T} H^{1,N}\left(\gamma^1_{p_1 r_1} \oplus \mathbb{1}\right), \qquad (53)$$

where $H^{2,N}$ is acting on $l_2, p_2, \dots, p_N$, and we relabel r_1 to l_2, as indicated on top of the equality sign. We can repeat this procedure,

$$H^{s+1,N} \overset{l_{s+1} \leftarrow r_s}{=} \left(\gamma^s_{l_s p_s | r_s} \oplus \mathbb{1}\right)^\mathsf{T} H^{s,N}\left(\gamma^s_{l_s p_s | r_s} \oplus \mathbb{1}\right), \qquad (54)$$

until we reach site $\hat{s}$. Similarly, we can contract H from the right, using

$$H^{\hat{s},s-1} \overset{r_{s-1} \leftarrow l_s}{=} \left[\gamma^s_{p_s r_s | l_s} \oplus (-\mathbb{1})\right]^\mathsf{T} H^{\hat{s},s}\left[\gamma^s_{p_s r_s | l_s} \oplus (-\mathbb{1})\right], \quad (55)$$

note the $-\mathbb{1}$ due to the opposite contraction order, until we reach $\hat{s}$.

In words, each step removes one site from the MPS and instead updates the way in which the Hamiltonian acts on the dangling virtual leg at the boundary (this update rule thus involves the correlations between this dangling leg and the physical leg in the bulk). This is illustrated for the left boundary in Fig. 8. In order to update H, we have to take the Hamiltonian part corresponding to the left+physical leg which is being removed (site s) and transform it with $\gamma^s_{l p | r}$,

which relates those legs to the new dangling leg (and correspondingly, from the right).

In order to compute the total energy, we must also keep track of the E contribution. It is updated as

$$E^{s+1,N} = \mathrm{tr}\big[H^{s,N}_{l_s p_s | l_s p_s}\gamma^s_{l_s p_s | l_s p_s}\big] + E^{s,N},$$
$$E^{\hat{s},s-1} = \mathrm{tr}\big[H^{\hat{s},s}_{p_s r_s | p_s r_s}\gamma^s_{p_s r_s | p_s r_s}\big] + E^{\hat{s},s}, \qquad (56)$$

with $E^{2,N} = \mathrm{tr}[H^{1,N}_{p_1 p_1}\gamma^1_{p_1 p_1}] = \mathrm{tr}[H_{p_1 p_1}\gamma^1_{p_1 p_1}]$, and accordingly, $E^{\hat{s},N-1} = \mathrm{tr}[H_{p_N p_N}\gamma^N_{p_N p_N}]$.

We are thus left with a single-site Hamiltonian $H^{\hat{s},\hat{s}}$ acting on $l_{\hat{s}} p_{\hat{s}} r_{\hat{s}}$ such that the total energy is

$$E_{\mathrm{tot}}(\gamma^{\hat{s}}) = \mathrm{tr}[H^{\hat{s},\hat{s}}\gamma^{\hat{s}}] + E^{\hat{s},\hat{s}}. \qquad (57)$$

C. Local Hamiltonians

Let us now consider the special case of local Hamiltonians—for simplicity, we consider nearest-neighbor Hamiltonians—and explain how the evaluation of the Hamiltonian in this case can be accomplished in linear time in the system size. Moreover, just as in conventional DMRG, it is possible to store intermediate terms at each cut such that the effort required to update this information when moving by one site is independent of system size.

In the following, it will be convenient to also carry an l_1 and r_N label, with the corresponding spaces being zero-dimensional (in the CM). In any step of the computation, the Hamiltonian will be of the banded form:

$$H^{s,N} = \begin{array}{c} \\ \end{array} \begin{pmatrix} \begin{array}{c} l_s \\ A^s_{\blacktriangleright} \end{array} & \begin{array}{c} p_s \\ B^s_{\blacktriangleright} \end{array} & p_{s+1} & p_{s+2} & \cdots \\ -(B^s_{\blacktriangleright})^\mathsf{T} & D^s_{\blacktriangleright} + h_{11} & h_{12} & & \\ & -h_{12}^\mathsf{T} & h_{22} + k_{11} & k_{12} & \\ & & -k_{12}^\mathsf{T} & k_{22} + \cdots & \ddots \\ & & & & \ddots & \ddots \end{pmatrix} \qquad (58)$$

Here, $A^s_{\blacktriangleright}$, $B^s_{\blacktriangleright}$, and $D^s_{\blacktriangleright}$ summarize the information from all sites to the left of s, and h (k) is the two-site Hamiltonian term acting on sites p_s, p_{s+1} (p_{s+1}, p_{s+2}), respectively. One update, step Eq. (54), maps this to

$$H^{s+1,N} = \begin{pmatrix} \begin{array}{c} l_{s+1} \\ A^{s+1}_{\blacktriangleright} \end{array} & \begin{array}{c} p_{s+1} \\ B^{s+1}_{\blacktriangleright} \end{array} & p_{s+2} & \cdots \\ -(B^{s+1}_{\blacktriangleright})^\mathsf{T} & D^{s+1}_{\blacktriangleright} + k_{11} & k_{12} & \\ & -k_{12}^\mathsf{T} & k_{22} + \cdots & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \qquad (59)$$

with the update rule

$$A^{s+1}_{\blacktriangleright} = \gamma_{lr}^\mathsf{T} A^s_{\blacktriangleright}\gamma_{lr} + \gamma_{pr}^\mathsf{T}(D^s_{\blacktriangleright} + h_{11})\gamma_{pr}$$
$$\qquad + \gamma_{lr}^\mathsf{T} B^s_{\blacktriangleright}\gamma_{pr} - \gamma_{pr}^\mathsf{T}(B^s_{\blacktriangleright})^\mathsf{T}\gamma_{lr}$$
$$B^{s+1}_{\blacktriangleright} = \gamma_{pr}^\mathsf{T} h_{12}$$
$$D^{s+1}_{\blacktriangleright} = h_{22}, \qquad (60)$$

where $\gamma \equiv \gamma^s$. We thus see that (i) the block-diagonal form of $H^{s,N}$ is preserved, (ii) the initial Hamiltonian is of this form with $A_{\blacktriangleright}^1 = 0$, $B_{\blacktriangleright}^1 = 0$, $D_{\blacktriangleright}^1 = 0$, and (iii) updating the information about the Hamiltonian requires only local updates.

Analogously, we can also keep the integrated Hamiltonian terms for all sites right of s through the update rule:

$$
\begin{aligned}
A_{\blacktriangleleft}^{s-1} &= \gamma_{rl}^{\mathsf{T}} A_{\blacktriangleleft}^s \gamma_{rl} + \gamma_{pl}^{\mathsf{T}}(h_{22} + D_{\blacktriangleleft}^s)\gamma_{pl} \\
&\quad + \gamma_{rl}^{\mathsf{T}} B_{\blacktriangleleft}^s \gamma_{pl} - \gamma_{pl}^{\mathsf{T}}(B_{\blacktriangleleft}^s)^{\mathsf{T}}\gamma_{rl}, \\
B_{\blacktriangleleft}^{s-1} &= \gamma_{pl}^{\mathsf{T}} h_{12}^{\mathsf{T}}, \\
D_{\blacktriangleleft}^{s-1} &= h_{11},
\end{aligned}
\tag{61}
$$

where again $\gamma = \gamma^s$, and h acts on sites p_{s-1}, p_s (with h_{11} corresponding to p_{s-1}, etc.), starting with $A_{\blacktriangleleft}^N = B_{\blacktriangleleft}^N = D_{\blacktriangleleft}^N = 0$.

