Matrix-product-state comparison of the numerical renormalization group and the variational formulation of the density-matrix renormalization group

Hamed Saberi, Andreas Weichselbaum, and Jan von Delft

Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience, Ludwig-Maximilians-Universität München, 80333 München, Germany

Wilson's numerical renormalization group (NRG) method for solving quantum impurity models yields a set of energy eigenstates that have the form of matrix product states (MPS). White's density-matrix renormalization group (DMRG) for treating quantum lattice problems can likewise be reformulated in terms of MPS. Thus, the latter constitute a common algebraic structure for both approaches. We exploit this fact to compare the NRG approach for the single-impurity Anderson model with a variational matrix product state approach (VMPS), equivalent to single-site DMRG. For the latter, we use an "unfolded" Wilson chain, which brings about a significant reduction in numerical costs compared to those of NRG. We show that all NRG eigenstates (kept and discarded) can be reproduced using VMPS, and compare the difference in truncation criteria, sharp vs smooth in energy space, of the two approaches. Finally, we demonstrate that NRG results can be improved upon systematically by performing a variational optimization in the space of variational matrix product states, using the states produced by NRG as input.

I. INTRODUCTION

Wilson's numerical renormalization group (NRG) is a highly successful method for solving quantum impurity models which allows the nonperturbative calculation of static and dynamic properties for a variety of impurity models.[1–6] NRG is formulated on a "Wilson chain," i.e., a tight-binding fermionic quantum chain with hopping matrix elements that decrease exponentially along the chain as $\Lambda^{-n/2}$, where $\Lambda > 1$ is a discretization parameter defined below and $n \geq 0$ is the chain's site index. It is thus not applicable to real-space quantum lattice problems featuring constant hopping matrix elements. For these, White's density-matrix renormalization group (DMRG) is the method of choice.[7–9] It has been known for some time[10,11] that the approximate ground states produced by DMRG have the form of matrix product states (MPS) [see Eq. (7) below] that had previously arisen in certain stochastic models[12] and quantum information processing.[13] This fact can be exploited to reinterpret the DMRG algorithm (more precisely, its one-site finite-size version) as a variational optimization scheme, in which the ground-state energy is minimized in the space of all matrix product states with specified dimensions.[11,14,15] To emphasize this fact, we shall refer to DMRG as "variational matrix product state" (VMPS) approach throughout this paper.

Quite recently it was understood[16] that NRG, too, in a natural way produces matrix product states. In other words, when applied to the same Wilson chain, NRG and VMPS produce approximate ground states of essentially the same MPS structure. The two approximate ground states are not identical, though, since the two methods use different truncation schemes to keep the size of the matrices involved manageable even for very long Wilson chains: NRG truncation relies on energy scale separation, which amounts to discarding the highest-energy eigenstates of a sequence of effective Hamiltonians, say $\mathcal{H}_n$, describing Wilson chains of increasing length n and yielding spectral information associated with the energy scale $\Lambda^{-n/2}$. This truncation procedure relies on the exponential decrease of hopping matrix elements along the Wilson chain, which ensures that adding a new site to the Wilson chain perturbs it only weakly. In contrast, VMPS truncation relies on singular value decomposition of the matrices constituting the MPS, which amounts to discarding the lowest-weight eigenstates of a sequence of reduced density matrices.[8] This procedure makes no special demands on the hopping matrix elements, and indeed works also if they are all equal, as is the case of standard quantum chain models for which DMRG was designed.

The fact that a *Wilson chain* model can be treated by two related but nonequivalent methods immediately raises an interesting and fundamental methodological question: How do the two methods compare? More precisely, to what extent and under which circumstances do their results agree or disagree? How do the differences in truncation schemes manifest themselves? VMPS, being a variational method operating in the same space of states as NRG, will yield a lower-energy ground state than NRG. However, it variationally targets *only* the ground state for the full Wilson chain, of length N, for example. In contrast, NRG produces a set of eigenenergies $\{E_\beta^n\}$ and eigenstates $\{|E_\beta^n\rangle\}$ for each of the sequence of effective Hamiltonians $\mathcal{H}_n$, with $n \leq N$, mentioned above. From these, a wealth of information about the RG flow, fixed points, relevant and irrelevant operators, their scaling dimensions, as well as static and dynamic physical properties can be extracted. Are these accessible to VMPS, too?

The goal of this paper is to explore such questions. We shall exploit the common matrix product state structure of the NRG and VMPS approaches to perform a systematic comparison of these two methods, as applied to the single-impurity Anderson model. It should be emphasized that our purpose is not to advocate using one method instead of the other. Instead, we hope to arrive at a balanced assessment of the respective strengths and weaknesses of each method.

In a nutshell, the main conclusion (which confirms and extends the results of Ref. 16) is the following: When applied to a Wilson chain with exponentially decreasing hopping, the VMPS approach is able to fully reproduce *all* information obtainable from NRG, despite being variationally optimized with respect to the ground state only. The reason is that the VMPS ground state is characterized by products of matrices of the form $\prod_{n=0}^{N} B^{[\sigma_n]}$ (details will be explained below), where the matrices with the same index n contain information about the energy scale $\Lambda^{-n/2}$. As will be shown below, this information can be used to construct eigenenergies $\{E_\beta^n\}$ and eigenstates $\{|E_\beta^n\rangle\}$ for a sequence of effective Hamiltonians $\mathcal{H}_n$ in *complete* analogy with (but not identical to) those of NRG. The agreement between NRG and VMPS results for these eigenenergies and eigenstates is excellent quantitatively, provided sufficient memory resources are used for both (and Λ is not too close to 1, see below). In this sense, NRG and VMPS can be viewed as yielding essentially equivalent results when applied to Wilson chains amenable to NRG treatment. In particular, *all* physical properties obtainable from the eigenspectra and eigenstates of NRG can likewise be obtained from those of VMPS.

