Variational methods for contracting projected entangled-pair states

Laurens Vanderstraeten,[1,*] Lander Burgelman [ORCID],[1] Boris Ponsioen,[2] Maarten Van Damme,[1] Bram Vanhecke,[1,3]
Philippe Corboz,[2] Jutho Haegeman [ORCID],[1] and Frank Verstraete[1]

[1]*Department of Physics and Astronomy, University of Ghent, Krijgslaan 281, 9000 Gent, Belgium*
[2]*Institute for Theoretical Physics and Delta Institute for Theoretical Physics, University of Amsterdam,*
Science Park 904, 1098 XH Amsterdam, The Netherlands
[3]*Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Wien, Austria*

The norms or expectation values of infinite projected entangled-pair states (PEPS) cannot be computed exactly, and approximation algorithms have to be applied. In the last years, many efficient algorithms have been devised—the corner transfer matrix renormalization group (CTMRG) and variational uniform matrix product state (VUMPS) algorithm are the most common—but it remains unclear whether they always lead to the same results. In this paper, we identify a subclass of PEPS for which we can reformulate the contraction as a variational problem that is algorithm independent. We use this variational feature to assess and compare the accuracy of CTMRG and VUMPS contractions. Moreover, we devise a new variational contraction scheme, which we can extend to compute general N-point correlation functions.

I. INTRODUCTION

Tensor networks provide excellent variational states for approximating strongly-correlated ground states of generic quantum lattice systems. This can be brought back to their special entanglement properties, which are characteristic for low-energy states of local Hamiltonians. The class of matrix product states (MPS) [1] and projected entangled-pair states (PEPS) [2] are the most versatile examples for simulating one-dimensional (1D) and two-dimensional (2D) systems, respectively. They have a built-in control parameter, the bond dimension D. For gapped systems, tensor networks are believed to yield essentially exact ground states as long as the bond dimension is chosen large enough [3], whereas for gapless systems, a finite-entanglement scaling gives access to the critical data [4–12].

For tensor networks to be useful for simulating realistic models, we need to be able to compute expectation values and perform variational optimizations efficiently. In the case of MPS, the density-matrix renormalization group [13] and its extensions [1] provide efficient numerical schemes that are used for simulating 1D quantum systems with high accuracy. For PEPS, the situation is less clear. First of all, the computation of expectation values can only be done approximately [14] and often requires large computational resources. Here a second control parameter enters, the environment bond dimension χ. There are different algorithms for performing this approximate evaluation (see below), and it is a priori not clear that they give identical results for the same value of χ. Secondly, the different algorithms for optimizing the PEPS tensors can give rise to different results: The most widely-used

PEPS algorithms rely on imaginary-time evolution with local truncations, where either the environment is approximated in a mean-field fashion (simple update) [15] or fully taken into account (full update) [16]. It is only by the advent of variational optimization methods for PEPS [17,18] that it was realized that even the full-update algorithm does not necessarily find the optimal PEPS tensors. As a result, it proved crucial to use PEPS results from variational optimization when extrapolating finite-D results for critical models [9,10,12].

In this paper, we show that under certain symmetry conditions on the PEPS tensors, the problem of contracting a PEPS can be turned into a variational problem itself. This implies that we can determine the approximate energy expectation value, as well as other observables, for a given value of the environment bond dimension χ, independently of the algorithm that is actually used for computing it. This solves some of the ambiguities in the practice of PEPS optimization, since it implies that we can formulate a variational principle for a PEPS class of states at a certain bond dimension by specifying the environment bond dimension χ. Moreover, this variational characterization allows us to use variational MPS algorithms for PEPS contractions and evaluating correlation functions. Finally, it allows us to compare the accuracy of different, possibly nonvariational, contraction schemes.

The paper is organized as follows. We first review (Sec. II) the different contraction algorithms that are currently used in PEPS simulations, and motivate the use of an algorithm-independent characterization of the environment. In the next section (Sec. III), we give symmetry constraints on the PEPS such that the transfer matrix becomes Hermitian. Then (Sec. IV) we show that this allows us to formulate a variational principle for the PEPS contraction, and we describe variational algorithms for finding the environment. We go on (Sec. V) with a new algorithm for computing two-point

*laurens.vanderstraeten@ugent.be

functions, using MPS techniques that resemble state-of-the-art methods for evaluating spectral functions in quantum spin chains. In Sec. VI, we benchmark these variational contraction and summation algorithms, and compare with other established methods.

II. REVIEW OF CONTRACTION ALGORITHMS

Let us first review the PEPS construction, and the most common contraction algorithms. A PEPS can be directly defined on an infinite lattice by repeating a unit cell of tensors. In the simplest case, the unit cell consists of just a single tensor A and the state can be represented diagrammatically as[1]

$$|\Psi(A)\rangle = \qquad . \qquad (1)$$

The four virtual indices of the PEPS tensors are contracted, whereas the physical index represents the physical degree of freedom. The state is thus parametrized by a single five-leg tensor A, where the bond dimension D of the four virtual indices can be increased to enlarge the variational class of states.

The first step in dealing with these infinite PEPSs consists of normalizing the state, i.e., we want to compute the norm

$$n(A, \bar{A}) = \langle \Psi(\bar{A}) | \Psi(A) \rangle \qquad (2)$$

directly in the thermodynamic limit. We can represent this norm as an infinite tensor network

$$n(A, \bar{A}) = \qquad , \qquad (3)$$

with the double-layer tensor O obtained by the contraction of the tensor A with its conjugate

$$\qquad . \qquad (4)$$

Similarly, a local expectation value is represented as

$$\qquad , \qquad (5)$$

---

[1]Here it is understood that the PEPS tensor network encodes the expansion coefficients of the state in the basis of spin configurations.

where the dashed lines indicates the location of a local operator squeezed between the two layers.

In practical PEPS algorithms, there are two approaches for computing such infinite tensor networks. The first one is based on the corner-transfer matrix renormalization group (CTMRG) [19–25], which finds a local environment in the form of

$$\qquad , \qquad (6)$$

where the C and T tensors[2] represent the infinite network surrounding this site; the dimension of the legs of these tensors is χ. The environment tensors are found numerically by iterating a real-space renormalization step until convergence: in each step (i) an O tensor is absorbed into the environment tensors and (ii) the bond dimension of the tensors is truncated to χ. There are different prescriptions for this truncation step, and they vary both in stability of the algorithm and accuracy of the results. The expectation value of a larger local operator is computed by enlarging the same environment in a trivial way, e.g.,

$$\qquad . \qquad (7)$$

The second approach relies on finding approximate MPS for representing the fixed point of the one-dimensional transfer matrix,

$$\mathcal{T} = \qquad . \qquad (8)$$

This fixed-point equation is written as

$$\approx \qquad , \qquad (9)$$

where χ is now the bond dimension of the boundary MPS. Different algorithms can be used to find the approximate fixed point such as the density-matrix renormalization group [26], a variant of the time-evolving block decimation [27] or the variational uniform MPS (VUMPS) algorithm [28–31]. Once such a boundary MPS is found from both directions parametrized by M and $\tilde{M}$ tensors,[3] the local environment is

---

[2]For generic PEPSs, the four C and T tensors are, in fact, different.

[3]In general there is no simple relation between M and $\tilde{M}$, but in many cases, the spatial symmetries of the PEPS allow to relate them in a simple way (see below).

found as

$$(10)$$

Again, larger environments can be found straightforwardly.

Evaluating nonlocal expectation values such as correlation functions can be done as well. For a two-point correlation function on the same row or column in the lattice, this is easily done by extending CTMRG environment,

$$(11)$$

or taking the boundary MPS

$$(12)$$

More general two-point correlation functions can be computed by extending boundary MPS to include corners, giving rise to corner-shaped environments [28] or by a summation routine based on the CTMRG scheme [17].

This review makes clear that there is a plethora of different types of environments and algorithms for contracting a given PEPS with a certain environment bond dimension χ. It is expected that they give different results for computing local expectation values, and it is not even clear that they all converge to the same result in the $\chi \to \infty$ limit. In the following sections, we aim at partially solving these ambiguities.

III. SYMMETRIES AND THE TRANSFER MATRIX

Imposing symmetry constraints on the PEPS tensors is a powerful tool for restricting the variational space and making algorithms more efficient. A first class of constraints involves on-site symmetries of the state: imposing that a PEPS wave function is invariant under a global on-site symmetry action can be imposed on the level of the local PEPS tensor. This way of constraining the PEPS tensor has been used successfully to simulate ground states with (Abelian) U(1) symmetries [32,33] and (non-Abelian) SU(2) [34–36] or SU(N) [37–39] symmetries.

Here we are interested in a second class of constraints, involving spatial symmetries of the PEPS wave function. Again, we can impose symmetry constraints on the level of the local PEPS tensor. One type of spatial symmetry that is often useful is rotation invariance of the PEPS tensor,

$$(13)$$

leading to a global state that is rotation invariant. A second spatial symmetry is reflection, which is typically combined with time-reversal symmetry,

$$(14)$$

where the bar denotes complex conjugation of the tensor elements. Note that an explicit breaking of reflection and time-reversal symmetry separately often leads to chiral wave functions [40,41]. The interplay between internal symmetries and spatial symmetries leads to nontrivial classes of PEPS wave functions [42].