For a state with working site $\hat{s}$, we find that

$$
\hat{H}^{\hat{s},\hat{s}} = \begin{pmatrix} A_{\blacktriangleright}^{\hat{s}} & B_{\blacktriangleright}^{\hat{s}} & 0 \\ -(B_{\blacktriangleright}^{\hat{s}})^{\mathsf{T}} & D_{\blacktriangleright}^{\hat{s}} + D_{\blacktriangleleft}^{\hat{s}} & -(B_{\blacktriangleleft}^{\hat{s}})^{\mathsf{T}} \\ 0 & B_{\blacktriangleleft}^{\hat{s}} & A_{\blacktriangleleft}^{\hat{s}} \end{pmatrix}
\tag{62}
$$

with mode ordering lpr. Similarly, we can compute $E^{\hat{s},\hat{s}} = E_{\blacktriangleright}^{\hat{s}} + E_{\blacktriangleleft}^{\hat{s}}$, where $E_{\blacktriangleright}^1 = E_{\blacktriangleleft}^N = 0$, and

$$
\begin{aligned}
E_{\blacktriangleright}^{s+1} &= E_{\blacktriangleright}^s + \mathrm{tr}\big[H_{l_s p_s | l_s p_s}^{s,N} \gamma_{l_s p_s | l_s p_s}^s\big] \\
&= E_{\blacktriangleright}^s + \mathrm{tr}\left[\begin{pmatrix} A_{\blacktriangleright}^s & B_{\blacktriangleright}^s \\ -(B_{\blacktriangleright}^s)^{\mathsf{T}} & D_{\blacktriangleright}^s + h_{11}^s \end{pmatrix}\gamma_{lp|lp}^s\right],
\end{aligned}
\tag{63}
$$

and accordingly,

$$
E_{\blacktriangleleft}^{s-1} = E_{\blacktriangleleft}^s + \mathrm{tr}\left[\begin{pmatrix} h_{22}^{s-1} + D_{\blacktriangleleft}^s & -(B_{\blacktriangleleft}^s)^{\mathsf{T}} B_{\blacktriangleleft}^s & A_{\blacktriangleleft}^s \end{pmatrix}\gamma_{pr|pr}^s\right],
\tag{64}
$$

where h_{11}^s (h_{22}^s) are the left (right) part of the Hamiltonian term acting between sites s and $s+1$. Overall, the dependence of the energy on $\gamma^{\hat{s}}$ is then given by

$$
E_{\mathrm{tot}}(\gamma^{\hat{s}}) = \mathrm{tr}[H^{\hat{s},\hat{s}} \gamma^{\hat{s}}] + E^{\hat{s},\hat{s}}.
\tag{65}
$$

Since all update rules for the A, B, D, and E are local, they can be updated (and thus the energy dependence on $\gamma^{\hat{s}}$ computed) at a cost independent of the system size when moving the working site.

The generalization beyond nearest-neighbor Hamiltonians follows the same pattern illustrated above, but the Hamiltonian of Eq. (58) will have additional off-diagonal terms. In particular, the first row/column, which contains coupling terms between sites $q < s$ and $p \geqslant s$, will be modified as follows: (i) the terms $A_{\blacktriangleright}^s$ and $B_{\blacktriangleright}^s$ will contain additional contributions from terms between $q < s$ and s, and (ii) there will be additional off-diagonal blocks similar to $B_{\blacktriangleright}^s$ for coupling terms between $q < s$ and $p > s$. The update rule (60) is adapted by applying the rule for $B_{\blacktriangleright}^{s+1}$ for all similar off-diagonal terms and including additional terms in $A_{\blacktriangleright}^{s+1}$. The update for the constant term can be adapted similarly. The total number of blocks that need to be treated in each step depends on the total number of Hamiltonian terms between sites $q < s$ and $p \geqslant s$.

D. Energy minimization

A core ingredient in DMRG algorithms is the optimization of the energy as a function of a single tensor $\gamma^{\hat{s}} \equiv \gamma$. As we have seen, this energy dependence can be summed up in an effective Hamiltonian $H^{\hat{s},\hat{s}} \equiv H$ and constant offset $E^{\hat{s},\hat{s}} \equiv E$. In order to minimize $E(\gamma) = \mathrm{tr}(H\gamma) + E$, we need to fill all modes with negative energy. This can be done by going to the eigenbasis of H,

$$
H \cong \bigoplus_i \begin{pmatrix} 0 & e_i \\ -e_i & 0 \end{pmatrix},
\tag{66}
$$

and choosing

$$
\gamma \cong \bigoplus_i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\tag{67}
$$

Numerically, this can be done by diagonalizing the antisymmetric matrix H (giving imaginary eigenvalues) and then replacing them by their sign, $\gamma = \mathrm{sign}(H)$, or by following the numerical approaches discussed in Ref. [48].

V. GFMPS OPTIMIZATION ALGORITHMS

We now describe some commonly used MPS algorithms specifically in the context of GFMPS. Most importantly, we will describe how to carry out the DMRG algorithm in its formulation as a variational method over GFMPS. In addition, we will discuss how to implement the time-dependent variational principle (TDVP), which is suitable for both time evolution and ground-state simulations, with GFMPS, as well as a translational invariant version of those methods.

A. Gaussian fermionic DMRG

The DMRG algorithm can be naturally described as a variational method over the family of MPS with a given bond dimension. The key idea is to sweep through the system and sequentially optimize a small number (usually one or two) of GFMPS tensors in each step. Each such optimization is a quadratic problem and can thus be efficiently solved. In the case where more than one tensor is optimized, the state can be brought back into MPS form using the singular value decomposition. While local convergence does not guarantee global convergence of the energy, this algorithm is empirically found to perform extremely well.

1. The algorithm

Gaussian fermionic DMRG, i.e., DMRG for a quadratic fermionic Hamiltonian and with GFMPS as a variational family, follows closely the conventional way of doing DMRG with MPS (see, e.g., [4,5]). It consists of an initialization step followed by a number of sweeps, where each sweep in turn consists of a right and a left sweep. For concreteness, we will consider the case of a nearest-neighbor Hamiltonian, but the same steps equally apply to more complex Hamiltonians (with suitable modifications regarding the computation of effective Hamiltonians, cf. Sec. IV). The notation follows that used in the previous sections. In addition, we will denote the complete set of left and right boundary terms that enter the effective Hamiltonian at position s by $\mathcal{B}_{\blacktriangleright}^s = \{A_{\blacktriangleright}^s, B_{\blacktriangleright}^s, D_{\blacktriangleright}^s, E_{\blacktriangleright}^s\}$ and

$\mathcal{B}_\blacktriangleleft^s = \{A_\blacktriangleleft^s, B_\blacktriangleleft^s, D_\blacktriangleleft^s, E_\blacktriangleleft^s\}$, respectively. We begin by describing the variant where only one tensor is optimized at a time ("single-site DMRG").

(i) *Initialization.*— Fix a bond mode number for each bond $(s, s+1)$, $s = 1, \ldots, N-1$, and choose a random (pure) initial tensor γ^s for each site s. Let $\mathcal{B}_\blacktriangleleft^N = \{\varnothing, \varnothing, \varnothing, \varnothing\}$, where by $\varnothing$ we denote a 0×0 matrix.
For $\hat{s} = N, N-1, \ldots, 2$ do:

(1) Bring $\gamma^{\hat{s}}$ into right-canonical form (Sec. III B).
(2) Compute $\mathcal{B}_\blacktriangleleft^{\hat{s}-1}$ as given by Eq. (61), and store the result.

After the initialization, we are left with a GFMPS with all tensors in right-canonical form.

(ii) *Right sweep.*— Let $\mathcal{B}_\blacktriangleright^1 = \{\varnothing, \varnothing, \varnothing, \varnothing\}$.
For $\hat{s} = 1, \ldots, N-1$ do:

(1) Compute $H^{\hat{s},\hat{s}}$ from $\mathcal{B}_\blacktriangleright^{\hat{s}}$ using Eq. (62) (or a suitable modification, based on Sec. IV, for longer ranged Hamiltonians).
(2) Choose $\gamma^{\hat{s}}$ such as to optimize the energy $\text{tr}[H^{\hat{s},\hat{s}} \gamma^{\hat{s}}]$.
(3) Bring $\gamma^{\hat{s}}$ into left-canonical form (updating $\gamma^{\hat{s}+1}$ correspondingly).
(4) (Re-)compute $\mathcal{B}_\blacktriangleright^{\hat{s}+1}$.

(iii) *Left sweep.*— The left sweep works analogous to the right sweep, just that one iterates over $\hat{s} = N, N-1, \ldots, 2$, and steps 3 and 4 are replaced by bringing $\gamma^{\hat{s}}$ into right-canonical form and updating $\mathcal{B}_\blacktriangleleft^{\hat{s}-1}$, respectively.

To find the ground state, one performs several left and right sweeps until convergence is reached. Convergence is typically tested for using the ground-state energy (see also our discussion in Sec. VI).