Nevertheless, NRG and VMPS do differ in performance, flexibility, and numerical cost. First, since NRG truncation relies on energy scale separation, it works well only if the discretization parameter Λ is not too close to 1 (although the continuum limit of the model is recovered only in the limit $\Lambda \to 1$). This restriction does not apply to VMPS. Indeed, we shall find that NRG and VMPS agree well for $\Lambda = 2.5$, but less well for $\Lambda = 1.5$. This in itself is not surprising. However it does illustrate the power of VMPS to get by without energy scale separation. This very useful feature can be exploited, for example, to obtain well-resolved sharp spectral features at high energies in dynamical correlation functions,[16] using projection operator techniques. However, the latter results go beyond the scope of the present paper and will be published separately.

Second, since VMPS does not rely on energy scale separation, it does not need to treat all terms in the Hamiltonian characterized by the same scale $\Lambda^{-n/2}$ at the same time, as is required for NRG. This allows VMPS to achieve a significant reduction in memory cost compared to NRG for representing the ground state. To be specific: For NRG, we use the standard "folded" representation of the Wilson chain, in which each site represents both spin-down and spin-up electrons, with the impurity site at one end [see Fig. 1(a)]. However, it turns out that apart from the first few sites of the folded chain, the spin-down and -up degrees of freedom of each site are effectively not entangled with each other at all (see Fig. 3 below). For VMPS, we exploit this fact by using an "unfolded" representation of the Wilson chain instead,[16,17] in which the spin-up and spin-down sites lie on opposite sides of the impurity site, which sits at the center of the chain [see Fig. 1(b)]. This unfolded representation greatly reduces the memory cost, as characterized by the dimensions, D for NRG or D' for VMPS, of the effective Hilbert spaces needed to capture the low energy properties with the same precision: We find that with the choice $D' = 2^m \sqrt{D}$, VMPS can reproduce the results of NRG in the following manner: (i) if $m = 0$, the NRG ground state is reproduced qualitatively; (ii) if m =1, all the "kept" states of NRG are reproduced quantitatively; and (iii) if $m = 2$ all the "kept" *and* "discarded" states of NRG are reproduced quantitatively. However, in cases (ii) and (iii) the reduction in memory costs of VMPS is somewhat offset by the fact that the calculation of the excited eigenstates needed for the sake of direct comparison with NRG requires diagonalizing matrices of effective dimension D'^2. Note, nevertheless, that all information needed for this comparison is already fully contained within the VMPS *ground state* characterized by dimension D', since its constituent matrices contain information from all energy scales represented by the Wilson chain.

The paper is organized as follows: Sec. II sets the scene by introducing a folded and an unfolded version of the Wilson chain. In Secs. III and IV we review the NRG and VMPS approaches for finding the ground state of a folded or unfolded Wilson chain, respectively, emphasizing their common matrix product state structure. We also explain how an unfolded MPS state may be "refolded," allowing it to be compared directly to folded NRG states. In Sec. V we compare the results of NRG and VMPS, for ground-state energies and overlaps (Sec. V A), excited-state eigenenergies and density of states (Sec. V B), and the corresponding energy eigenstates themselves (Sec. V C). This allows us, in particular, to obtain very vivid insights into the differences in the truncation criteria used by the NRG and VMPS approaches, being sharp or smooth in energy space, respectively (Figs. 8–10). In Sec. VI we demonstrate that NRG results for the ground state can be improved upon systematically by first producing an unfolded "clone" of a given NRG ground state, and subsequently lowering its energy by performing variational energy minimization sweeps in the space of variational matrix product states. Finally, Sec. VII contains our conclusions and an assessment of the relative pros and cons of NRG and VMPS in relation to each other.

II. FOLDED AND UNFOLDED REPRESENTATIONS OF WILSON CHAIN

For definiteness, we consider the single-impurity Anderson model. It describes a spinful fermionic impurity level with energy ϵ_d and double occupancy cost U (with associated creation operators $f_{0\mu}^\dagger$, where $\mu = \downarrow, \uparrow$ denotes spin), which acquires a level width Γ due to being coupled to a spinful fermionic bath with bandwidth $W = 1$. Since the questions studied in this paper are of a generic nature and do not depend much on the specific parameter values used, we consider only the symmetric Anderson model and take $U = \frac{1}{2}$, $U/\pi\Gamma = 1.013$ and $\epsilon_d = -\frac{1}{2}U$ throughout this paper. To achieve a separation of energy scales, following Wilson,[1,2] the bath is represented by a set of discrete energy levels with logarithmically spaced energies Λ^{-n} (with associated creation operators $f_{n\mu}^\dagger$), where $n \geq 1$, $\Lambda > 1$ is a "discretization parameter," and the limit $\Lambda \to 1$ reproduces a continuous bath spectrum. The discretized Anderson model Hamiltonian can then be represented as

$$\mathcal{H}_{\mathrm{AM}} = \lim_{N \to \infty} \mathcal{H}_N, \tag{1}$$

where $\mathcal{H}_N$ describes a Wilson chain of "length N" (i.e., up to and including site N):

$$\mathcal{H}_N = \mathcal{H}_{N\downarrow} + \mathcal{H}_{N\uparrow} + U\left(f_{0\uparrow}^\dagger f_{0\uparrow} f_{0\downarrow}^\dagger f_{0\downarrow} + \tfrac{1}{2}\right), \qquad (2a)$$