Here we are interested in the latter symmetry, and in the relation with the transfer matrix. Indeed, if we take the Hermitian conjugate of the transfer matrix

$$(15)$$

we see that this is invariant when the symmetry condition in Eq. (14) is imposed on the A tensor. So this symmetry implies that the transfer matrix is Hermitian, but the Hermiticity requirement can be met by a more general condition. Indeed, under the condition

$$(16)$$

with U a unitary matrix and X an invertible matrix, the transfer matrix is hermitian by construction.

This hermiticity condition can be extended to larger unit cells, where we require that the transfer matrix of a full unit cell is hermitian. This requirement can be illustrated in the case of a PEPS with a two-by-two unit cell,

$$|\Psi(A_i)\rangle =$$ $$.\qquad(17)$$

Indeed, if we require

$$(18)$$

$$(19)$$

and similar conditions on A_3 and A_4, then the two-row transfer matrix

$$\mathcal{T} = \qquad\qquad\qquad\qquad (20)$$

is Hermitian.

IV. VARIATIONAL CONTRACTION

In the previous section, we have explained under what conditions the PEPS transfer matrix is Hermitian. In this section, we show why this feature is extremely useful for characterizing and finding a boundary MPS.

A. Single-row transfer matrix

The norm of a PEPS $n(A, \bar{A})$ can be interpreted as an infinite product of transfer matrices and, therefore, reduces to an infinite power of the leading eigenvalue. The corresponding eigenvalue equation is given by

$$\mathcal{T}|\Psi\rangle = \Lambda|\Psi\rangle, \qquad (21)$$

where Λ scales exponentially with the number of sites in the x direction, i.e., $\Lambda \propto \lambda^{N_x}$. In analogy with transfer-matrix approaches in statistical mechanics [29], we can associate a free-energy density to the norm of the infinite PEPS as

$$f(A, \bar{A}) = -\lim_{N_x, N_y \to \infty} \frac{1}{N_x N_y} \ln n(A, \bar{A})$$
$$= -\ln \lambda. \qquad (22)$$

Since the transfer matrix is a Hermitian operator, the above eigenvalue equation can be reformulated as a variational problem

$$|\Psi\rangle = \arg\min_{|\Psi\rangle} \left(-\frac{1}{N_x} \ln \left(\frac{\langle\Psi|\mathcal{T}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \right) \right). \qquad (23)$$

We can now solve the eigenvalue equation variationally by approximating the fixed point as an infinite MPS, parametrized by a single tensor M

$$|\Psi_M\rangle = \cdots \qquad\qquad\qquad\qquad \cdots. \qquad (24)$$

The variational characterization of the MPS tensor is given by the following variational optimization problem:

$$M = \arg\min_M \left(-\frac{1}{N_x} \ln \Lambda(M, \bar{M}) \right) \qquad (25)$$

with

$$\Lambda(M, \bar{M}) = \left(\frac{\langle\Psi_{\bar{M}}|\mathcal{T}|\Psi_M\rangle}{\langle\Psi_{\bar{M}}|\Psi_M\rangle} \right). \qquad (26)$$

Diagrammatically, the numerator is represented as an infinite channel

$$\langle\Psi_{\bar{M}}|T|\Psi_M\rangle = \cdots \qquad\qquad \cdots, \qquad (27)$$

which scales as $\Lambda \propto \lambda^{N_x}$. The value λ is determined as the leading eigenvalue of the channel operator

$$\lambda = \rho_{\max} \left(\qquad\qquad \right). \qquad (28)$$

Given this variational characterization of the boundary MPS, we can use gradient-based optimization techniques to find it. The gradient of the objective function in Eq. (25) is given by

$$g = 2\partial_{\bar{M}} \left(-\frac{1}{N_x} \ln \Lambda(M, \bar{M}) \right)$$
$$= -\frac{2}{N_x} \frac{1}{\Lambda} \frac{\langle\partial_{\bar{M}}\Psi_{\bar{M}}|(\mathcal{T} - \Lambda)|\Psi_M\rangle}{\langle\Psi_{\bar{M}}|\Psi_M\rangle}. \qquad (29)$$

Diagrammatically, the gradient is given by

$$g = -2 \left(\frac{1}{\lambda} \cdots \qquad\qquad \cdots \right.$$
$$\left. - \cdots \qquad\qquad \cdots \right), \qquad (30)$$

assuming the MPS is normalized. We can now use efficient optimization techniques over MPS manifolds [43] to optimize the boundary MPS, exploiting the gauge degrees of freedom in the MPS ansatz.

Instead of using gradient-based techniques, we can reformulate the variational characterization in Eq. (25) as a fixed-point equation, and, starting from a random MPS, iterate this equation until it converges to the variational optimum. This procedure is known as the VUMPS algorithm [30,44]. First, we bring the MPS into center-gauged form as

$$|\Psi_M\rangle = \cdots \qquad\qquad\qquad\qquad \cdots, \qquad (31)$$

where M^l and M^r are related to the original MPS tensor via a gauge transform and satisfy the left- and right-canonical

gauge condition

$$\tag{32}$$

and are related via the matrix C

$$\tag{33}$$

In this MPS representation, the condition of a vanishing gradient is equivalent to the fixed-point equation for M^c

$$\tag{34}$$

and a similar equation for the matrix C is easily derived:

$$\tag{35}$$

These equations [Eqs. (32)–(35)] fully characterize the variationally optimal boundary MPS; in Refs. [30,31] the iterative scheme is explained in detail.

B. Two-row transfer matrix

Suppose we have a PEPS with a two-by-two unit cell (the more general case of a two-by-N unit cell is treated similarly), giving rise to a two-row transfer matrix $\mathcal{T} = \mathcal{T}_2 \mathcal{T}_1$. As explained above, we parametrize the PEPS such that the two-row transfer matrix is Hermitian, i.e., we have $\mathcal{T}_2^\dagger = \mathcal{T}_1$. We characterize the fixed point of this two-row transfer matrix as

$$\mathcal{T}_1 |\Psi_1\rangle = \Lambda_1 |\Psi_2\rangle,$$
$$\mathcal{T}_2 |\Psi_2\rangle = \Lambda_2 |\Psi_1\rangle. \tag{36}$$

in terms of two normalized states $|\Psi_1\rangle$ and $|\Psi_2\rangle$. Under the Hermiticity condition, $|\Psi_1\rangle$ can be characterized variationally as the state which optimizes the eigenvalue of the two-row transfer matrix

$$|\Psi_1\rangle = \arg\min_{|\Psi_1\rangle}\left(-\frac{1}{N_x}\ln\frac{\langle\Psi_1|\mathcal{T}_2\mathcal{T}_1|\Psi_1\rangle}{\langle\Psi_1|\Psi_1\rangle}\right). \tag{37}$$

Here, however, we can go one step further. The eigenvectors $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are, in fact, the left and right leading singular vectors of the matrix $\mathcal{T}_1$, and we can use the variational characterization for singular vectors to reformulate the optimality

condition as

$$|\Psi_1\rangle, |\Psi_2\rangle = \arg\min_{|\Psi_1\rangle, |\Psi_2\rangle}\left(-\frac{1}{N_x}\ln\Lambda\right), \tag{38}$$

with

$$\Lambda = \frac{\langle\Psi_2|\mathcal{T}_1|\Psi_1\rangle\langle\Psi_1|\mathcal{T}_2|\Psi_2\rangle}{\langle\Psi_1|\Psi_1\rangle\langle\Psi_2|\Psi_2\rangle}. \tag{39}$$

The Hermiticity condition $\mathcal{T}_2^\dagger = \mathcal{T}_1$ ensures that this cost function is real.

We can now approximate the fixed point as a set of two MPSs, which we also take to have a two-site unit cell and are, therefore, each described by two MPS tensors

$$|\Psi_1\rangle \approx |\Psi_{M_{11},M_{12}}\rangle, \quad |\Psi_2\rangle \approx |\Psi_{M_{21},M_{22}}\rangle. \tag{40}$$

The above cost function can be expressed as (with $\Lambda_1 = \bar{\Lambda}_2$)

$$\Lambda = \Lambda_1\Lambda_2, \tag{41}$$

with

$$\tag{42}$$

Again, we can use optimization techniques over MPS manifolds to optimize these two MPSs efficiently.

Analoguously to the above case of a one-row transfer matrix, we can switch to a center-gauge representation for the MPSs and derive fixed-point equations for the MPS tensors. Here, we find

$$\tag{43}$$

for the center-site MPS tensors (the other three equations follow similarly). These equations constitute the multisite version of the VUMPS algorithm, which was explained in detail in Ref. [45]. Starting from a set of random MPSs, the VUMPS fixed-point equations can be iterated in order to find the optimal MPSs.