A straightforward generalization of this algorithm is to the "two-site DMRG" algorithm, which optimizes two adjacent tensors in each step. This coincides with the original DMRG algorithm due to White [3]. Given a GFMPS in canonical form around a center site $\hat{s}$ during a sweep to the right, the tensors on sites $\hat{s}, \hat{s}+1$ are optimized by first using the steps laid out in Sec. IV to obtain an effective Hamiltonian $H^{(\hat{s},\hat{s}+1),(\hat{s},\hat{s}+1)}$. One then finds the ground state of this Hamiltonian on the modes $l_{\hat{s}}, p_{\hat{s}}, p_{\hat{s}+1}, r_{\hat{s}+1}$. Using the singular value decomposition of Sec. II D 1 and absorbing the middle tensor Λ into the right tensor, one can put the state back into canonical GFMPS form with a new center site $\hat{s}' = \hat{s}+1$; therefore, no extra steps are required to keep the state canonical. When sweeping to the left, the same steps are performed on sites $\hat{s}-1, \hat{s}$, and in the final SVD Λ is absorbed into the left tensor. During the singular value decomposition, the bond dimension can be truncated as appropriate.

In conventional DMRG, the single-site algorithm needs to be augmented with the technique described in Ref. [54]. This addresses two issues of the single-site algorithm, namely, that the number of states on a bond is difficult to adapt dynamically and a tendency to become stuck in local minima. While this technique does not have a simple generalization to the GFMPS case, we find in practice that for a fixed bond dimension, the convergence of single- and two-site DMRG on GFMPS is very similar; furthermore, given the much lower cost of the GFMPS-based algorithms, it is often not necessary to dynamically adapt the bond number.

2. Scaling and efficiency

All steps in the Gaussian fermionic DMRG (GFDMRG) algorithm can be carried out at a cost of $\mathcal{O}[(\chi + p)^3]$, where χ is the number of Majorana modes on each bond and p the number of physical Majorana modes on each site. In practical numerical implementations, the most costly step will usually be the inversion required to contract two tensors, which can be optimized using the block formulas of Eqs. (19) and (20). The scaling of a single sweep with system size is $\mathcal{O}(N)$. The number of sweeps required to reach convergence heavily depends on the underlying problem and cannot in general be bounded.

Additional cost is incurred if the Hamiltonian is not nearest neighbor. In particular, the number of additional blocks that need to be included when computing the effective Hamiltonian at site s depends on the number of Hamiltonian terms that couples sites $p < s$ and $q \geqslant s$. Therefore, the scaling depends crucially on details of the Hamiltonian. In a simple case, such as periodic boundary conditions leading to a single additional term between the left and the right end of the system, the additional cost is constant, while in the most general case of coupling between all sites an additional factor N^2 could be incurred.

In contrast to conventional DMRG algorithms, the performance can be improved significantly by blocking a group of n_b sites together. To understand this, consider a system of N sites with p Majorana modes on each site and a GFMPS description with Majorana bond number χ. In this case, the matrices of the MPS will be of size $(2\chi + p) \times (2\chi + p)$, and the computational cost of a sweep will scale as $L(2\chi + p)^3$. If we group n_b sites into 1, we instead obtain a scaling of $(L/n_b)(2\chi + n_b p)^3$. In general, one has $\chi \gg p$, and optimal performance will be achieved for $n_b > 1$. For local Hamiltonians, the optimal choice is $n_b = \chi/2$. For nonlocal Hamiltonians, an additional advantage can arise because the Hamiltonian may become less long-ranged if sites are blocked together.

Finally, in the discussion so far we have considered general Gaussian states rather than states with a fixed particle number. In principle, Gaussian states with fixed particle number can be described by matrices that are smaller by a factor of 2, thus potentially leading to a speedup of order 2^3.

We will perform a more detailed analysis of the performance for specific models in Sec. VI B 2.

B. Time-dependent variational principle

The time-dependent variational principle (TDVP) can be adapted to use MPS [26] as a variational manifold to approximately solve the time-dependent Schrödinger equation both in real and imaginary time. In Ref. [27], it is shown that TDVP can be reformulated as a modification of the DMRG method, with only a few modifications. In the following, we discuss how to modify the GFDMRG algorithm described above to obtain TDVP for GFMPS following the recipe of Ref. [27].

To perform the evolution for some sufficiently small time step t (where real time evolution corresponds to real t, and imaginary time evolution to $t = i\tau$, $\tau > 0$), steps 2, 3, and 4 in step (ii) (right sweep) are replaced by the following sequence:

(2a) Evolve $\gamma^{\hat{s}}$ under $H^{\hat{s},\hat{s}}$ for time t.

(3a) Perform step (I) of the elementary move to bring $\gamma^{\hat{s}}$ into left-canonical form, see Sec. III B and Fig. 6, i.e., bring $\gamma^{\hat{s}}$ into left-canonical form $\hat{\gamma}^{\hat{s}}$ by splitting off the SVD Λ, but do not absorb Λ in $\gamma^{\hat{s}+1}$ yet.

(4) (Re-)compute $\mathcal{B}_{\blacktriangleright}^{\hat{s}+1}$ using the new (left-canonical) $\hat{\gamma}^{\hat{s}}$. (Since $\mathcal{B}^{\hat{s}+1}$ does not depend on $\gamma^{\hat{s}+1}$, this is already the correct $\mathcal{B}^{\hat{s}+1}$.)

(2b) Compute the effective Hamiltonian H_Λ for Λ, which describes the dependence of the energy on Λ:

$$H_\Lambda = \begin{pmatrix} A_{\blacktriangleright}^{\hat{s}+1} & B_{\blacktriangleright}^{\hat{s}+1}\gamma_{pl}^{\hat{s}+1} \\ -\left(B_{\blacktriangleright}^{\hat{s}+1}\gamma_{pl}^{\hat{s}+1}\right)^{\mathsf{T}} & A_{\blacktriangleleft}^{\hat{s}} \end{pmatrix}, \quad (68)$$

and evolve Λ with H_Λ for time $-t$.

(3b) Absorb the evolved Λ into $\gamma^{\hat{s}+1}$.

A core ingredient used in steps 2a and 2b is to evolve a state ρ with covariance matrix Γ with a time-independent Hamiltonian $\mathcal{H} = i\sum H_{ij}c_i c_j$ either in real or imaginary time. In the covariance matrix formalism, the Schrödinger equation in real time takes the form

$$\Gamma(0) \mapsto \Gamma(t) = O(t)\Gamma(0)O(t)^{\mathsf{T}}, \quad (69)$$

with $O(t) = \exp(4Ht)$ [55]. For imaginary time, this is replaced by integrating [56]

$$\dot{\Gamma}(t) = -4[H + \Gamma(t)H\,\Gamma(t)]. \quad (70)$$

For details on the implementation, such as the optimal choice of time discretization as well as the generalization to a two-site version of the algorithm that allows a dynamical choice of the bond dimension, we refer the reader to Ref. [27]. Note that in the case where $t \to i\infty$, we recover the DMRG algorithm.

C. Infinite systems

It was already pointed out in Ref. [3] that DMRG is in principle suitable for infinite systems. While this approach initially enjoyed some success, for a long time DMRG for finite systems was considered more accurate. More recently, however, DMRG for infinite systems has been reinterpreted in the language of matrix product states, which has allowed for accurate and efficient methods for infinite, translationally invariant MPS to be developed [5,57–59]. The underlying idea of this reinterpretation is to consider an MPS made up of a unit cell of K sites that repeats indefinitely; in the simplest case, $K = 1$, the tensor is the same on each site of the lattice. It turns out that most MPS algorithms can be generalized to such an infinite, translationally invariant ansatz by reformulating them in terms of eigenvectors of transfer operators along the MPS.

While noninteracting systems can be solved directly in the thermodynamic limit by expressing the problem in momentum space, such an approach may scale unfavorably with the number of sites in the unit cell and does not give access to real-space correlation functions without solving for all momenta and transforming back to real space. Therefore, for a large number of sites in the unit cell, a GFMPS approach directly in the thermodynamic limit will be advantageous.

To illustrate how GFMPS can be generalized to the infinite case, we will limit ourselves to the infinite DMRG (iDMRG) algorithm and restrict the discussion to the aspects which need to adapted in the GFMPS case. However, similar generalizations are also possible for other infinite MPS methods, such as variational uniform MPS (VUMPS) [59]. We begin by briefly reviewing the conventional iDMRG algorithm of Ref. [3], where we specialize to the case of a two-site Hamiltonian H. Our description follows closely Ref. [5], which we refer to for details. The algorithm proceeds in the following steps:

(1) Compute the covariance matrix $\Gamma^{n=1}$ corresponding to the ground state of H on two sites and perform an SVD such that Γ is given as the contraction of three tensors A^n, Λ^n, and B^n, where A^n and B^n are left- and right-canonical MPS tensors, respectively.

(2) Extend the lattice by two sites in the middle and consider an MPS where the tensor immediately to the left (right) of the newly inserted sites is given by A^n (B^n). Follow the steps outlined in Sec. IV, especially Eqs. (60)–(62) (adapted for two center sites), to obtain the effective Hamiltonian for the two sites newly added in the center. Use Eqs. (63) and (64) to obtain the constant part of the energy.