$$\mathcal{H}_{N\mu} = \epsilon_d f_{0\mu}^\dagger f_{0\mu} + \sum_{n=0}^{N-1} t_n(f_{n\mu}^\dagger f_{(n+1)\mu} + \text{H.c.}), \qquad (2b)$$

with hopping coefficients given by

$$t_n \equiv \begin{cases} \sqrt{\dfrac{2\Gamma}{\pi}} & \text{for } n=0, \\[2ex] \dfrac{1}{2}(1+\Lambda^{-1})\Lambda^{-(n-1)/2}\xi_n & \text{for } n \geq 1, \end{cases}$$

$$\xi_n = (1-\Lambda^{-n})(1-\Lambda^{-2n+1})^{-1/2}(1-\Lambda^{-2n-1})^{-1/2}. \qquad (3)$$

In passing, we note that for our numerics we have found it convenient (following Refs. 17 and 16) to keep track of fermionic minus signs by making a Jordan-Wigner transformation[18] of the Wilson chain to a spin chain, using $f_{n\mu}^\dagger = P_{n\mu}s_{n\mu}^+$ and $f_{n\mu} = P_{n\mu}s_{n\mu}^-$. Here $s_{n\mu}^\pm$ are a set of spin-$\tfrac{1}{2}$ raising and lowering operators, that for equal indices satisfy $\{s_{n\mu}^-, s_{n\mu}^+\} = 1$, $(s_{n\mu}^-)^2 = (s_{n\mu}^+)^2 = 0$, but commute if their indices are unequal. The fermionic anticommutation relations for the $f_{n\mu}$ are ensured by the operators $P_{n\mu} = (-1)^{\sum_{(\bar{n}\bar{\mu})<(n\mu)} s_{\bar{n}\bar{\mu}}^+ s_{\bar{n}\bar{\mu}}^-}$, where $<$ refers to some implicitly specified ordering for the composite index $(n\mu)$. The $P_{n\mu}$ need to be kept track of when calculating certain correlation functions, but do not arise explicitly in the construction of the matrix product states that are the focus of this paper. This transformation will implicitly be assumed to have been implemented throughout the ensuing discussion.

For the Anderson model, site n of the Wilson chain represents the set of four states $|\sigma_n\rangle$, with $\sigma_n = (\sigma_{n\downarrow}, \sigma_{n\uparrow}) \in \{(00),(10),(01),(11)\}$, where $\sigma_{n\mu} \in \{0,1\}$, to be viewed as eigenvalue of $s_{n\mu}^+ s_{n\mu}^-$, gives the occupancy on site n of electrons with spin μ. Thus, the dimension of the spinful index σ_n is $d=4$, and that of the spin-resolved index $\sigma_{n\mu}$ is $d'=2$. As a general rule, we shall use the absence or presence of primes, d vs d' (and D vs D' below), to distinguish dimensions referring to spinful or spin-resolved indices, respectively, and correspondingly to folded or unfolded representations of the Wilson chain. For other quantum impurity models, such as the Kondo model or multilevel Anderson models, the dimension of the local impurity site, say d_0, differs from that of the bath sites, $d_0 \neq d$. It is straightforward to generalize the discussion below accordingly.

The Hamiltonian $\mathcal{H}_N$ of a Wilson chain of length N is defined on a Hilbert space of dimension d^{N+1}. It is spanned by an orthonormal set of states that, writing $|\sigma_n\rangle = |\sigma_{n\downarrow}\rangle|\sigma_{n\uparrow}\rangle$, can be written in either spinful or spin-resolved form,

$$|\boldsymbol{\sigma}^N\rangle = |\sigma_0\rangle|\sigma_1\rangle \dots |\sigma_N\rangle, \qquad (4a)$$

$$= |\sigma_{0\downarrow}\rangle|\sigma_{0\uparrow}\rangle|\sigma_{1\downarrow}\rangle|\sigma_{1\uparrow}\rangle \dots |\sigma_{N\downarrow}\rangle|\sigma_{N\uparrow}\rangle, \qquad (4b)$$

corresponding to a "folded" or "unfolded" representation of the Wilson chain, illustrated by Figs. 1(a) or 1(b), respectively. The unfolded representation of Fig. 1(b) makes explicit that the Anderson Hamiltonian of Eqs. (2a) and (2b)
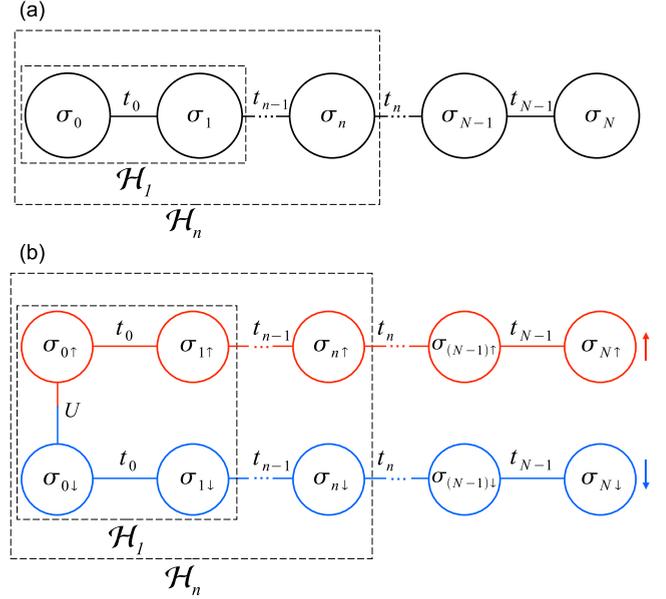
FIG. 1. (Color online) (a) The standard spinful or "folded" representation of the Wilson chain of the single-impurity Anderson model, and (b) its spin-resolved or "unfolded" representation. The latter makes explicit that spin-down and -up states are coupled *only* at the impurity sites and not at any of the bath sites. The dashed boxes indicate the chains described by $\mathcal{H}_1$ and $\mathcal{H}_n$, respectively.