C. More than two rows

Let us now consider the case of a three-row transfer matrix, with the Hermiticity conditions $\mathcal{T} = \mathcal{T}_1^\dagger \mathcal{T}_2 \mathcal{T}_1$ and $\mathcal{T}_2^\dagger = \mathcal{T}_2$. The fixed point is given by

$$\mathcal{T}_1 |\Psi_1\rangle = \Lambda_1 |\Psi_2\rangle,$$
$$\mathcal{T}_2 |\Psi_2\rangle = \Lambda_2 |\Psi_3\rangle,$$
$$\mathcal{T}_1^\dagger |\Psi_3\rangle = \Lambda_3 |\Psi_1\rangle, \tag{44}$$

and the state $|\Psi_1\rangle$ can be characterized variationally as

$$|\Psi_1\rangle = \arg\min_{|\Psi_1\rangle}\left(-\frac{1}{N_x}\ln\frac{\langle\Psi_1|T_1^\dagger T_2 T_1|\Psi_1\rangle}{\langle\Psi_1|\Psi_1\rangle}\right). \quad (45)$$

Here, however, we *cannot* take the step of reformulating this as a variational principle where we optimize all layers simultaneously,

$$|\Psi_1\rangle,|\Psi_2\rangle,|\Psi_3\rangle \neq \arg\min_{|\Psi_1\rangle,|\Psi_2\rangle,|\Psi_3\rangle}\left(-\frac{1}{N_x}\ln\Lambda\right) \quad (46)$$

with

$$\Lambda = \frac{\langle\Psi_2|\mathcal{T}_1|\Psi_1\rangle\langle\Psi_3|\mathcal{T}_2|\Psi_2\rangle\langle\Psi_1|\mathcal{T}_1^\dagger|\Psi_3\rangle}{\langle\Psi_1|\Psi_1\rangle\langle\Psi_2|\Psi_2\rangle\langle\Psi_3|\Psi_3\rangle}, \quad (47)$$

as this cost function is not real, since it is not guaranteed that $|\Psi_3\rangle$ and $|\Psi_2\rangle$ are proportional.

In order to still obtain a boundary MPS algorithm, we instead start from the set of equations in Eq. (44). Assuming the first state can be approximated by an MPS, $|\Psi_1\rangle \approx |\Psi_{\{M_1\}}\rangle$, we can find an MPS for the second state by applying the operator $\mathcal{T}_1$ to $|\Psi_{\{M_1\}}\rangle$ and approximating the result by an MPS with smaller bond dimension. This approximation $|\Psi_{\{M_2\}}\rangle$ can, again, be characterized by a variational principle, where we now optimize the (normalized) fidelity

$$\{M_2\} = \arg\max_{\{M_2\}}\frac{|\langle\Psi_{\{M_2\}}|\mathcal{T}_1|\Psi_{\{M_1\}}\rangle|^2}{\langle\Psi_{\{M_2\}}|\Psi_{\{M_2\}}\rangle}, \quad (48)$$

and we have similar equations for the $\{M_1\}$ and $\{M_3\}$ tensors. In Ref. [46], it was shown that these optimization problems can be reformulated in center-gauge representation to obtain fixed-point equations characterizing each set of MPS tensors $\{M_i\}$. This naturally leads to a variational boundary MPS algorithm consisting of a power method with a variational optimization of the individual fidelities in each step.

If one instead solves all these fixed-point equations simultaneously by optimizing over all MPS at once, this gives rise to a single set of equations which is similar to the ones we have given for one- and two-row transfer matrices [45]. Thus, by resorting to a method which optimizes all eigenvalue equations [Eq. (44)] simultaneously, we essentially find the VUMPS fixed-point equations that would follow had Eq. (46) been a valid variational principle. Given that this is not the case, and there is no variational principle underlying those fixed point equations, we expect that iterating the multisite VUMPS fixed-point equations will, in general, not always work for transfer matrices with three or more rows, even if the total transfer operator is Hermitian. In Appendix B, we provide a detailed diagnostic of this issue.

We conclude by noting that the computational complexity of this multisite VUMPS algorithm scales linearly in L_x and L_y, where L_x is the MPS unit-cell size and L_y is the number of rows in the transfer matrix.

V. SUMMATION OF TWO-POINT FUNCTIONS

Besides the norm and local expectation values, nonlocal two-point functions—or N-point functions, in general—are often relevant observables of a given PEPS wave function. In Ref. [28] an extension of the boundary-MPS approach

was introduced for computing the most general two-point and three-point functions, and in Ref. [17], the CTMRG approach was extended to perform the same task; finally, in Ref. [47], a method was introduced for summing two-point functions using generating functions. In this section, we propose a different scheme that works for generic N-point functions by translating standard MPS algorithms to this context.

A. Static structure factor

Let us focus on a generic momentum-resolved two-point function, or static structure factor, of the form

$$s(\vec{k}) = \sum_{mn} e^{i(k_x m + k_y n)}\langle\Psi(A)|S_{m,n}S_{0,0}|\Psi(A)\rangle, \quad (49)$$

with $S_{m,n}$ a local operator acting on site (m, n), which requires resumming an infinite number of two-point correlators. We assume that we have found a normalized boundary MPS, which we can bring into center-gauged form

$$ \quad (50)$$

and that we have normalized the PEPS.

The first contribution to the structure factor is the one where both operators act on the same site, which can be easily computed as a local expectation value

$$ \quad (51)$$

where, as before, the dashed lines indicate the presence of the local operators. We can introduce the left- and right fixed points of the channel operator,

$$ \quad (52)$$

and

$$ \quad (53)$$

such that this expectation value is given by the network contraction

$$ \quad (54)$$

Summing contributions where the two operators live on the same row of the lattice is also straightforward. Indeed, all these contributions can be grouped under an expression of the

form

$$\mathrm{e}^{ik_x} \sum_n \left[\cdots \right]. \quad (55)$$

The infinite series can be taken explicitly, and we find

$$s_x(k_x) = \mathrm{e}^{ik_x} \left[\cdots \right]^{-1} \left[\cdots \right] + \mathrm{H.c.} \quad (56)$$

Summing up contributions where the two operators live on different lines is less straightforward. Indeed, these all have the form

$$(57)$$

with a power of the transfer matrix in between the two operators. If we call the state

$$|\Psi_M(S_i)\rangle = \cdots, \quad (58)$$

where the operator acts on site i, the above diagram can be denoted as

$$\langle \Psi_M(S_j) | \mathcal{T}^{n-1} | \Psi_M(S_i) \rangle, \quad (59)$$

with n the vertical distance between the two operators. Summing up all these different contributions amounts to

$$s_{xy}(\vec{k}) = \sum_{mn} \mathrm{e}^{i(k_x m + k_y n)} \langle \Psi_M(S_m) | \mathcal{T}^{n-1} | \Psi_M(S_0) \rangle$$

$$= \sum_m \mathrm{e}^{ik_x m} \langle \Psi_M(S_m) | \mathrm{e}^{ik_y} (1 - \mathrm{e}^{ik_y} \mathcal{T})^{-1} | \Psi_M(S_0) \rangle. \quad (60)$$

This expression is very similar to a frequency-resolved dynamical correlation function for a 1D spin chain, typically of

the form

$$S(q, \omega) = \int dt\, \mathrm{e}^{i\omega t} \sum_m \mathrm{e}^{iqm} \langle \Psi_0 | S_m(t) S_0(0) | \Psi_0 \rangle$$

$$= \sum_m \mathrm{e}^{iqm} \langle \Psi_0 | S_m (\omega - (H - E_0))^{-1} S_0 | \Psi_0 \rangle, \quad (61)$$

where $|\Psi_0\rangle$ would be the ground state of the spin-chain Hamiltonian H with ground-state energy E_0. Many different approaches exist in the MPS literature to compute the latter overlap explicitly, including Lanczos methods [48,49] the correction vector approach [50–53], Chebyshev expansion [54], etc. A more straightforward approach consists of time evolving the state according to the spin-chain Hamiltonian, and transforming the time-dependent correlator back to frequency space [55–59]. Here, in order to evaluate Eq. (60), we take the discretized version of the latter approach, i.e., we take the state $|\Psi_M(O_0)\rangle$ and sequentially apply the transfer matrix. In order to keep the problem tractable, we represent the state as a window MPS.

A window MPS [60–62] is an MPS with a finite window of nontranslation-invariant tensors, embedded in a translation-invariant MPS to the left and the right of this window. Such an MPS is given by

$$|\Psi_M(N_i)\rangle = \cdots \qquad \cdots \quad (62)$$

and represents a local perturbation of the infinite MPS.