(3) Find the ground state Γ^{n+1} of this new effective Hamiltonian and again perform a singular value decomposition to obtain A^{n+1}, Λ^{n+1}, and B^{n+1}. Return to step 2.

All of the steps outlined here can be performed using techniques already discussed. Note that this approach effectively simulates a finite system of length $2n$, where n is the number of iterations performed. The energy obtained as the sum of the constant term and the effective Hamiltonian for the center sites is variational for this system of $2n$ sites. One generally finds that for sufficiently large n, the tensors converge.

One can now take the tensors obtained in the last step and use them to form an infinite, translationally invariant MPS [5,57]. Let A^n, Λ^n, and B^n be the tensors obtained in the last iteration of the infinite DMRG algorithm, and Λ^{n-1} the central tensor from the previous iteration. We then define, analogous to Ref. [60],

$$\Gamma^A = (\Lambda^{n-1})^{-1} \rhd A^n \quad \lambda^A = \Lambda^n \quad (71)$$

$$\Gamma^B = A^n \rhd (\Lambda^{n-1})^{-1} \quad \lambda^B = \Lambda^{n-1}. \quad (72)$$

(In the following, we use the notation $A \rhd B$ to denote the contraction of tensors; which indices are to be contracted follows from the structure of the tensor network.) Here, Λ^{-1} denotes a square rank-2 tensor such that $\Lambda \rhd \Lambda^{-1}$ yields the maximally entangled state between the two sets of modes. From the SVD, Λ is of the form $\Lambda = \bigoplus W(\lambda_k)$ [see Eq. (30)]; for a tensor of this form, the inverse of this sense is given by exchanging the left and right modes (i.e., mapping $\lambda_k \to -\lambda_k$). This can be verified explicitly using Eq. (18).

If the state is well converged and the bond dimension large enough, $\gamma^{L,1} = \lambda^B \rhd \Gamma^A \rhd \lambda^A \rhd \Gamma^B$ and $\gamma^{L,2} = \lambda^A \rhd \Gamma^B \rhd \lambda^B \rhd \Gamma^A$ will be approximately left-canonical, and $\gamma^{R,1} = \Gamma^A \rhd \lambda^A \rhd \Gamma^B \rhd \lambda^B$ and $\gamma^{R,2} = \Gamma^B \rhd \lambda^B \rhd \Gamma^A \rhd \lambda^A$ will be approximately right-canonical. Expectation values of local observables, such as the energy on a bond, can

be obtained from the observation that the reduced density matrix for the two-site unit cell AB is given by $(\lambda^B \rhd \Gamma^A \rhd \lambda^A \rhd \Gamma^B \rhd \lambda^B)_{p^A p^B | p^A p^B}$, where the physical indices in the subscript refer to the physical indices of Γ^A and Γ^B. Note that for the average energy, one must compute the energy for the AB and BA unit cells separately and average.

VI. NUMERICAL RESULTS

Below we test our numerical implementation of the algorithms presented in the previous sections for three relevant cases: We first study a one-dimensional system that exhibits nontrivial spatial correlation functions but can be exactly solved, thus providing a simple comparison. Then we turn to a two-dimensional system, which we solve on quasi-two-dimensional cylinder geometries, to assess how accurately our approach works beyond a strictly one-dimensional system and when it performs better than exact approaches. Finally, we study charge transport through a quantum point contact in a simple geometry at finite bias voltage to demonstrate the capability of TDVP to compute dynamical properties of nanostructures.

While the models discussed in this section preserve particle number, we use the Majorana formalism described throughout this manuscript without exploiting particle number conservation.

A. The resonant-level model

First, we consider a one-dimensional system which still exhibits nontrivial real-space correlation functions: the resonant-level model, which appears as the noninteracting limit of a number of impurity models (see, e.g., Refs. [61–64]). Here, we follow the discussion of Ref. [65]. We consider a fermionic impurity with a corresponding creation operator $d^\dagger$ coupled to a bath of fermions, which here is represented by a periodic chain of L fermions with associated creation operators $c_i^\dagger$. The Hamiltonian is given by

$$H = J'(d^\dagger c_1 + c_1^\dagger d) + JT_{\text{chain}}, \tag{73}$$

$$T_{\text{chain}} = \sum_{i}^{L-1}(c_i^\dagger c_{i+1} + c_{i+1}^\dagger c_i) + (c_L^\dagger c_1 + c_1^\dagger c_L). \tag{74}$$

Here, T is the kinetic energy operator on the chain and J sets the bandwidth; we set the unit of energy to $J = 1$ and fix the system to half filling, leaving J' as the only free parameter aside from system size.

The quantity of interest is the correlation between the impurity site and the sites in the bath, $C(r) = \langle d^\dagger c_r \rangle / J'$. Owing to the noninteracting nature of the problem, this correlation function can be computed exactly in the thermodynamic limit (assuming a constant density of states around the Fermi energy) and at zero temperature is found to be given by

$$C(r) = Af(BJ'^2 r), \tag{75}$$

$$f(\kappa) = -\pi \int_{-\infty}^{\infty} dx \frac{x\cos(\kappa x) + \sin(\kappa x)}{[x^2+1][1+\delta(x)]}. \tag{76}$$
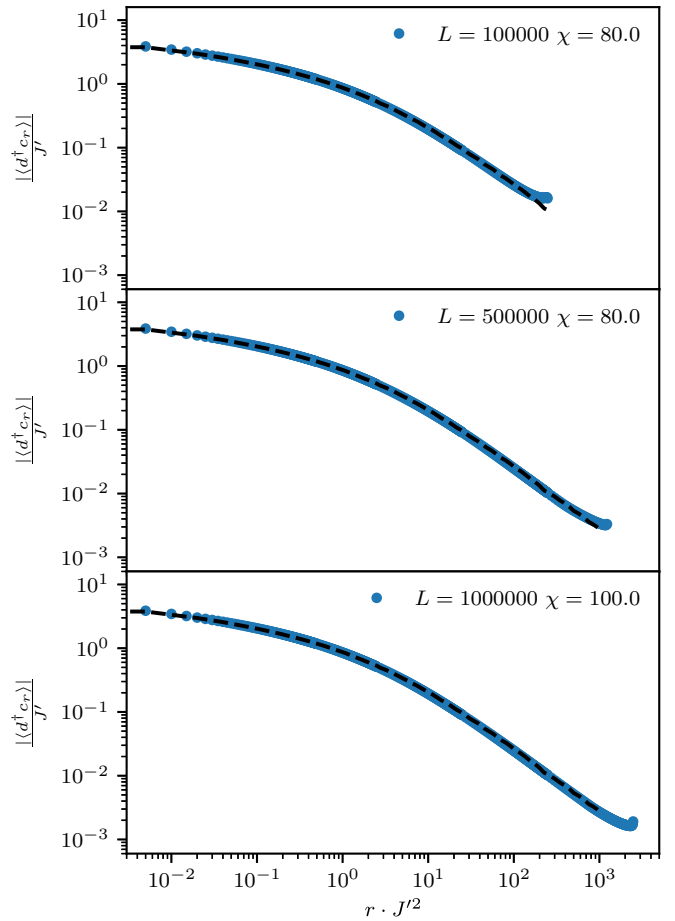

FIG. 9. Rescaled real-space correlation function $\langle d^\dagger c_r \rangle$ of the ground state of Eq. (73) as a function of (rescaled) distance $r \cdot J'^2$. Other parameters are $J' = 0.05J$, and the number of sites per block $n_b = \chi/2$ (see Sec. V A 2). χ denotes the maximum Majorana bond number, and L is the size of the system including the impurity. Dashed lines indicate a fit to the exact solution (75).

The constants A and B are independent of L and J' but depend on the microscopic choice for the bath, such as the Fermi velocity and density of states at the Fermi energy, and are taken as fit parameters here; however, they can be computed exactly for certain choices of the bath [65]. The relevant universal behavior is that $C(r)$ shows a characteristic crossover from short-distance to long-distance behavior around $\kappa = BJ'^2 r \approx 1$; for $\kappa \ll 1$, $f(\kappa) \sim \ln(\kappa) + \gamma$, while for $\kappa \gg 1$, $f(\kappa) \sim 1/\kappa$.