has the form of two separate Wilson chains of specified spin, described by $\mathcal{H}_{N\downarrow}$ and $\mathcal{H}_{N\uparrow}$, *which interact only at site zero*. This fact will be exploited extensively below. Note that the ordering chosen for the $|\sigma_{n\mu}\rangle$ states in Eqs. (4a) and (4b) fixes the structure of the many-body Hilbert space once and for all. The fact that the sites of the unfolded chain in Fig. 1 are connected in a different order than that specified in Eqs. (4a) and (4b) is a statement about the dynamics of the model and of no consequence at this stage, where we simply fix a basis.

III. NUMERICAL RENORMALIZATION GROUP TREATMENT OF FOLDED WILSON CHAIN

A. Numerical renormalization group matrix product state arises by iteration

Wilson proposed to diagonalize the folded Wilson chain numerically using an iterative procedure, starting from a short chain and adding one site at a time. Consider a chain of length n, sufficiently short that $\mathcal{H}_n$ can be diagonalized exactly numerically. Denote its eigenstates by $|E_\alpha^n\rangle_f$, ordered by increasing energy $(E_\alpha^n)_f$, with $\alpha = 1, \dots, D_n$ and $D_n = d^{n+1}$. (We use subscripts f and u to distinguish quantities obtained from a folded or unfolded Wilson chain, respectively; similarly, in later parts of the paper we will use the subscripts r and c for "refolded" and "cloned.") E.g., for a chain consisting of only the impurity site, $n=0$, the d eigenstates can be written as linear combinations of the form $|E_\alpha^0\rangle_f = \sum_{\sigma_0} |\sigma_0\rangle A_{1\alpha}^{[\sigma_0]}$, where the coefficients have been arranged into d matrices $A^{[\sigma_0]}$ of dimensions $1 \times d$ (i.e., d-dimensional vectors), with matrix elements $A_{1\alpha}^{[\sigma_0]}$. Then add to the chain

the site $n+1$ and diagonalize $\mathcal{H}_{n+1}$ in the enlarged Hilbert space spanned by the $(D_n d)$ states $|E_\alpha^n\rangle_f |\sigma_{n+1}\rangle$. The new orthonormal set of eigenstates, with energies $(E_\beta^{n+1})_f$, can be written as linear combinations of the form

$$|E_\beta^{n+1}\rangle_f = \sum_{\sigma_{n+1}=1}^{d} \sum_{\alpha=1}^{D_n} |E_\alpha^n\rangle_f |\sigma_{n+1}\rangle A_{\alpha\beta}^{[\sigma_{n+1}]}, \qquad (5)$$

with $\beta = 1, \ldots, (D_n d)$. Here the coefficients specifying the linear combination have been arranged into a set of d matrices $A^{[\sigma_{n+1}]}$ of dimension $D_n \times D_{n+1}$, with matrix elements $A_{\alpha\beta}^{[\sigma_{n+1}]}$. The orthonormality of the eigenstates at each stage of the iteration, $_f\langle E_\beta^n | E_{\beta'}^n \rangle_f = \delta_{\beta\beta'}$, implies that the A-matrices automatically satisfy the orthonormality condition

$$\sum_{\sigma_n} A^{[\sigma_n]\dagger} A^{[\sigma_n]} = \mathbb{1}. \qquad (6)$$

We remark that it is possible to exploit symmetries of $\mathcal{H}_n$ (e.g., under particle-hole transformation) to cast A in block-diagonal form to make the calculation more time- and memory-efficient. However, for the purposes of the present paper, this was not required.

Iterating the above procedure by adding site after site and repeatedly using Eq. (5), we readily find that the NRG eigenstates of $\mathcal{H}_N$ on the folded Wilson chain can be written in the form of a so-called *matrix product state*,[16]

$$|E_\beta^N\rangle_f = \sum_{\{\boldsymbol{\sigma}^N\}} |\boldsymbol{\sigma}^N\rangle (A^{[\sigma_0]} A^{[\sigma_1]} \ldots A^{[\sigma_N]})_{1\beta}, \qquad (7)$$

illustrated in Fig. 2(a). Here matrix multiplication is implied in the product, $(A^{[\sigma_n]} A^{[\sigma_{n+1}]})_{\alpha\beta} = \sum_\gamma A_{\alpha\gamma}^{[\sigma_n]} A_{\gamma\beta}^{[\sigma_{n+1}]}$, and $\{\boldsymbol{\sigma}^N\}$ denotes the set of all sequences $\sigma_0, \sigma_1, \ldots, \sigma_N$. This matrix multiplication generates entanglement between neighboring sites, with the capacity for entanglement increasing with the dimension D_n of the index being summed over.