Approximating the initial state as a finite-window MPS can be done in a simple variational step. Indeed, we can approximate it as a window MPS with a single site as

$$|\Psi_M(S_0)\rangle \approx \cdots \qquad \cdots, \quad (63)$$

where the tensor N_1 is found by solving the variational problem

$$N_1^{\mathrm{opt}} = \arg\max_{N_1} \frac{|\langle \Psi_M(N_1) | \Psi_M(S_0) \rangle|^2}{\langle \Psi_M(N_1) | \Psi_M(N_1) \rangle}. \quad (64)$$

Working in the center gauge reduces the norm to the identity

$$\langle \Psi_M(N_1) | \Psi_M(N_1) \rangle = \left(\cdots \right), \quad (65)$$

so that differentiating the above variational objective function gives the solution for N_1,

$$N_1^{\mathrm{opt}} = \cdots = \cdots. \quad (66)$$

We can continue to represent the state

$$\mathcal{T}^n | \Psi_M(S_0) \rangle \quad (67)$$

as a finite-window MPS by iterating the same procedure. In every step, we start from a window MPS with tensors $\tilde{N}_i$, apply the transfer matrix and approximate the state again as a window MPS with new tensors N_i. Every approximation step is done variationally, i.e., we optimize the objective function

$$N_i^{\text{opt}} = \arg\max_{N_i} \frac{|\langle \Psi_M(N_i) | \mathcal{T} | \Psi_M(\tilde{N}_i) \rangle|^2}{\langle \Psi_M(N_i) | \Psi_M(N_i) \rangle}. \qquad (68)$$

We can efficiently solve this optimization problem by a sweeping algorithm, where we sequentially optimize the window tensors separately. Optimizing a single tensor N_i is done by bringing the window MPS in center gauge around site i, and optimizing as we did above. After we have found a new window MPS representing the state (67), we can evaluate all contributions to $s_{xy}(\vec{k})$ from that row, and proceed to the next row.

In the case of unitary time evolution according to a spin-chain Hamiltonian, one is severely limited due to the growth of entanglement in the time-evolved state. Here this is not the case, since the transfer matrix is an operator with a spectral gap[4]; as a result, the evolved state will converge again to the fixed-point MPS after a number of iterations. At that point, the contributions to the structure factor have become negligible, and we can terminate the algorithm.

The size of the window-MPS is a tunable parameter: We expect that the effect of applying the operator will grow in size as layers of the transfer matrix are applied, very similar to the spreading of correlations in the unitary real-time evolution of a quantum spin chain. In Sec. VI we will investigate the effect of the window size on the accuracy.

B. Energy gradient

The motivation behind the first PEPS summation algorithms [17,18] was finding an efficient method for evaluating the gradient of the PEPS energy objective function, which can be used in gradient-based PEPS optimization routines. In Appendix A, we explain in detail how our new summation scheme can be extended for evaluating the gradient of a plaquette interaction term.

We should note that, besides summation schemes, the idea of backwards differentiation can be used to evaluate the energy gradient [33,63] efficiently. One important advantage of the latter approach is that it computes the gradient of the energy function at a given value of environment χ, and can therefore be used to perform a gradient optimization of a PEPS with a fixed value of χ. Such a finite-χ optimization might be interesting to study finite-χ effects; these can be used for simulating critical systems, where a finite-χ scaling behavior can be expected [12].

In contrast, the approximate summation schemes do not exactly provide the gradient for a given value of χ: only in the infinite-χ limit can one expect that the energy and gradient

---

[4]The gap in the transfer matrix is directly related to the correlation length of the PEPS. The latter is finite for generic variational PEPS, but there are fine-tuned PEPS wave functions that exhibit critical correlations. For the latter PEPSs, a summation scheme is expected to converge more slowly.
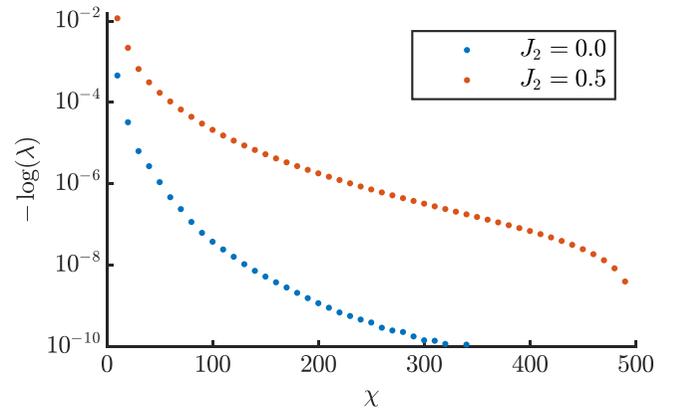
FIG. 1. The log-norm of the PEPS wave function at $J_2 = 0$ (blue) and $J_2 = 0.5$ (red) as a function of boundary MPS bond dimension χ. We have rescaled the PEPS tensor such that $\lambda_{\chi=500} = 1$ exactly, and do not show results below a cutoff of 10^{-10} because they contain numerical noise.

evaluation become fully compatible. Therefore the summation approach can only be used in a PEPS optimization in the infinite-χ regime—i.e., the regime where χ is chosen large enough such that finite-χ effects are negligible.

VI. BENCHMARKS

In order to benchmark and compare different contraction schemes, we take optimized PEPS tensors for the J_1-J_2 model on the square lattice, after a sublattice rotation with $J_2 = 0$ and $J_2 = 1/2$, each time with bond dimension $D = 5$. After a sublattice rotation, we can represent the ground state by a PEPS with a single-site unit cell. We have imposed both rotation invariance and reflection/time-reversal symmetry [Eqs. (13) and (14)] on the PEPS tensors. The latter symmetry implies that we can use our variational contraction schemes.

We have optimized these PEPS tensors by a gradient-based quasi-Newton optimization, where we have applied the gradient evaluation from Sec. A. We find energies $e = -0.669376$ and $e = -0.495912$ respectively, which are comparable to the results in Ref. [33].

A. Variational contraction

First of all, we explicitly demonstrate the variational nature of the contraction by computing the leading eigenvalue of the PEPS transfer matrix, obtained using an optimized boundary MPS with increasing bond dimension χ. In Fig. 1, one immediately observes that $f = -\ln(\lambda)$ monotonically decreases as the bond dimension is increased. The value for the unfrustrated PEPS ($J_2 = 0$) decreases significantly faster than for the frustrated case ($J_2 = 1/2$), which points to the fact that the frustrated ground state is more correlated than the unfrustrated one and therefore requires a larger value of χ for the same accuracy.

Next, we assess the performance of a direct optimization of the variational cost function in Eq. (25) by comparing with the VUMPS algorithm. In Fig. 2, we plot the norm of the gradient [Eq. (30)] after each iteration as a function of the wall time. For the VUMPS algorithm, each iteration consists of (i)
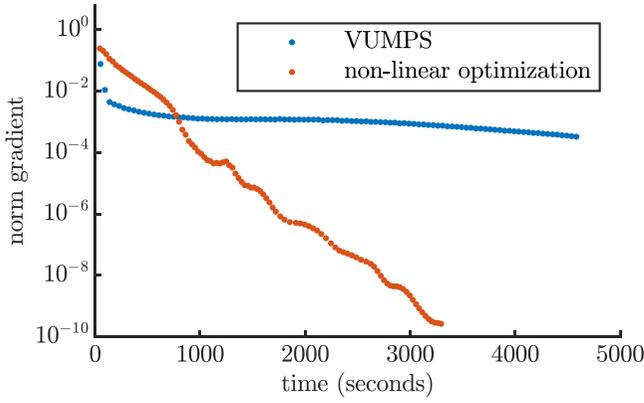
FIG. 2. The convergence of the boundary MPS with the VUMPS algorithm and by direct variational optimization. We plot the norm of the gradient in Eq. (30) for each iteration, as a function of the wall time.

FIG. 3. The error in evaluating the local energy of the PEPS as a function of the norm of the gradient during the process of optimizing the boundary MPS; we plot this for three different values of χ for the $J_2 = 1/2$ PEPS. The error is defined as the absolute difference with the fully optimized boundary MPS.

computing left- and right environments [Eqs. (52) and (53)], (ii) solving the eigenvalue equations for M^c and C [Eqs. (34) and (35)], and (iii) finding new M^l and M^r; the two first steps are the most expensive ones. For the direct optimization, we use a quasi-Newton scheme for optimizing over isometric tensor networks [43]; each iteration consists of an evaluation of the cost function and the gradient, which again requires computing the left- and right environments [Eqs. (52) and (53)]. As compared to VUMPS, the direct optimization strategy does not require solving the eigenvalue equations for M^c and C, so each iteration is faster by roughly a factor of 1.5 (the eigenvalue equation for C is not that costly, typically). As VUMPS is a fixed-point method, it can take larger steps in each iteration, allowing to converge much faster initially than a direct-optimization strategy. In the tail of the optimization, however, the manifold optimization clearly outperforms VUMPS fixed-point iterations, and converges more quickly. A hybrid strategy that switches between the two algorithms would be optimal.

One might wonder how well the boundary MPS should be converged in order to obtain accurate local PEPS expectation values. In Fig. 3, we have plotted the error in the energy expectation value for the J_1-J_2 Hamiltonian during the process of optimizing the boundary MPS. From this plot, it is clear that pushing the norm of the gradient to very small values is often not necessary in practice.