In the previous literature, this crossover was illustrated by collapsing many datasets for different values of J', each covering a different range of κ. Since we are able to study systems almost 3 orders of magnitude larger than previously possible, we can exhibit the full scaling form in a *single* dataset. This is shown in Fig. 9. Here, χ denotes the maximum number of Majorana modes on the bonds of the MPS. This corresponds to a conventional bond dimension of $D(\chi) = \sqrt{2}^\chi$; for $\chi = 80$, this works out to $D \approx 10^{12}$ and for $\chi = 100$ to $D \approx 10^{15}$. The dashed line indicates the exact solution of Eq. (75), with A, B chosen as best fit to the MPS data. We see that agreement is essentially exact up to the largest distances, which are weakly affected by finite-size effects.
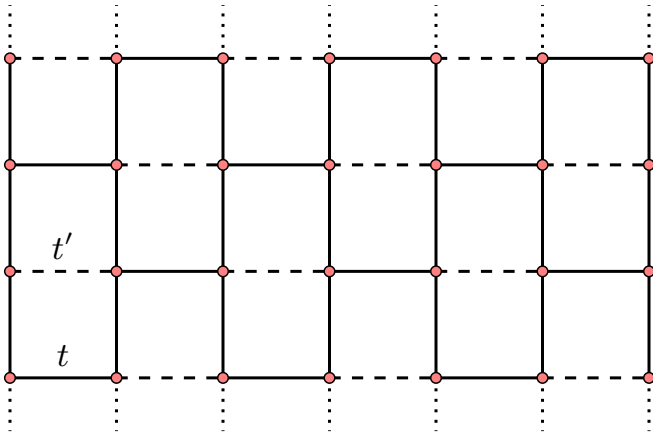
FIG. 10. Lattice used to define the two-dimensional Hamiltonian, Eq. (77). Solid lines indicate bonds with strength t, and dashed lines indicate bonds with strength t'. For $t' = t$, the square lattice is recovered, while for $t' = 0$ the brick-wall representation of the honeycomb lattice is obtained. The dotted lines indicate periodic boundary conditions in one direction, effectively wrapping the system onto a cylinder.

B. Quasi-two-dimensional systems

For a more challenging test, we turn to a model of spinless fermions hopping on a two-dimensional lattice,

$$H_{\text{hop}} = -\sum t_{xy}^{x'y'} (c_{xy}^\dagger c_{x'y'} + c_{x'y'}^\dagger c_{xy}), \qquad (77)$$

where $c_{xy}^\dagger$ creates a fermion on the xth (yth) site in the horizontal (vertical) direction, and the hopping elements $t_{xy}^{x'y'}$ are chosen to give rise to the brick-wall lattice as sketched in Fig. 10. There are three particular limits of this model that we are interested in: (1) the limit where $t' = t$, where the system simply becomes the square lattice; (2) the limit of $t' = 0$, where it becomes the brick-wall lattice, whose connectivity equals that of the honeycomb lattice; and (3) the limit of $t = 0$, where it becomes a lattice of isolated dimers.

From the point of view of a matrix product state description, the most relevant property of the system is its entanglement structure, which differs significantly between these three limits. Most easily understood is the case of isolated dimers: in this case, the spectrum is fully gapped with a gap of $2t$, and there is only entanglement between the sites that together form a dimer, and no entanglement (or correlations) otherwise. The state thus has a very simple MPS description. Upon adding the t term and interpolating to the square lattice limit $t = t'$, the gap decreases and finally the system becomes gapless at $t = t'$. Correspondingly, the entanglement increases. The square lattice hosts a gapless state with one-dimensional Fermi surface. This leads to a logarithmic violation of the area law [66–69] and thus represents the most difficult case for a tensor network state approach. Finally, in the case of the brick-wall/honeycomb lattice, the system is gapless but with Dirac points instead of Fermi surfaces. In this limit, the entanglement follows an area law despite the gapless nature of the system.

In the following we will focus on the case of the honeycomb/brick-wall lattice and defer a discussion of

the other cases to Appendix A. Due to the area law in the honeycomb/brick-wall case, the maximum entropy that must be captured in the MPS scales linearly in its width W. We therefore choose χ proportional to the width of system W.

For the numerical simulations using DMRG methods, we study this lattice on cylinders of circumference W and length L, i.e., with a total of $W \cdot L$ sites. We map the sites to a chain following each rung (of length W) from bottom to top. We have found that the single-site and two-site optimization algorithm generally lead to comparable results and thus use the single-site algorithm throughout. We perform at least five sweeps and terminate when the absolute change in energy between sweeps drops below 10^{-3}.

1. Convergence of real-space correlation functions

To characterize how well the MPS ground state approximates the exact one, we study the real-space correlations. Figure 11 shows the decay of the equal-time Green's function along the long direction of the cylinder in the limit of a brick-wall/honeycomb lattice, $t' = 0$. Here, we scale the Majorana bond number with the width of the cylinder, showing $\chi = 2W, 4W, 6W$. This choice is motivated by the area law.

It is well known that for finite bond dimension, correlations in an MPS decay exponentially at sufficiently long distances [1]. At shorter distances, the correlations can approximate a polynomial decay. This is clearly borne out in the data of Fig. 11: for small Majorana bond number χ, the correlations exhibit an artificial and unphysical exponential decay, while for sufficiently large χ the exact polynomial decay $\mathcal{O}(1/l)$ is recovered to high accuracy.

2. Performance comparison against exact methods

We now examine the runtime of the MPS algorithm for quasi-one-dimensional systems compared to exact, full diagonalization methods. Of course, absolute runtimes are implementation specific and particularly in the case of MPS methods, may depend sensitively on details such as the choice of initial states, convergence criterion, and many others. Therefore an analysis of the scaling of the method is more pertinent.

Exact methods based on fully diagonalizing the hopping matrix of the underlying problem will scale as the cube of the matrix size, i.e., in the case at hand as $\mathcal{O}(W^3 L^3)$. Estimating the cost of the MPS method is more challenging since, in general, we cannot make any strong assertions on the convergence with the number of sweeps [70] or the bond dimension. However, we can quantify the cost of a single sweep for fixed bond dimension. As stated in Sec. V A 2, the cost is $(N/n_b)(2\chi + n_b p)^3 D$, where N is the number of sites, p the number of Majorana modes per site, n_b the number of physical sites that are grouped into one site of the MPS, χ the Majorana bond number, and D the number of operators that must be tracked to accommodate long-range hopping terms, which depends on the details of the Hamiltonian and n_b. For the case of the quasi-one-dimensional system at hand, as we have seen above, to achieve fixed accuracy we scale $\chi \sim W$ and block a number of sites $n_b = \chi/2$ for optimal performance; therefore, $n_b \sim W$. For this case, it turns out
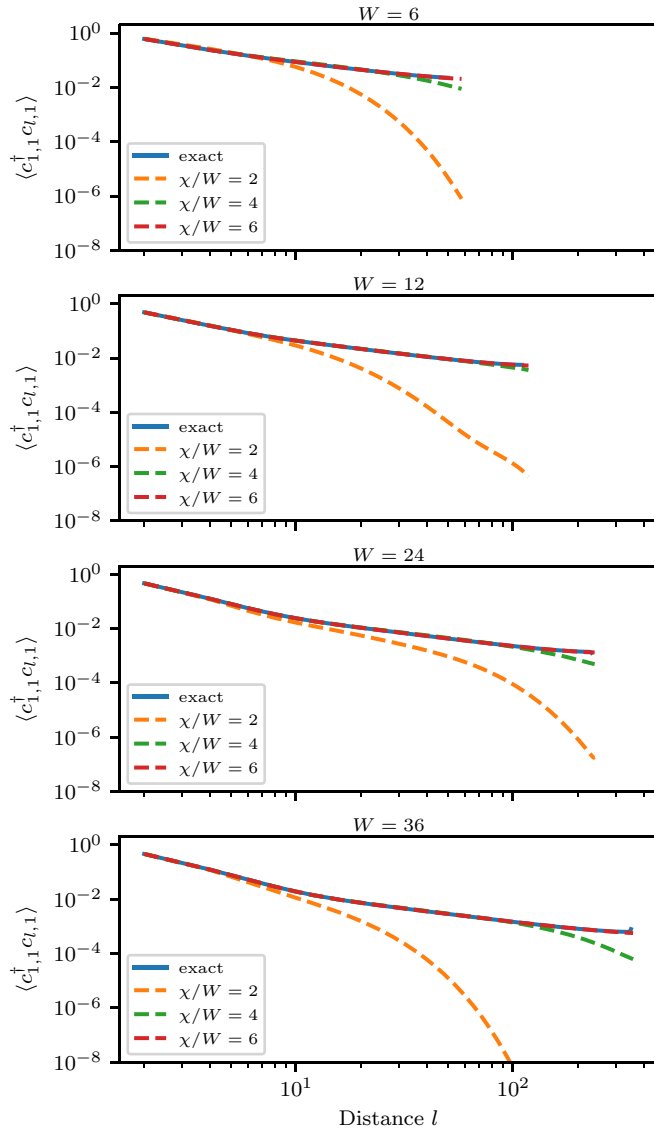
FIG. 11. Real-space correlation functions for the ground state of Eq. (77) with $t = 1, t' = 0$. Correlations are shown between two sites of the same (odd) vertical position and different horizontal positions, i.e., probing the decay of correlations along the cylinder. These results are obtained using the single-site optimization algorithm with W sites blocked into one for systems of aspect ratio 10, i.e., $L = 10W$. Exact results are obtained from a full diagonalization of the same system.

that D is a constant independent of W. Therefore, the overall scaling becomes LW^3. For the case where we fix the aspect ratio, $L \sim W$, we thus obtain a scaling of $\mathcal{O}(W^6)$ for the exact methods and $\mathcal{O}(W^4)$ for the MPS method.