B. Numerical renormalization group truncation

In practice, it is of course not possible to carry out the above iteration strategy explicitly for chains longer than a few sites, because the size of the A-matrices grows exponentially with N. Hence Wilson proposed the following NRG truncation procedure: Once D_n becomes larger than a specified value, say D, only the lowest D eigenstates $|E_\alpha^n\rangle_f$, with $\alpha = 1, \ldots, D$, are retained or kept at each iteration, and all higher-lying ones are discarded.[19] Explicitly, the upper limit for the sum over α in Eq. (5) is redefined to be

$$D_n = \min(d^{n+1}, D). \qquad (8)$$

As a result, the dimensions of the $A^{[\sigma_n]}$ matrices occurring in the matrix product state (7) start from $1 \times d$ at $n=0$ and grow by a factor of d for each new site until they saturate at $D \times D$ after truncation has set in. The structure of the resulting states $|E_\beta^N\rangle_f$ is schematically depicted in Figs. 2(a) and 2(b), in which the site index is viewed as a single or composite index, σ_n or $(\sigma_{n\downarrow}, \sigma_{n\uparrow})$, respectively.

Wilson showed that this truncation procedure works well in practice, because the hopping parameters t_n of Eq. (3)
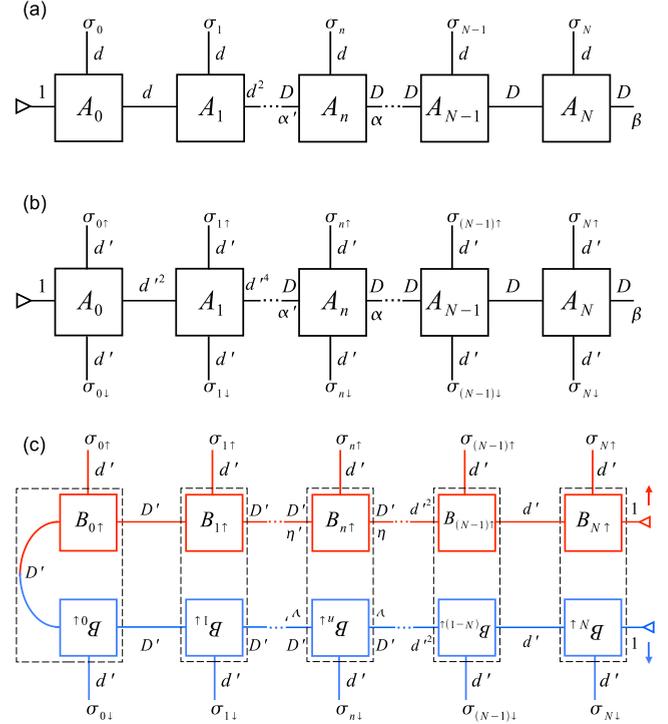
FIG. 2. (Color online) (a) and (b) show the matrix product structure of the state $|E_\beta^N\rangle_f$ of Eq. (7), depicting the site index as a single or composite index, σ_n or $(\sigma_{n\downarrow}, \sigma_{n\uparrow})$, respectively. (c) shows the matrix product structure of the state $|\Psi^N\rangle_u$ of Eq. (15). [For the sake of illustrating Eq. (A9) of Appendix A2, the labels $(B_{n\downarrow})_{\nu\nu'}$ in the bottom row are purposefully typeset "upside down," so that they would be right-side up if the chain of boxes were all drawn in one row in the order indicated by Eq. (15). Thus, the latter contains the factors $\ldots (B_{n\downarrow})_{\nu\nu'} \ldots (B_{n\uparrow})_{\eta'\eta} \ldots$, in that order, compare Eq. (A9).] Each matrix A or B is represented by a box, summed-over indices by links, free indices by terminals, and dummy indices having just a single value, namely 1, by ending in a triangle. The dimensions $(d, D, d', D',$ etc.) next to each link or terminal give the number of possible values taken on by the corresponding index, assuming Wilsonian truncation for (a) and (b), and VMPS truncation for (c). Note the similarity in structure between (c) and (b): the dashed boxes in the former, containing $B_{\nu\nu'}^{[\sigma_{n\downarrow}]} \otimes B_{\eta'\eta}^{[\sigma_{n\uparrow}]}$, play the role of the $A_{\alpha'\alpha}^{[(\sigma_{n\downarrow}, \sigma_{n\uparrow})]}$ matrices in the latter. Their capacity for entangling neighboring sites is comparable if one chooses $D'^2 \propto D$ [cf. Eq. (23)], since neighboring dashed boxes in (c) are connected by two links of combined dimension D'^2, whereas neighboring A-matrices in (b) are connected by only a single link of dimension D.

decrease exponentially with n: The resulting separation of energy scales along the chain ensures that high-lying eigenstates from iteration n make only a small contribution to the low-lying eigenstates of iteration $n+1$, so that discarding the former hardly affects the latter. The output of the NRG algorithm is a set of eigenstates $|E_\beta^n\rangle_f$ and eigenenergies $(E_\beta^n)_f$ for each iteration, describing the physics at energy scale $\Lambda^{-n/2}$. The NRG eigenenergies are usually plotted in rescaled form,

$$(\epsilon_\beta^n)_f = (E_\beta^n - E_1^n)_f / \Lambda^{-n/2}, \qquad (9)$$

as functions of n, to obtain a so-called NRG flow diagram; it converges to a set of fixed-point values as $n \to \infty$. Figure 7 in

Sec. V B below shows some examples. The ground-state energy of the entire chain is given by the lowest energy of the last iteration, $(E_G^N)_f = (E_1^N)_f$.

Despite the great success of NRG, Wilsonian truncation does have some drawbacks. First, its errors grow systematically as Λ tends to 1, because then the separation of energy scales on which it relies becomes less efficient. Second, it is not variational, and hence it is not guaranteed to produce the best possible approximation for the ground state within the space of all matrix product states of similar form and size. We shall return to this point later in Sec. VI and study quantitatively to what extent the NRG ground-state wave function can be improved upon by further variational optimization.