B. Comparing boundary MPS and CTMRG

We can now compare the accuracy of boundary-MPS and CTMRG contractions for evaluating local expectation values of PEPS wave functions. For the boundary MPS, we use a direct variational optimization as outlined above. For CTMRG, due to the symmetries of the PEPS tensor we can choose between the directional CMTRG [25] and the symmetric CTMRG [19]. Here we choose the former, because it can be applied more generally. Both schemes are slightly different, also for symmetric tensors, because in the directional scheme the growing of the environment is done in an asymmetric way and the projectors are constructed from a product of C tensors.
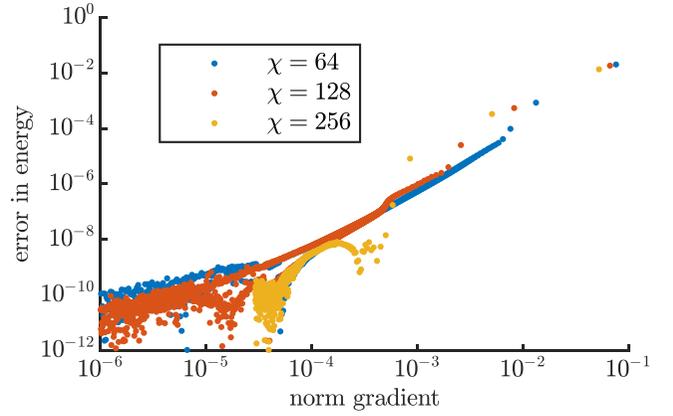
Nonetheless, we find that these differences are very small, and that both give rise to the same expectation values.

In Fig. 4, we have plotted the energy expectation value for the two optimized PEPSs as a function of the environment bond dimension χ, evaluated with a boundary-MPS and a CTMRG environment. For a nearest-neighbour Hamiltonian ($J_2 = 0$) the energy can be evaluated by a single-row boundary-MPS or a one-by-two CTMRG environment; in the top panel of Fig. 4, we observe that both approaches yield the same values. For a next-nearest-neighbour Hamiltonian ($J_2 = 1/2$) evaluating the energy from the boundary MPS requires a two-step process [Eq. (A4)] or a two-by-two CTMRG environment [Eq. (7)]. For small values of χ the boundary-MPS contraction is slightly more precise, but both methods clearly converge to the same value for larger χ.

Note that the energy converges a lot slower for the frustrated PEPS, and therefore requires a larger value of χ for the same accuracy. This is in agreement with the result in Fig. 1.

C. Structure factor

Finally, we benchmark our new summation scheme based on the window-MPS approach to compute the static spin structure factor of the PEPS wave function

$$s_\alpha(\vec{k}) = \sum_{x,y} e^{i(k_x x + k_y y)} \langle \Psi(A)|S^\alpha_{(x,y)} S^\alpha_{(0,0)}|\Psi(A)\rangle. \qquad (69)$$

The window-MPS approach has two control parameters, the MPS bond dimension and the window size. In Fig. 5, we plot the value of the structure factor as a function of the window size, for five different values of χ. We observe that a larger value of N is needed to converge the result as χ is lower. This can be understood from the fact that a larger-χ MPS contains longer-range correlations, and that a local operation can influence the state over a larger region—therefore, a small window can represent a larger perturbation when its bond dimension is higher. We also observe that the result for small χ does not converge to the correct result, as the initial boundary MPS does not carry the correct correlations—this is

FIG. 5. The value of the structure factor as a function of window size, for different values of χ.

factor is not a variational quantity itself, in contrast to the norm of the PEPS wave function in Fig. 1. The window-MPS approach, which exploits variational principles through its evaluation of all contributions to the structure factor, is clearly the most accurate; for summing the contributions of each row, however, it requires us to perform a number of optimization sweeps. The CTMRG summation is less accurate, but only comes with little extra computational cost compared to the

FIG. 4. The value for the energy expectation value as a function of χ, computed with the variational MPS and CTMRG environments for $J_2 = 0$ (top) and $1/2$ (middle). We also plot the relative error of the energy expectation value at $J_2 = 1/2$ with respect to the converged value at $\chi = 500$ (bottom).

analogous to simulating the time evolution of a local operator on a ground-state MPS with a too small bond dimension.

We can now compare our new summation scheme with the corner-environment approach [28] and CTMRG summation [17,64]. Figure 6 shows that all three approaches converge to the same result, but with a clear difference in accuracy for smaller values of χ. We observe that all schemes approach the converged value from below, but we have no good a priori reason to expect that this is always the case—the window-MPS scheme takes variational steps, but the summed-up structure

FIG. 6. The value of the structure factor as a function of χ, computed with the window-MPS approach with window size $N = 10$ (blue), the CTMRG resummation scheme (red) and the channel environments (yellow). $J_2 = 0$ (top) and $1/2$ (bottom).

standard CTMRG scheme (one needs to keep track of separate environment tensors for the norm and the summed operator terms). The corner environment, finally, is the least accurate, but evaluating the structure factor just requires contracting a single one-dimensional tensor network and, therefore, comes at almost no cost on top of a boundary-MPS contraction. Choosing the summation scheme, therefore, depends a lot on the application, the required accuracy and the available computational resources.

VII. DISCUSSION AND OUTLOOK

In this paper, we have identified a subclass of PEPS wave functions for which the transfer matrix is a Hermitian operator. This has allowed us to formulate the contraction as a variational problem, yielding an optimality condition that is independent of the algorithm that is actually used in practice. We have used this feature to compare state-of-the-art contraction algorithms such as the CTMRG and VUMPS algorithms, showing that they give comparable results for e.g. evaluating the energy of a PEPS. In addition, we have formulated a new scheme for computing general N-point functions, which is significantly more accurate as compared to existing alternative schemes. It would be instructive to compare our scheme to the generating-function approach of Ref. [47].

We believe that our new scheme will prove very useful in future PEPS methods and applications that go beyond ground-state properties. Indeed, a recent method for computing excitation spectra on top of a PEPS [28,64–66] relies heavily on summing two- and three-point functions. In addition, the time-dependent variational principle for uniform time evolution will require summing up two-point functions: the tangent vector that dictates the TDVP flow equations [31] takes on a similar form as the energy gradient (see Appendix A). Finally, the idea of using window MPS in PEPS calculations will prove useful to simulate local dynamics against a uniform background—simulating the response of the action of a local operator on the ground state for obtaining spectral functions.

For the Hermitian subclass of PEPS, we can unambiguously define the variationally optimal PEPS at a certain bond dimension D and contracted with an environment bond dimension χ. In Ref. [12], this feature was exploited to formulate a scaling hypothesis in terms of a single effective length scale that results from both finite-D and finite-χ effects.

Given the power of this subclass of PEPS, the question naturally arises under what conditions such a restriction can safely be imposed on the PEPS ansatz. One important restriction concerns the fact that the transfer matrix only has real eigenvalues. In the case of MPS, the eigenvalues of the transfer matrix are directly related to the wavevector of the dominant correlations in the state and, indirectly, to the minima of the quasiparticle dispersion relation on top of this state [67]. This implies that a ground state with dominant incommensurate correlations is expected to be poorly described by our subclass of PEPS—critical states with an incommensurate filling would satisfy this condition. It would be very interesting to track the limitations of the Hermitian subclass of PEPS for these systems, and what this implies for the contraction of their PEPS ground-state approximations.

In this paper, we have only looked at PEPS on a square lattice, and it would be interesting to generalize the construction to other lattices such as the triangular or the kagome lattice. Also, we would like to investigate the case of highly nontrivial PEPS unit cells, which represent, e.g., stripe phases in the Hubbard model with a commensurate filling [68,69] or complicated magnetization patterns [70] in spin systems in an external magnetic field.