The memory usage of the exact method scales as $(WL)^2$, while the GFMPS approach scales as $(WL/n_b)(2\chi + n_b p)^2$. For the choices made above ($W \sim L, \chi, n_b$), this leads to the exact method scaling as W^4 and the GFMPS approach scaling as W^3. In practice, given our choice of aspect ratio, the memory requirements of the exact approach are much greater than the GFMPS approach and turn out to be the limiting factor.
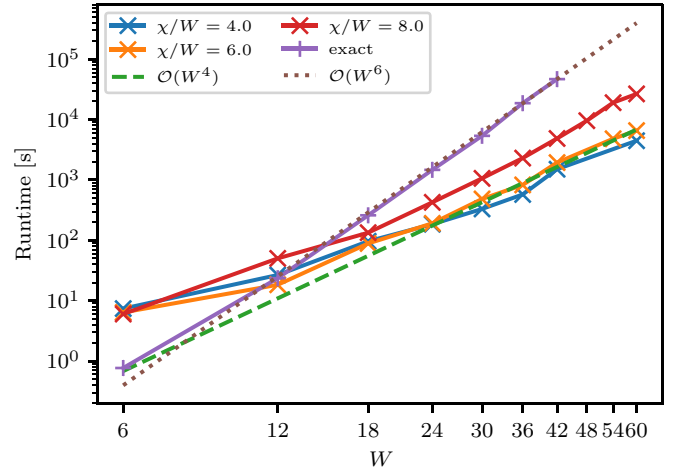

FIG. 12. Run time to a converged solution for the ground state of Eq. (77) with $t = 1, t' = 0$ for aspect ratio $L = 20W$, i.e., the largest system has $60^2 \times 20 = 72\,000$ sites. The dashed line indicates a best fit of the GFMPS data for $\chi/W = 6$ to aW^4, with a a fit parameter, and the dotted line is a best fit of the exact method to aW^6. These simulations are performed on four cores of an Intel(R) Xeon(R) E5-2690 v3 @ 2.60 GHz.

We numerically confirm the scaling of the runtime in Fig. 12, where we show the time to achieve a converged solution instead of the time to perform a single sweep. The GFMPS approach shows polynomial scaling consistent with W^4 across a range of bond dimensions. This is consistent with the empirical expectation that the number of sweeps scales at most very weakly with the size of the system.

For the parameters chosen here and our implementation, the crossover where the GFMPS method becomes faster than exact approaches is around $W = 14$. This is most strongly influenced by the aspect ratio: upon approaching the one-dimensional limit, the GFMPS approach will become more and more favorable.

C. Transport through a quantum point contact

To illustrate the power of the time-dependent variational principle applied to GFMPS, we study a simple transport problem, namely, the quantization of conductance through a quantum point contact in a multichannel wire [71]. We model the system as a square lattice of spinless fermions of size $W \times L$, i.e., as shown in Fig. 10 with hopping Hamiltonian (77) for $t' = t = 1$. We add an on-site potential

$$H_{\text{pot}} = \sum \left[-\mu_0 + V_B \exp \left(-\frac{(x - L/2)^2}{d^2} \right) \right] n_{xy}, \quad (78)$$

where $n_{xy} = c_{xy}^\dagger c_{xy}$. This corresponds to a Gaussian potential barrier of height V_B and width d at the center of the system in x direction and independent of y.

We initialize the system in the ground state of $H_{\text{hop}} + H_{\text{pot}}$ using the single-site DMRG algorithm and then apply a voltage V by evolving under the Hamiltonian $H_{\text{hop}} + H_{\text{pot}} + H_{\text{bias}}$
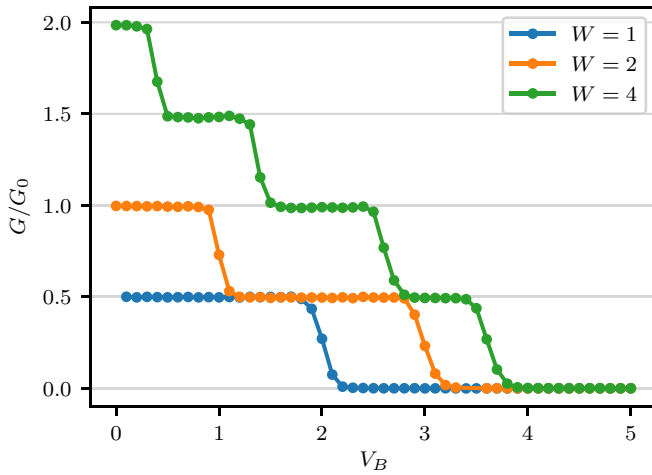
FIG. 13. GFMPS simulation of the quantized conductance across a Gaussian potential barrier of height V_B in a $W \times 500$ system. For details, see main text.

with

$$H_{\text{bias}} = \frac{V}{2} \sum_y \left(\sum_{x=1}^{L/2-4d} n_{xy} - \sum_{x=L/2+4d}^{L} n_{xy} \right). \quad (79)$$

Here, we use the sites that are more than $4d$ away from the center of the barrier as leads and apply a symmetric bias voltage, and consider the region between the leads as the quantum point contact. We then measure the charge in the left half of the system, $N_{\text{left}} = \sum_y \sum_{x=1}^{L/2} c_{xy}^\dagger c_{xy}$, as a function of time. Due to conservation of charge, we can compute the current flowing across the junction as $J(t) = dN_{\text{left}}/dt$. Alternatively, we could evaluate the charge operator across the junction.

After a transient time of order d, the current settles to a plateau value that will last until the charge transport reaches the end of the leads after a time of order L. We compute the average current $\bar{J}$ over this plateau and define the conductance as $G = \bar{J}/V$. According to well-known theoretical calculations, in the limit of a smooth barrier the conductance should be quantized as $G = NG_0/2$, where N is the number of conducting *spin* channels (i.e., for a usual spin-degenerate system, $N = 2K$ with K the number of spinful channels, but here we consider a spinless system) and $G_0 = 2e^2/h$. We use units where $e = \hbar = 1$, and thus $G_0 = 1/\pi$.

For $V_B = 0$, we expect the system to have W conducting channels and thus $G = WG_0/2$; as we increase the barrier height, the number of channels is reduced one by one as the lowest-momentum channels are cut off first. We confirm this behavior in Fig. 13 for a system with $W = 4$, $L = 500$, $d = 10$, and applied bias voltage $V = 0.1$. Our simulations are performed with a GFMPS of Majorana bond number $\chi = 50$ and with $n_b = W$, and TDVP is run with a time step of $dt = 0.05$. We also note that since the Hamiltonian is time-independent after the quench, a time t can be reached in one step for exact simulations while requiring t/dt steps in TDVP; however, one can easily generalize the TDVP

simulation to a time-dependent Hamiltonian, as would arise, for example, when computing the ac conductance, without incurring additional computational cost.

VII. CONCLUSIONS

In this paper, we have studied how to apply the DMRG algorithm and other MPS-based algorithms to the simulation of systems of noninteracting fermions. By combining the advantages of the exponentially compressed description of noninteracting fermionic states in terms of second moments and the efficient description of states with bounded (area-law) entanglement through matrix product states we were able to simulate systems far beyond what is possible with either method alone. The central insight of our GFDMRG algorithm is the use of a suitable canonical form which allows us to express the minimization of the energy as a function of a given tensor as an eigenvalue problem, much like in conventional DMRG. This enabled us to realize the variational algorithm for noninteracting fermions in close analogy to the conventional MPS-based formulation of DMRG, as well as to generalize the method to other MPS-based algorithms such as TDVP for time evolution or iDMRG for infinite systems.