C. Mutual information of opposite spins on site n

A crucial feature of the folded Wilson chain is that all degrees of freedom associated with the same energy scale, $\Lambda^{-n/2}$, are represented by one and the same site and hence are all added during the same iteration step. Since the spin-down and -up degrees of freedom associated with each site are thus treated on an equal footing, the resulting matrix product state provides comparable amounts of resources for encoding entanglement between local states of the same spin, involving $|\sigma_{n\mu}\rangle|\sigma_{n+1\mu}\rangle$, or between states of opposite spin (indicated by the bar), involving $|\sigma_{n\mu}\rangle|\sigma_{n\bar{\mu}}\rangle$ or $|\sigma_{n\mu}\rangle|\sigma_{n+1\bar{\mu}}\rangle$. However, it turns out that for the Anderson model this feature, though *a priori* attractive, is in fact an unnecessary (and memory-costly) luxury: Since the Anderson model Hamiltonian (2a) and (2b) couples spin-down and -up electrons only at the impurity site, the amount of entanglement between states of opposite spin rapidly decreases with n.

To illustrate and quantify this claim, it is instructive to calculate the so-called mutual information $M_n^{\downarrow\uparrow}$ of the spin-down and -up degrees of freedom of a given site n. This quantity is defined via the following general construction:[20] Let C denote an arbitrary set of degrees of freedom of the Wilson chain, represented by the states $|\sigma^C\rangle$. Let ρ^C be the reduced density matrix obtained from the ground-state density matrix by tracing out all degrees of freedom except those of C, denoted by $N \backslash C$:

$$\rho^C = \sum_{\{\sigma^{N\backslash C}\}} \langle \sigma^{N\backslash C}|E_G^N\rangle_{ff}\langle E_G^N|\sigma^{N\backslash C}\rangle. \qquad (10)$$

For example, if C represents the spin-down and -up degrees of freedom of site n, its matrix elements are

$$\rho^C_{\sigma_n \sigma'_n} = \sum_{\{\sigma^{N\backslash n}\}} (A^{[\sigma_N]\dagger} \ldots A^{[\sigma_n]\dagger} \ldots A^{[\sigma_0]\dagger})_{G1}$$

$$\times (A^{[\sigma_0]} \ldots A^{[\sigma'_n]} \ldots A^{[\sigma_N]})_{1G}. \qquad (11)$$

If C represents only the spin-μ degree of freedom of site n, a similar expression holds, with n replaced by $n\mu$. The entropy associated with such a density matrix is given by

$$S^C = -\sum_i w_i^C \ln w_i^C, \qquad (12)$$

where w_i^C are the eigenvalues of ρ^C, with $\sum_i w_i^C = 1$. Now, consider the case that $C = AB$ is a combination of the degrees
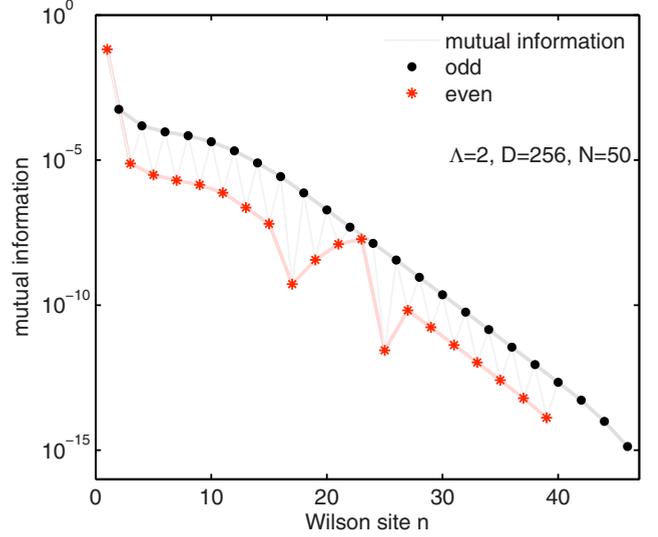
FIG. 3. (Color online) NRG result for the mutual information $M_n^{\downarrow\uparrow}$ between spin-down and -up degrees of freedom of site n of a folded Wilson chain of lenght $N=50$. The Anderson model parameters are fixed at $U=\frac{1}{2}$, $U/\pi\Gamma=1.013$, $\epsilon_d=-\frac{1}{2}U$ throughout this paper. Lines connecting data points are guides for the eye. The slight differences in behavior observed for even or odd n are reminiscent of the well-known fact (Ref. 1) that the ground-state degeneracy of a Wilson chain is different for even or odd N.

of freedom of two distinct subsets A and B, represented by states of the form $|\sigma^C\rangle = |\sigma^A\rangle|\sigma^B\rangle$. Then the mutual information of A and B, defined by

$$M^{AB} = S^A + S^B - S^{AB}, \qquad (13)$$

characterizes the information contained in ρ^{AB} beyond that contained in $\rho^A \otimes \rho^B$. The mutual information $M^{AB}=0$ if there is no entanglement between the degrees of freedom of A and B, since then $\rho^{AB} = \rho^A \otimes \rho^B$ and its eigenvalues have a product structure, $w_{ij}^{AB} = w_i^A w_j^B$.