APPENDIX A: THE ENERGY GRADIENT AS A TWO-POINT FUNCTION

In this Appendix, we write out the expressions for evaluating the energy and the gradient for a given PEPS tensor, based on the summation scheme from Sec. V. We consider plaquette interactions on the square lattice. The energy expectation value is then given by

$$e = \quad , \qquad (A1)$$

but evaluating this expression requires computing the fixed points of the two-row channel. We propose a different contraction technique, based on the window-MPS approach. We split up the Hamiltonian into two parts, and apply a first layer of the above diagram and approximate it by introducing a two-site tensor P as

$$\approx \qquad . \qquad (A2)$$

Finding the optimal tensor P is straightforward since the cost function is linear in P. We find

$$\qquad (A3)$$

The energy expectation value is then obtained as

$$e = \qquad . \qquad (A4)$$

The gradient is then computed in a similar fashion as the structure factor. Indeed, the gradient for the energy expectation value

$$g = 2\partial_{\bar{A}} \frac{\langle \Psi(\bar{A})|H|\Psi(A)\rangle}{\langle \Psi(\bar{A})|\Psi(A)\rangle} \tag{A5}$$

$$= 2 \frac{\langle \partial_{\bar{A}}\Psi(\bar{A})|H - e(A,\bar{A})|\Psi(A)\rangle}{\langle \Psi(\bar{A})|\Psi(A)\rangle} \tag{A6}$$

is an infinite sum where in each term one differentiates with respect to one tensor in the bra layer. The first terms are the ones where the differential is on the Hamiltonian

$$\tag{A7}$$

Then we continue in the same row, with the term

$$\tag{A8}$$

Finally we have all the terms of the form

$$\tag{A9}$$

which can, again, be computed similarly to an out-of-time correlator using a finite-window MPS. Indeed, denoting the Hamiltonian MPS as

$$\tag{A10}$$

and the MPS with the differential

$$\tag{A11}$$

we find

$$g_2 = \sum_{n=1}^{\infty} \langle \Psi_{\partial}(M)|T^{n-1}|\Psi_H(M)\rangle \tag{A12}$$

$$= \langle \Psi_{\partial}(M)|(1-T)^{-1}|\Psi_H(M)\rangle. \tag{A13}$$

This correlator is most easily computed by a finite-window MPS, i.e., we sequentially apply the transfer matrix and approximate the state as

$$T^n|\Psi_H(M)\rangle \approx |\Psi(M; \{N_i\})\rangle \tag{A14}$$

and compute the overlaps

$$\langle \Psi_{\partial}(M)|\Psi(M; \{N_i\})\rangle. \tag{A15}$$

APPENDIX B: VARIATIONAL CHARACTERIZATION OF BOUNDARY MPS METHODS FOR MULTI-ROW TRANSFER MATRICES

In this Appendix, we give a detailed explanation of variational principles and contraction routines for larger unit cells. Such an explanation is intended to complement the discussion in Ref. [45], where it was missed that the multisite VUMPS algorithm might not converge for unit cells with more than two rows.

1. Discussion

The VUMPS fixed point equations [Eqs. (32)–(35)] are derived through the projection of the eigenvalue equation Eq. (21) onto the tangent space to the manifold of MPS,

$$\mathcal{P}_M(T|\Psi_M\rangle - \Lambda|\Psi_M\rangle) = 0, \tag{B1}$$

where $\mathcal{P}_M$ is the projector onto the tangent space of the MPS manifold at the point parameterized by M. For the case of a transfer matrix with a one-by-one unit cell where the tensor parametrizing T is Hermitian, the solution of Eq. (B1) directly corresponds to the variational optimum of a real-valued objective function representing the free-energy density Eq. (23). Thus the VUMPS algorithm for a transfer matrix with a one-by-one Hermitian unit cell gives an efficient method for obtaining the approximate leading eigenvector in a variationally optimal way.

For transfer matrices with a nontrivial unit cell, the cost of a naive application the VUMPS algorithm would scale exponentially with the size of the unit cell. It is therefore desirable to work with a variational algorithm that directly takes into account the nontrivial structure of the unit cell. One such algorithm is the multisite generalization of the VUMPS algorithm put forward in Ref. [45]. In this approach, one finds the approximate boundary MPS for an n-row Hermitian transfer matrix $T = T_n \ldots T_2 T_1$, where $T_i^{\dagger} = T_{n-i+1}$, in the following way. Instead of directly finding a single MPS $|\Psi_{\{M\}}\rangle$ which satisfies $T|\Psi_{\{M\}}\rangle = \Lambda|\Psi_{\{M\}}\rangle$, one formulates a recursive eigenvalue problem in terms of a set of MPSs $|\Psi_{\{M_i\}}\rangle$ that satisfy

$$\begin{aligned} T_1|\Psi_{\{M_1\}}\rangle &= \Lambda_1|\Psi_{\{M_2\}}\rangle, \\ T_2|\Psi_{\{M_2\}}\rangle &= \Lambda_2|\Psi_{\{M_3\}}\rangle, \\ &\cdots \\ T_n|\Psi_{\{M_n\}}\rangle &= \Lambda_n|\Psi_{\{M_1\}}\rangle, \end{aligned} \tag{B2}$$

where $\Lambda = \prod_{i=1}^{n} \Lambda_i$. The appropriate fixed point equations for a multirow boundary MPS algorithm are obtained by projecting each of these equations onto the tangent space

at the appropriate point on the MPS manifold,

$$\mathcal{P}_{\{M_{i+1}\}}(T_i|\Psi_{\{M_i\}}\rangle - \Lambda_i|\Psi_{\{M_{i+1}\}}\rangle)) = 0. \qquad (B3)$$

Whereas in the case of a single-row transfer matrix this tangent-space criterion is in direct correspondence with the variational optimum of a real-valued cost function, this is no longer necessarily the case when dealing with larger unit cells. For a two-row transfer matrix, there is no issue, as the conditions [Eqs. (B3)] still correspond to the variational optimum of Eq. (38). However, for transfer matrices with three or more rows this characterization in terms of optimizing the free-energy density breaks down, as mentioned in Sec. IV C.

The conditions [Eqs. (B3)] can however still be characterized in terms of an optimality criterion: they are in direct correspondence with the variational optimum of the normalized fidelity between consecutive MPS. Specifically, if

$$\{M_{i+1}\} = \underset{\{M_{i+1}\}}{\arg\max} \frac{|\langle\Psi_{\{M_{i+1}\}}|T_i|\Psi_{\{M_i\}}\rangle|^2}{\langle\Psi_{\{M_{i+1}\}}|\Psi_{\{M_{i+1}\}}\rangle}, \qquad (B4)$$

then Eq. (B3) must hold. This variational characterization in terms of the normalized fidelity naturally leads to a multirow boundary MPS algorithm in the form of a power method, as originally described in Sec. A 2 of Ref. [45]. Start with some initial MPS $|\Psi_{\{M_1^{(0)}\}}\rangle$ and consecutively apply the individual transfer matrix layers T_i, each time approximating the resulting MPS by some other MPS of a given bond dimension by variationally optimizing the normalized fidelity Eq. (B4) [46]. This procedure is repeated until convergence is reached, where the result will then correspond to the fixed point characterized by [Eqs. (B2)]. In Appendix B 2, we argue that this different variational characterization of the multirow algorithm in terms of an optimal normalized fidelity is compatible with the characterization in terms of optimizing the free-energy density, for systems that are sufficiently gapped, by considering the example of the two-dimensional classical Ising model. Note that the computational complexity of this power method is again linear in both linear dimensions of the unit cell L_x and L_y.

The algorithm proposed in Ref. [45], however, goes beyond this power method. The multisite VUMPS algorithm does not optimize [Eqs. (B4)] sequentially for one MPS at a time, but in fact iteratively solves a global fixed-point equation which is obtained by writing down all eigenvalue equations [Eqs. (B2)] in their tangent space projected form simultaneously, thereby optimizing over all MPS at once. This gives rise to a more efficient algorithm compared to the simple power method. As the conditions [Eqs. (B3)] are contained within the solution of this global fixed point equation, if the multisite VUMPS algorithm converges then the result is guaranteed to be optimal in the sense of the normalized fidelity.

An important point, however, is that the multisite algorithm may not converge for certain transfer matrices with three or more rows. This is caused by the fact that the algorithm optimizes over all MPS at the same time in order to solve a single global fixed point equation, instead of optimizing each fidelity individually as is actually implied by the tangent space conditions [Eqs. (B3)]. As stated in Sec. IV C, this phenomenon is related to the fact that the multisite VUMPS fixed point equations can actually be related to an invalid

variational principle of the form Eq. (46) in this case. So while the more efficient multisite VUMPS algorithm usually converges in which case the obtained solution automatically satisfies [Eqs. (B3)] in an optimal way, the algorithm can fail in which case the more robust power method should be used to obtain the desired fixed point.

2. Numerical results for the 2D Ising model

We now provide some numerical results in support of this discussion. The model we study is the classical ferromagnetic Ising model on the infinite two-dimensional square lattice with Hamiltonian

$$H = -J\sum_{\langle i,j\rangle} s_i s_j. \qquad (B5)$$

The corresponding partition function

$$\mathcal{Z} = \sum_{\{s_i\}} \mathrm{e}^{-\beta H} \qquad (B6)$$

can be written as the contraction of a 2D tensor network,

$$(B7)$$

where the black dots represent δ tensors

$$(B8)$$

and Boltzmann weights

$$-\!\!\!\!\boxed{t}\!\!\!\!- = \begin{pmatrix} \mathrm{e}^{\beta J} & \mathrm{e}^{-\beta J} \\ \mathrm{e}^{-\beta J} & \mathrm{e}^{\beta J} \end{pmatrix} \qquad (B9)$$

are place on each bond. Instead of the conventional way of simulating this model using a single-row transfer matrix, we will resort to a three-row transfer matrix here. Our approach is inspired by problems for which the use of a two-by-two unit cell transfer matrix is required, either explicitly or as a result of breaking translational invariance. For these models it often occurs that, while the total partition function has reflection and time-reversal symmetry, the corresponding transfer matrix unit cell is non-Hermitian and the variational characterization of boundary MPS contraction is therefore not valid. This occurs, for example, for certain classical lattice gauge theories in two spatial dimensions. In such cases, the partition function can often be transformed into an equivalent partition function with a larger Hermitian three-by-three unit cell by means of an identity for δ tensors, which in the current case

reads

$$+\!\!-\!\!\!\bigcirc\!\!t\!\!-\!\!+ = \quad .$$ (B10)