The computational cost of the GFDMRG method scales as $L\chi^3$ per sweep, with L the length of the chain and χ the number of modes per bond. Due to the exponential compression of the representation in terms of second moments, the effective bond dimension is $D = 2^{\chi/2}$. This allows us to simulate even systems with an amount of entanglement that grows faster than an area law, such as a Fermi surface or states after a quantum quench, efficiently and to high accuracy for very large system sizes. We have demonstrated the power of our method for both one-dimensional and quasi-2D systems and obtained results for up to one million lattice sites and effective bond dimensions of 10^{15}.

An interesting follow-up question is how to generalize the method to two-dimensional systems using projected entangled pair states (PEPS). A key point to be addressed is that in our one-dimensional algorithm, the canonical form serves a more central role than in conventional MPS simulations. While in the latter case the canonical form is primarily used to stabilize the method, in our case it is the key ingredient which allows us to express the dependence of the energy on a tensor as an eigenvalue problem in the first place. Therefore, generalizing the method to two dimensions requires either to understand how to solve the energy optimization problem without using a canonical form or consider more specialized two-dimensional tensor networks that have a canonical form.
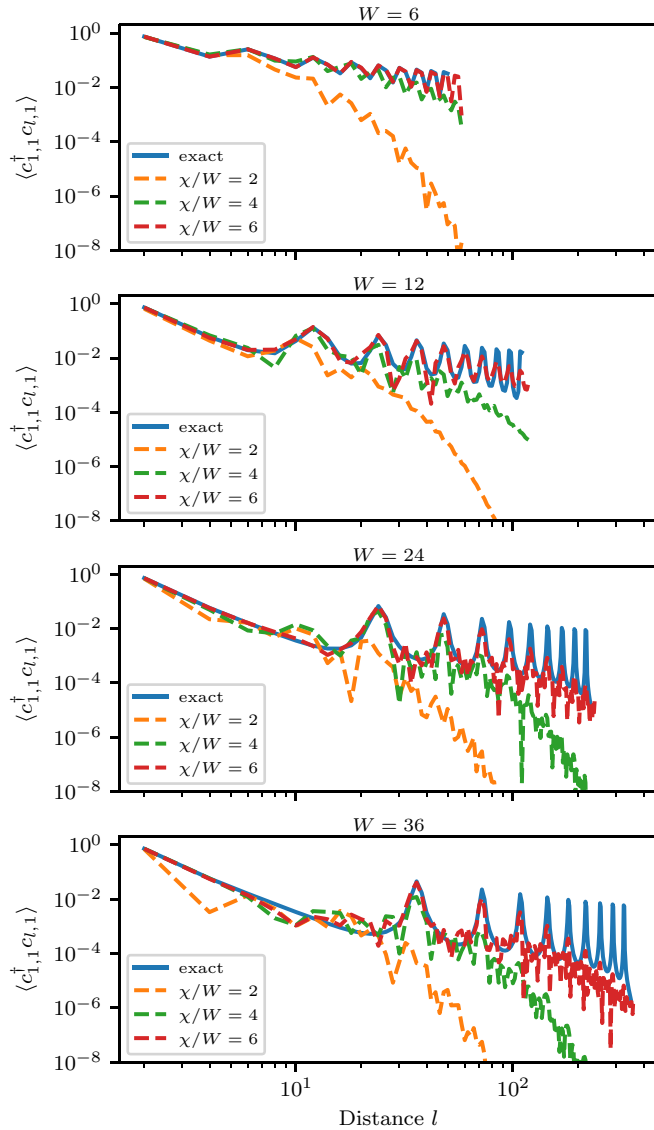
FIG. 14. Convergence of real-space correlations functions for the ground state of Eq. (77) in the limit of a square lattice, $t = t'$.

APPENDIX A: REAL-SPACE CORRELATIONS IN THE PRESENCE OF A FERMI SURFACE

In Fig. 14, we show the real-space correlation functions, analogous to Fig. 11, for the case of fermion hopping on the square lattice, i.e., $t = t'$. While these exhibit much more structure than the correlations in the honeycomb case, which is due to the interplay of finite Fermi momentum k_F and the quantization of momentum around the periodic direction of the cylinder, they still have an envelope that decays as a power law. Importantly, in contrast to the honeycomb case of Fig. 11, the system violates the area law and thus an accurate description requires a χ that grows faster than W. This can be seen in the data: while for $W = 6$, $\chi/W = 6$ is sufficient to accurately describe the correlations, for $W = 36$ the same χ/W cannot capture the correlations accurately beyond a distance of roughly 10^2.

APPENDIX B: COMPUTATION OF OVERLAPS OF GFMPS

In this Appendix, we describe how to efficiently evaluate the absolute value squared of the overlap of two GFMPS. We focus on this instead of the overlap itself, since the phase of the overlap is not fully determined by a Gaussian covariance matrix; to see this, note that the covariance matrix of two Gaussian states $|\phi\rangle$ and $e^{i\theta}|\phi\rangle$ is the same. For a more detailed discussion of this issue and possible solutions, see Ref. [46].

We begin by slightly generalizing the contraction formalism introduced in Sec. II B. In particular, Eq. (18) can be generalized to the contraction of two indices b, c on a given covariance matrix γ even when there is correlation between the two, i.e., $\gamma_{bc|bc} \neq 0$. This should be contrasted with the case of Eq. (18), where the input covariance matrix is the direct sum of two parts and the indices being contracted have no correlations. Starting from the covariance matrix $\gamma_{abc|abc}$, the CM after contracting $b \rhd c$ is given by

$$K = \gamma_{aa} + \gamma_{a|bc}\left[\gamma_{bc|bc} + \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}\right]^{-1}\gamma_{a|bc}^{\mathsf{T}}. \quad (B1)$$

Given two states ρ and σ with CMs γ_ρ and γ_σ, their overlap is $\mathrm{tr}(\rho\sigma) = 2^{-\chi/2}\sqrt{\det(\mathbb{1} - \gamma_\rho\gamma_\sigma)}$, where χ is the number of Majorana modes [47], cf. also Eq. (10). From this we can immediately infer that the change in normalization induced by contracting two indices $b \rhd c$ of a state with CM $\gamma_{abc|abc}$—which is equivalent to projecting bc on a maximally entangled state—is given by

$$\mathcal{N} = 2^{-\chi/2}\sqrt{\left|\det\left[\gamma_{bc|bc} + \begin{pmatrix} & \mathbb{1} \\ -\mathbb{1} & \end{pmatrix}\right]\right|}, \quad (B2)$$

with $\chi = |b| = |c|$ (this can be seen by multiplying with $\begin{pmatrix} & \mathbb{1} \\ -\mathbb{1} & \end{pmatrix}$ in the determinant); note that this exactly corresponds to the determinant of the inverse in the Schur complement in the contraction Eq. (B1). This normalization also holds for the case $|a| = 0$, i.e., when all indices are contracted.

In order to compute the overlap of two Gaussian states with covariance matrices γ and Γ, we can form the joint covariance matrix $\begin{pmatrix} \gamma & \\ & -\Gamma \end{pmatrix}$ and contract all indices using Eq. (B2); note that in that case $\mathcal{N}$ is independent of the direction of the contraction (i.e., replacing $\mathbb{1} \leftrightarrow -\mathbb{1}$). The minus sign reflects the fact that contracting as defined here involves a particle-hole transformation on Γ (corresponding to particle number conservation of the maximally entangled state ω onto which the bond is projected). To compute the overlap of two GFMPS (cf. Fig. 15) we can perform the same procedure iteratively by contracting the two GFMPS together from left to right. In each step, the change in normalization must be tracked according to Eq. (B2). Furthermore, the particle-hole transformation on the second MPS requires the direction of the contraction of its virtual indices to be reversed. A convenient way to carry out the procedure is to perform the following contractions:

(1) $L^{(1)} = (-\Gamma^1) \rhd \gamma^1$
(2) $L' = L^{(n)} \rhd \gamma^{(n+1)}$
(3) $L^{(n+1)} = (-\Gamma^{(n+1)}) \rhd L'$ and return to step 2.