We define the mutual information between spin-down and -up degrees of freedom of site n of the folded chain, $M_n^{\downarrow\uparrow}$, by Eq. (13), taking $A=n\downarrow$ and $B=n\uparrow$. Figure 3 shows this quantity as function of n for the symmetric Anderson model. Evidently $M_n^{\downarrow\uparrow}$ is very small for all but the first few sites, and decreases exponentially with n. This implies that for most of the folded chain, there is practically no entanglement between the spin-down and -up degrees of freedom. Consequently, the corresponding matrices occurring in Eq. (7) for $|E_G^N\rangle_f$ in effect have a direct product structure: loosely speaking, we may write $A^{[\sigma_n]} \simeq B^{[\sigma_{n\downarrow}]} \otimes B^{[\sigma_{n\uparrow}]}$. In Sec. IV A, we will exploit this fact to achieve a significant reduction in memory cost, by implementing the effective factorization in an alternative matrix product ansatz [see Eq. (15) below], defined on an *unfolded* Wilson chain which represents $n\downarrow$ and $n\uparrow$ of freedom by two separate sites.

IV. DENSITY-MATRIX RENORMALIZATION GROUP TREATMENT OF UNFOLDED WILSON CHAIN

A. Variational matrix product state ansatz

As pointed out by Verstraete *et al.*,[16] an alternative approach for finding a numerical approximation for the ground

state can be obtained by *variationally minimizing* the ground-state energy in the space of all VMPS of fixed norm. Implementing the latter constraint via a Lagrange multiplier λ, one thus considers the following minimization problem,

$$\min_{|\Psi\rangle \in \{|\Psi^N\rangle_u\}} [\langle\Psi|\mathcal{H}_N|\Psi\rangle - \lambda(\langle\Psi|\Psi\rangle - 1)]. \quad (14)$$

The minimization is to be performed over the space of all variational matrix product states $|\Psi^N\rangle_u$ having a specified structure (see below), with specified dimensions D'_n for the matrices, whose matrix elements are now treated as variational parameters. This minimization can be performed by a "sweeping procedure," which optimizes one matrix at a time while keeping all others fixed, then optimizing the neighboring matrix, and so forth, until convergence is achieved. The resulting algorithm is equivalent to a single-site DMRG treatment of the Wilson chain. Our main goal is to analyze how the energies and eigenstates so obtained compare to those produced by NRG.

Having decided to use a variational approach, it becomes possible to explore matrix product states having different, possibly more memory-efficient structures than those of Eq. (7) and Fig. 2(a). In particular, we can exploit[17] the fact that the Anderson model Hamiltonian (2a) and (2b) couples spin-down and -up electrons only at the impurity site, as emphasized in Eq. (2a) and (2b) and Fig. 1(b). For such a geometry, it is natural to consider matrix product states defined on the *unfolded* Wilson chain (subscript u) and having the following form, depicted schematically in Fig. 2(c):

$$|\Psi^N\rangle_u = \sum_{\{\boldsymbol{\sigma}^N\}} |\boldsymbol{\sigma}^N\rangle (B^{[\sigma_{N\downarrow}]} \dots B^{[\sigma_{0\downarrow}]} B^{[\sigma_{0\uparrow}]} \dots B^{[\sigma_{N\uparrow}]})_{11}. \quad (15)$$

The order in which the $B^{[\sigma_{n\mu}]}$ matrices occur in the product mimics the order in which the sites are connected in the unfolded Wilson chain. (The fact that this order differs from the order in which the basis states $|\sigma_{n\mu}\rangle$ for each site are arranged in the many-body basis state $|\boldsymbol{\sigma}^N\rangle$ [see Eq. (4b)] does not cause minus signs complications, because we work with Jordan-Wigner-transformed effective spin chains.) Each $B^{[\sigma_{n\mu}]}$ stands for a set of $d' = 2$ matrices with matrix elements $B_{\nu\eta}^{[\sigma_{n\mu}]}$, with dimensions $D'_n \times D'_{n-1}$ for $B^{[\sigma_{n\downarrow}]}$ and $D'_{n-1} \times D'_n$ for $B^{[\sigma_{n\uparrow}]}$, where

$$D'_n = \min(d'^{N-n}, D'), \quad (16)$$

as indicated on the links connecting the squares in Fig. 2(c). This choice of matrix dimension allows the outermost few sites at both ends of the unfolded chain to be described exactly (similarly to the first few sites of the folded Wilson chain for NRG), while introducing truncation, governed by D', for the matrices in the central part of the chain. The first index on $B_{1\nu}^{[\sigma_{N\downarrow}]}$ and the second index on $B_{\nu 1}^{[\sigma_{N\uparrow}]}$ are dummy indices taking on just a single value, namely 1, since they represent the ends of the chain. The triangles in Fig. 2(c) are meant to represent this fact. As a result, Eq. (15) represents just a single state, namely the ground state, in contrast to Eq. (7), which represents a set of states, labeled by the index β. Moving inward from the end points by decreasing n, the

matrix dimension parameter D'_n increases by one factor of d' for each site, in such a way that the resulting matrices are of just the right size to describe the outside ends of the chain (from n to N) *exactly*, i.e., without truncation. After a few sites, however, truncation sets in and the matrix dimensions saturate at $D' \times D'$ for the central part of the chain.

To initialize the variational search for optimal B-matrices, it turns out to be sufficient to start with a set of random matrices with normally distributed random matrix elements. Next, singular value decomposition is used to orthonormalize the B-matrices in such a way [see Eq. (A1)] that the matrix product state Eq. (15) has norm 1 (see Appendix A1 for details). Thereafter, variational optimization sweeps are performed to minimize Eq. (14) one B-matrix at a time.[16,21] After a sweeping back and forth through the entire chain a few times, the variational state typically converges (as illustrated by Fig. 13 in Sec. IV B), provided that D' is sufficiently large. We shall denote the resulting converged variational ground state by $|E_G^N\rangle_u$. Its variational energy, $(E_G^N)_u$, turns out to be essentially independent of the random choice of initial matrices.