Here the black dots on the far left and right of the right-hand side represent three-leg δ tensors just as in Eq. (B8) and

$$-\!\!\bigcirc\!\!p\!\!- = \begin{pmatrix} e^{\frac{\beta J}{2}} & e^{-\frac{\beta J}{2}} \\ e^{-\frac{\beta J}{2}} & e^{\frac{\beta J}{2}} \end{pmatrix} .$$ (B11)

Applied to Eq. (B7), this gives rise to a partition function with a Hermitian three-by-three unit cell, indicated by the dashed square,

$$\mathcal{Z} = \quad .$$ (B12)

If we absorb the bond matrices into the sites symmetrically this gives a Hermitian three-row transfer matrix,

$$T = \quad ,$$ (B13)

where the bond and plaquette tensors of the unit cell are defined as

$$\bigcirc = -\!\!\bigcirc\!\!p\!\!-\!\!\bullet\!\!-\!\!\bigcirc\!\!p\!\!- , \qquad \square = \quad .$$ (B14)

Given this three-row transfer matrix, we now illustrate the points made in the above discussion.

We approximate the leading eigenvector of Eq. (B13) as a boundary MPS by directly using multisite algorithms, as well as by blocking the unit cell and employing the single-site VUMPS algorithm. In the following, we will use the term "single-row error" to denote the deviation of the boundary MPS with respect to the constraint Eq. (B1) where all rows of the transfer matrix are blocked, indicating its distance from the variational optimum of the free-energy density (25). The term "multirow error" will be used to denote the deviation with respect to the set of constraints [Eqs. (B3)], indicating its distance from the variational optimum of the normalized

FIG. 7. Multi- and single-row error at each iteration using the power method (top) and the multisite VUMPS algorithm (bottom) for determining the boundary MPS of Eq. (B13), at $T = 0.95 \, T_c$.

fidelity [Eqs. (B4)]. We set $J = 1$ in Eq. (B5) and consider a temperature range around the critical temperature $T_c = 2/\ln(1 + \sqrt{2})$). Using a bond dimension of $\chi = 50$ for the boundary MPS, we apply the power method described above, the multisite VUMPS algorithm [45] as well as the single-site VUMPS algorithm for the blocked unit cell [30] and analyze the results.

FIG. 8. Multi- and single-row error at each iteration using the single-site VUMPS algorithm for determining the boundary MPS of (B13) when blocking all layers, at $T = 0.95 \, T_c$.

FIG. 10. Increase of the multisite error after 40 additional multisite VUMPS iterations starting from the power method result for a temperature range $(.95 : .00125 : 1.05) \cdot T_c$. The dashed vertical line indicates $T = T_c$. Differences smaller than 10^{-10} are ignored since these are tainted by numerical noise.

FIG. 9. Multirow error at each iteration using the multisite VUMPS algorithm (top) and the power method (bottom) for determining the boundary MPS of (B13) at $T = T_c$.

In order to motivate the validity of the variational characterization in terms of an optimal fidelity per site, we take the MPS obtained with the multirow power method and the multisite VUMPS algorithm at each iteration and use it to compute the single-row error with respect to the blocked transfer matrix. The results are shown in Fig. 7, for $T = 0.95\,T_c$. The superior efficiency of the multisite VUMPS algorithm over the power method is immediately apparent here. Conversely, we take the MPS obtained with the single-site VUMPS algorithm for the blocked transfer matrix and use it to compute the multirow error with respect to the multirow transfer matrix by splitting the state into a three-site unit cell MPS by means of SVD. The result is shown in Fig. 8 for $T = 0.95\,T_c$. We see that both variational characterizations are compatible at a sufficient distance from criticality, i.e., if one characterization

is used to optimize the boundary MPS then the other characterization is automatically optimized as well.

As stated above, the multisite VUMPS algorithm [45] can suffer from stability issues for certain transfer matrices with three or more rows. In order to illustrate this, we monitor the multirow error for the first 50 iterations when determining the boundary MPS of Eq. (B13) at $T = T_c$ using both the multisite VUMPS algorithm and the multirow power method. The result is shown in Fig. 9. It can clearly be seen from the top panel that the behavior of the deviation with respect to Eq. (B3) is entirely erratic after a few iterations of the multisite VUMPS algorithm: the algorithm becomes unstable and will not converge. In contrast, this behavior does not occur when using the power method: the error measure keeps decreasing with the number of iterations and the algorithm will converge.

To further illustrate this instability, we take the fixed point obtained after 200 iterations of the power method and use this as the input for the multisite VUMPS algorithm. The increase of the deviation with respect to [Eqs. (B3)] for the multisite VUMPS result after 40 additional iterations compared to the power method result is shown in Fig. 10. It can clearly be seen that the multisite VUMPS algorithm becomes unstable in a finite temperature region around the critical temperature. Our simulations suggest that the size of this window is largely independent of the MPS bond dimension, while its position shifts very slightly towards lower temperatures when the bond dimension is increased.

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

[2] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, arXiv:cond-mat/04070663.

[3] F. Verstraete and J. I. Cirac, Matrix product states represent ground states faithfully, Phys. Rev. B **73**, 094423 (2006).

[4] T. Nishino, K. Okunishi, and M. Kikuchi, Numerical renormalization group at criticality, Phys. Lett. A **213**, 69 (1996).

[5] L. Tagliacozzo, T. R. de Oliveira, S. Iblisdir, and J. I. Latorre, Scaling of entanglement support for matrix product states, Phys. Rev. B **78**, 024410 (2008).

[6] F. Pollmann, S. Mukerjee, A. M. Turner, and J. E. Moore, Theory of Finite-Entanglement Scaling at One-Dimensional Quantum Critical Points, Phys. Rev. Lett. **102**, 255701 (2009).

[7] B. Pirvu, G. Vidal, F. Verstraete, and L. Tagliacozzo, Matrix product states for critical spin chains: Finite-size versus finite-entanglement scaling, Phys. Rev. B **86**, 075117 (2012).

[8] B. Vanhecke, J. Haegeman, K. Van Acoleyen, L. Vanderstraeten, and F. Verstraete, Scaling Hypothesis for Matrix Product States, Phys. Rev. Lett. **123**, 250604 (2019).

[9] M. Rader and A. M. Läuchli, Finite Correlation Length Scaling in Lorentz-Invariant Gapless iPEPS Wave Functions, Phys. Rev. X **8**, 031030 (2018).

[10] P. Corboz, P. Czarnik, G. Kapteijns, and L. Tagliacozzo, Finite Correlation Length Scaling with Infinite Projected Entangled-Pair States, Phys. Rev. X **8**, 031031 (2018).

[11] P. Czarnik and P. Corboz, Finite correlation length scaling with infinite projected entangled pair states at finite temperature, Phys. Rev. B **99**, 245107 (2019).

[12] B. Vanhecke, J. Hasik, F. Verstraete, and L. Vanderstraeten, A scaling hypothesis for projected entangled-pair states, arXiv:2102.03143.

[13] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[14] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Computational Complexity of Projected Entangled Pair States, Phys. Rev. Lett. **98**, 140506 (2007).

[15] H. C. Jiang, Z. Y. Weng, and T. Xiang, Accurate Determination of Tensor Network State of Quantum Lattice Models in Two Dimensions, Phys. Rev. Lett. **101**, 090603 (2008).

[16] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Classical Simulation of Infinite-Size Quantum Lattice Systems in Two Spatial Dimensions, Phys. Rev. Lett. **101**, 250602 (2008).

[17] P. Corboz, Variational optimization with infinite projected entangled-pair states, Phys. Rev. B **94**, 035133 (2016).

[18] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Gradient methods for variational optimization of projected entangled-pair states, Phys. Rev. B **94**, 155123 (2016).

[19] R. J. Baxter, Dimers on a rectangular lattice, J. Math. Phys. **9**, 650 (1968).

[20] R. J. Baxter, Variational approximations for square lattice models in statistical mechanics, J. Stat. Phys. **19**, 461 (1978).

[21] T. Nishino and K. Okunishi, Corner transfer matrix renormalization group method, J. Phys. Soc. Jpn. **65**, 891 (1996).

[22] T. Nishino and K. Okunishi, Corner transfer matrix algorithm for classical renormalization group, J. Phys. Soc. Jpn. **66**, 3040 (1997).

[23] R. Orús and G. Vidal, Simulation of two-dimensional quantum systems on an infinite lattice revisited: Corner transfer matrix for tensor contraction, Phys. Rev. B **80**, 094403 (2009).

[24] P. Corboz, J. Jordan, and G. Vidal, Simulation of fermionic lattice models in two dimensions with projected entangled-pair states: Next-nearest neighbor hamiltonians, Phys. Rev. B **82**, 245119 (2010).