Here, in slight abuse of previous notation, we use the symbol $\rhd$ to denote the (directional) contraction of two tensors (rather than indices), where the correct indices to contract are
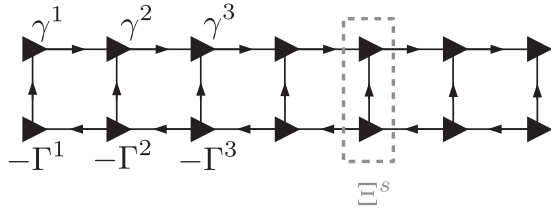
FIG. 15. Overlap of GFMPS. The overlap is obtained by contracting the GFMPS with tensors γ^s and the particle-hole transformed GFMPS with tensors $-\Gamma^s$. The contraction can be carried out efficiently by contracting, say, from left to right, in each step contracting first the upper and then the lower MPS tensor to a left boundary tensor. In each step, the opposite contraction order in the lower part must be observed.

specified in Fig. 15. In each contraction, including the very last one, the normalization factor $\mathcal{N}$ is tracked; the overall

normalization of the overlap is then given by the product of all normalizations. We denote this overlap as $\mathcal{N}(\gamma^s, \Gamma^s)$.

One can then obtain the normalized overlap of the two states as

$$|\langle\gamma|\Gamma\rangle|^2 = \frac{\mathcal{N}(\gamma^s, \Gamma^s)}{\sqrt{\mathcal{N}(\gamma^s, \gamma^s)\mathcal{N}(\Gamma^s, \Gamma^s)}}. \quad (B3)$$

The computation of the denominator is greatly simplified by transforming each GFMPS into the left- or right-canonical form. If the left-canonical gauge is chosen for all tensors, then $\gamma_{rr}^s = 0$, yielding a normalization $2^{-\chi/2}$ for the contraction of a bond with χ Majorana modes, and thus

$$\mathcal{N}(\gamma^s, \gamma^s) = \prod_i 2^{-\chi_{i,i+1}^\gamma/2}, \quad (B4)$$

with $\chi_{i,i+1}^\gamma$ the bond number at link $(i, i+1)$.

[1] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).
[2] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
[3] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[4] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[5] U. Schollwöck, Ann. Phys. **326**, 96 (2011).
[6] G. Sierra and M. A. Martin-Delgado, in *The Exact Renormalization Group–Proceedings of the Workshop*, edited by A. Krasnitz, Y. A. Kubyshin, R. Potting, and P. Sa (World Scientific, Singapore, 1999).
[7] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **67**, 3066 (1998).
[8] T. Nishino, K. Okunishi, Y. Hieida, N. Maeshima, and Y. Akutsu, Nucl. Phys. B **575**, 504 (2000).
[9] Y. Nishio, N. Maeshima, A. Gendiar, and T. Nishino, arXiv:cond-mat/0401115.
[10] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.
[11] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).
[12] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).
[13] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
[14] R. Orús, Ann. Phys. **349**, 117 (2014).
[15] R. Orús, Eur. Phys. J. B **87**, 280 (2014).
[16] J. C. Bridgeman and C. T. Chubb, J. Phys. A: Math. Theor. **50**, 223001 (2017).
[17] C. V. Kraus, N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. A **81**, 052338 (2010).
[18] G. Evenbly and G. Vidal, Phys. Rev. B **81**, 235102 (2010).
[19] N. Schuch, M. Wolf, and J. I. Cirac, in *Proceedings of the Conference on Quantum Information and Many Body Quantum Systems*, edited by M. Ericsson and S. Montangero (Edizioni della Normale, Pisa, 2008), p. 129.
[20] J. Dubail and N. Read, Phys. Rev. B **92**, 205307 (2015).
[21] J. Haegeman, T. J. Osborne, H. Verschelde, and F. Verstraete, Phys. Rev. Lett. **110**, 100402 (2013).
[22] M. T. Fishman and S. R. White, Phys. Rev. B **92**, 075132 (2015).
[23] G. Evenbly and S. R. White, Phys. Rev. Lett. **116**, 140403 (2016).

[24] J. Haegeman, B. Swingle, M. Walter, J. Cotler, G. Evenbly, and V. B. Scholz, Phys. Rev. X **8**, 011003 (2018).
[25] A. Jahn, M. Gluza, F. Pastawski, and J. Eisert, Sci. Adv. **5**, eaaw0092 (2019).
[26] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Phys. Rev. Lett. **107**, 070601 (2011).
[27] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Phys. Rev. B **94**, 165116 (2016).
[28] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, Rev. Mod. Phys. **78**, 275 (2006).
[29] D. J. Thouless and S. Kirkpatrick, J. Phys. C **14**, 235 (1981).
[30] C. H. Lewenkopf and E. R. Mucciolo, J. Comput. Electron. **12**, 203 (2013).
[31] S. Datta, Superlatt. Microstruct. **28**, 253 (2000).
[32] B. Gaury, J. Weston, M. Santin, M. Houzet, C. Groth, and X. Waintal, Phys. Rep. **534**, 1 (2014).
[33] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[34] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[35] M. Hastings, J. Stat. Mech. (2007) P08024.
[36] I. Arad, Z. Landau, U. Vazirani, and T. Vidick, Commun. Math. Phys. **356**, 65 (2017).
[37] Y. Huang, arXiv:1403.0327.
[38] S. R. White, Phys. Rep. **301**, 187 (1998).
[39] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).
[40] M. B. Hastings, Phys. Rev. B **76**, 035114 (2007).
[41] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 030504 (2008).
[42] A. Botero and B. Reznik, Phys. Rev. A **67**, 052311 (2003).
[43] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).
[44] I. Peschel, J. Phys. A: Math. Gen. **36**, L205 (2003).
[45] A. Botero and B. Reznik, Phys. Lett. A **331**, 39 (2004).
[46] S. Bravyi and D. Gosset, Commun. Math. Phys. **356**, 451 (2017).
[47] S. Bravyi, Quantum Inf. Comput. **5**, 216 (2005).
[48] M. Wimmer, ACM Trans. Math. Software (TOMS) **38**, 30 (2012).

[49] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[50] G. Giedke and J. I. Cirac, Phys. Rev. A **66**, 032316 (2002).

[51] J. Eisert, S. Scheel, and M. B. Plenio, Phys. Rev. Lett. **89**, 137903 (2002).

[52] J. Fiurášek, Phys. Rev. Lett. **89**, 137904 (2002).

[53] This can also be seen without calculation by noting that this amounts to a postselected teleportation protocol (i.e., attaching a maximally entangled state and projecting onto the maximally entangled state).

[54] S. R. White, Phys. Rev. B **72**, 180403(R) (2005).

[55] This can be shown by observing that Eq. (69) is equivalent to $c_k(t) = \sum O(t)_{kl} c_l(0)$ and thus $\dot{c}_k = 4 \sum H_{kl} c_l$, and on the other hand, comparing it with the evolution $\dot{c}_k = i[\mathcal{H}, c_k]$ of c_k in the Heisenberg picture.

[56] C. V. Kraus and J. I. Cirac, New J. Phys. **12**, 113004 (2010).

[57] I. P. McCulloch, arXiv:0804.2509.

[58] G. M. Crosswhite, A. C. Doherty, and G. Vidal, Phys. Rev. B **78**, 035116 (2008).

[59] L. Vanderstraeten, J. Haegeman, and F. Verstraete, SciPost Phys. Lect. Notes **7** (2019).

[60] G. Vidal, Phys. Rev. Lett. **98**, 070201 (2007).

[61] P. W. Anderson, Phys. Rev. **124**, 41 (1961).

[62] G. Toulouse, Phys. Rev. B **2**, 270 (1970).

[63] P. Schlottmann, Phys. Rev. B **22**, 613 (1980).

[64] V. Filyov, A. Tzvelik, and P. Wiegmann, Phys. Lett. A **81**, 175 (1981).

[65] S. Ghosh, P. Ribeiro, and M. Haque, J. Stat. Mech. (2014) P04011.

[66] M. M. Wolf, Phys. Rev. Lett. **96**, 010404 (2006).

[67] D. Gioev and I. Klich, Phys. Rev. Lett. **96**, 100503 (2006).

[68] T. Barthel, M.-C. Chung, and U. Schollwöck, Phys. Rev. A **74**, 022329 (2006).

[69] W. Li, L. Ding, R. Yu, T. Roscilde, and S. Haas, Phys. Rev. B **74**, 073103 (2006).

[70] See Ref. [72] for an MPS-based variational method where convergence can be guaranteed in certain cases.

[71] Y. V. Nazarov, Y. Nazarov, and Y. M. Blanter, *Quantum Transport: Introduction to Nanoscience* (Cambridge University Press, Cambridge, UK, 2009).

[72] Z. Landau, U. Vazirani, and T. Vidick, Nat. Phys. **11**, 566 (2015).