[25] P. Corboz, T. M. Rice, and M. Troyer, Competing States in the *t-j* Model: Uniform *d*-Wave State Versus Stripe State, Phys. Rev. Lett. **113**, 046402 (2014).

[26] T. Nishino, Density matrix renormalization group method for 2D classical models, J. Phys. Soc. Jpn. **64**, 3598 (1995).

[27] R. Orús and G. Vidal, Infinite time-evolving block decimation algorithm beyond unitary evolution, Phys. Rev. B **78**, 155117 (2008).

[28] L. Vanderstraeten, M. Mariën, F. Verstraete, and J. Haegeman, Excitations and the tangent space of projected entangled-pair states, Phys. Rev. B **92**, 201111(R) (2015).

[29] J. Haegeman and F. Verstraete, Diagonalizing transfer matrices and matrix product operators: A medley of exact and computational methods, Annu. Rev. Condens. Matter Phys. **8**, 355 (2017).

[30] M. T. Fishman, L. Vanderstraeten, V. Zauner-Stauber, J. Haegeman, and F. Verstraete, Faster methods for contracting infinite two-dimensional tensor networks, Phys. Rev. B **98**, 235148 (2018).

[31] L. Vanderstraeten, J. Haegeman, and F. Verstraete, Tangent-space methods for uniform matrix product states, SciPost Phys. Lect. Notes, 7 (2019).

[32] B. Bauer, P. Corboz, R. Orús, and M. Troyer, Implementing global Abelian symmetries in projected entangled-pair state algorithms, Phys. Rev. B **83**, 125106 (2011).

[33] J. Hasik, D. Poilblanc, and F. Becca, Investigation of the Néel phase of the frustrated Heisenberg antiferromagnet by differentiable symmetric tensor networks, SciPost Phys. **10**, 012 (2021).

[34] M. Mambrini, R. Orús, and D. Poilblanc, Systematic construction of spin liquids on the square lattice from tensor networks with SU(2) symmetry, Phys. Rev. B **94**, 205124 (2016).

[35] D. Poilblanc and M. Mambrini, Quantum critical phase with infinite projected entangled paired states, Phys. Rev. B **96**, 014414 (2017).

[36] J.-Y. Chen, L. Vanderstraeten, S. Capponi, and D. Poilblanc, Non-Abelian chiral spin liquid in a quantum antiferromagnet revealed by an iPEPS study, Phys. Rev. B **98**, 184409 (2018).

[37] I. Kurečić, L. Vanderstraeten, and N. Schuch, Gapped SU(3) spin liquid with Z_3 topological order, Phys. Rev. B **99**, 045116 (2019).

[38] O. Gauthé, S. Capponi, and D. Poilblanc, SU(4) topological resonating valence bond spin liquid on the square lattice, Phys. Rev. B **99**, 241112(R) (2019).

[39] J.-Y. Chen, S. Capponi, A. Wietek, M. Mambrini, N. Schuch, and D. Poilblanc, $SU(3)_1$ Chiral Spin Liquid on the Square Lattice: A View from Symmetric Projected Entangled Pair States, Phys. Rev. Lett. **125**, 017201 (2020).

[40] D. Poilblanc, N. Schuch, and I. Affleck, $SU(2)_1$ chiral edge modes of a critical spin liquid, Phys. Rev. B **93**, 174414 (2016).

[41] A. Hackenbroich, A. Sterdyniak, and N. Schuch, Interplay of SU(2), point group, and translational symmetry for projected entangled pair states: Application to a chiral spin liquid, Phys. Rev. B **98**, 085151 (2018).

[42] S. Jiang and Y. Ran, Symmetric tensor networks and practical simulation algorithms to sharply identify classes of quantum phases distinguishable by short-range physics, Phys. Rev. B **92**, 104414 (2015).

[43] M. Hauru, M. Van Damme, and J. Haegeman, Riemannian optimization of isometric tensor networks, SciPost Phys. **10**, 040 (2020).

[44] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, Variational optimization algorithms for uniform matrix product states, Phys. Rev. B **97**, 045145 (2018).

[45] A. Nietner, B. Vanhecke, F. Verstraete, J. Eisert, and L. Vanderstraeten, Efficient variational contraction of two-dimensional tensor networks with a non-trivial unit cell, Quantum **4**, 328 (2020).

[46] B. Vanhecke, M. V. Damme, J. Haegeman, L. Vanderstraeten, and F. Verstraete, Tangent-space methods for truncating uniform MPS, SciPost Phys. Core **4**, 004 (2021).

[47] W.-L. Tu, H.-K. Wu, N. Schuch, N. Kawashima, and J.-Y. Chen, Generating function for tensor network diagrammatic summation, Phys. Rev. B **103**, 205155 (2021).

[48] K. A. Hallberg, Density-matrix algorithm for the calculation of dynamical properties of low-dimensional systems, Phys. Rev. B **52**, 9827(R) (1995).

[49] P. E. Dargel, A. Wöllert, A. Honecker, I. P. McCulloch, U. Schollwöck, and T. Pruschke, Lanczos algorithm with matrix product states for dynamical correlation functions, Phys. Rev. B **85**, 205119 (2012).

[50] S. Ramasesha, S. K. Pati, H. R. Krishnamurthy, Z. Shuai, and J. L. Brédas, Symmetrized density-matrix renormalization-group method for excited states of Hubbard models, Phys. Rev. B **54**, 7598 (1996).

[51] T. D. Kühner and S. R. White, Dynamical correlation functions using the density matrix renormalization group, Phys. Rev. B **60**, 335 (1999).

[52] E. Jeckelmann, Dynamical density-matrix renormalization-group method, Phys. Rev. B **66**, 045114 (2002).

[53] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, Variational matrix-product-state approach to quantum impurity models, Phys. Rev. B **80**, 165117 (2009).

[54] A. Holzner, A. Weichselbaum, I. P. McCulloch, U. Schollwöck, and J. von Delft, Chebyshev matrix product state approach for spectral functions, Phys. Rev. B **83**, 195115 (2011).

[55] S. R. White and I. Affleck, Spectral function for the S=1 Heisenberg antiferromagetic chain, Phys. Rev. B **77**, 134437 (2008).

[56] R. Pereira, S. White, and I. Affleck, Exact Edge Singularities and Dynamical Correlations in Spin-1/2 Chains, Phys. Rev. Lett. **100**, 027206 (2008).

[57] T. Barthel, U. Schollwöck, and S. R. White, Spectral functions in one-dimensional quantum systems at finite temperature using the density matrix renormalization group, Phys. Rev. B **79**, 245101 (2009).

[58] J. A. Kjäll, F. Pollmann, and J. E. Moore, Bound states and E_8 symmetry effects in perturbed quantum Ising chains, Phys. Rev. B **83**, 020407(R) (2011).

[59] L. Seabra and F. Pollmann, Exotic Ising dynamics in a Bose-Hubbard model, Phys. Rev. B **88**, 125103 (2013).

[60] A. Milsted, J. Haegeman, T. J. Osborne, and F. Verstraete, Variational matrix product ansatz for nonuniform dynamics in the thermodynamic limit, Phys. Rev. B **88**, 155116 (2013).

[61] H. N. Phien, G. Vidal, and I. P. McCulloch, Infinite boundary conditions for matrix product state calculations, Phys. Rev. B **86**, 245107 (2012).

[62] V. Zauner, M. Ganahl, H. G. Evertz, and T. Nishino, Time evolution within a comoving window: Scaling of signal fronts and magnetization plateaus after a local quench in quantum spin chains, J. Phys.: Condens. Matter **27**, 425602 (2015).

[63] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Differentiable Programming Tensor Networks, Phys. Rev. X **9**, 031041 (2019).

[64] B. Ponsioen and P. Corboz, Excitations with projected entangled pair states using the corner transfer matrix method, Phys. Rev. B **101**, 195109 (2020).

[65] L. Vanderstraeten, J. Haegeman, and F. Verstraete, Simulating excitation spectra with projected entangled-pair states, Phys. Rev. B **99**, 165121 (2019).

[66] B. Ponsioen, F. F. Assaad, and P. Corboz, Automatic differentiation applied to excitations with projected entangled pair states, SciPost Phys. **12**, 006 (2022).

[67] V. Zauner, D. Draxler, L. Vanderstraeten, M. Degroote, J. Haegeman, M. M. Rams, V. Stojevic, N. Schuch, and F. Verstraete, Transfer matrices and excitations with matrix product states, New J. Phys. **17**, 053002 (2015).

[68] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Stripe order in the underdoped region of the two-dimensional Hubbard model, Science **358**, 1155 (2017).

[69] B. Ponsioen, S. S. Chung, and P. Corboz, Period 4 stripe in the extended two-dimensional Hubbard model, Phys. Rev. B **100**, 195141 (2019).

[70] P. Corboz and F. Mila, Crystals of Bound States in the Magnetization Plateaus of the Shastry-Sutherland Model, Phys. Rev. Lett. **112**, 147203 (2014).