B. Variational matrix product state truncation

Since $D' \times D'$ is the maximal dimension of B-matrices, D' is the truncation parameter determining the effective size of the variational space to be searched and hence the accuracy of the results. Its role can be understood more explicitly using a technique that is exceedingly useful in the VMPS approach, namely singular value decomposition: Any rectangular matrix $\mathcal{B}$ of dimension $m \times m'$ can be written as

$$\mathcal{B} = \mathcal{U}\mathcal{S}\mathcal{V}^\dagger, \quad \text{with} \quad \mathcal{U}^\dagger\mathcal{U} = \mathcal{V}^\dagger\mathcal{V} = \mathbb{1}, \quad (17)$$

where $\mathcal{S}$ is a diagonal matrix of dimension $\min(m, m')$, whose diagonal elements, the so-called "singular values," can always be chosen to be real and non-negative, and $\mathcal{U}$ and $\mathcal{V}^\dagger$ are column- and row-unitary matrices (with dimensions $m \times \min(m, m')$ and $\min(m, m') \times m'$, respectively). Due to the latter fact, the matrix norm of $\mathcal{B}$ is governed by the magnitude of the singular values.

For any given site of the unfolded Wilson chain, this decomposition can be applied in one of two ways (depending on the context, see Appendix A) to the set of matrices with elements $B_{\nu\eta}^{[\sigma_{n\mu}]}$: Introduce a composite index $\bar{\nu} = (\sigma_{n\mu}, \nu)$ [or $\bar{\eta} = (\sigma_{n\mu}, \eta)$] to arrange their matrix elements into a rectangular matrix carrying only two labels, with matrix elements $\mathcal{B}_{\bar{\nu}\eta} = B_{\nu\eta}^{[\sigma_{n\mu}]}$ (or $\tilde{\mathcal{B}}_{\nu\bar{\eta}} = B_{\nu\eta}^{[\sigma_{n\mu}]}$), and decompose this new matrix as $\mathcal{B} = \mathcal{U}\mathcal{S}\mathcal{V}^\dagger$.

Now, if this is done for any site for which the set of matrices $B^{[\sigma_{n\mu}]}$ have maximal dimensions $D' \times D'$, the corresponding matrix $\mathcal{S}$ will likewise have dimensions $D' \times D'$. Let its diagonal elements, the singular values s_ν (with $\nu = 1, \dots, D'$), be labeled in order of decreasing size. (Their squares, s_ν^2, correspond to the eigenvalues of the density matrix constructed in the course of the single-site DMRG algorithm.[8]) If D' is sufficiently large, the s_ν are typically found to decrease with increasing ν roughly as some negative power of ν, as illustrated in Fig. 4(a). The last and smallest of the singular values, $s_{D'}^2$ (squared, following Ref. 8),
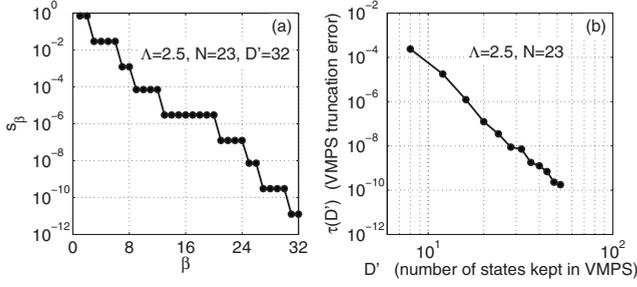
FIG. 4. (a) Typical singular value spectrum for site $5\downarrow$ of the unfolded Wilson chain, obtained by singular value decomposition of $B^{[\sigma_{5\downarrow}]}$. It shows, roughly, power-law decrease for large enough β, modulo steps due to degeneracies in the singular value spectrum. (b) D'-dependence of the truncation error $\tau(D')$ [Eq. (18)].

thus indicates the weight of the information that is lost at that site due to the given (finite) choice of D': By choosing D' larger, less information would be lost since more singular values (though of smaller size) would be retained. Repeating such an analysis for all sites of the unfolded Wilson chain, one may define the largest of the $s_{D'}^2$ parameters of the entire chain,

$$\tau(D') = \max_{\{n\mu\}}(s_{D'}^2), \qquad (18)$$

as "truncation error" characterizing the maximal information loss for a given value of D'. Typically, $\tau(D')$ decreases as some negative power of D', as illustrated in Fig. 4(b). In this way, D' assumes the role of a cutoff parameter that directly governs the accuracy of the VMPS approach, in a way analogous to the parameter D of NRG.

C. Refolding

The VMPS approach purposefully focuses on finding an optimal description of the variational *ground* state $|E_G^N\rangle_u$. Nevertheless, the B-matrices from which the latter is constructed contain information about all energy scales of the model, due to the logarithmic discretization of the Wilson chain. In particular, information about the scale $\Lambda^{-n/2}$ is encoded in the set of matrices $B^{[\sigma_{n\mu}]}$ associated with the two site $n\downarrow$ and $n\uparrow$. From these, it is possible to extract excited-state eigenspectra and energy flow diagrams in complete analogy to those produced by NRG. In this section we explain how this can be accomplished by a technique to be called "refolding," which combines the two matrices $B^{[\sigma_{n\downarrow}]}$ and $B^{[\sigma_{n\uparrow}]}$ into a single matrix, say $B^{[\sigma_n]}$, and thereby recasts unfolded matrix product states into folded ones. It should be emphasized that this procedure simply amounts to an internal reorganization of the representation of the VMPS ground state.

Consider a given matrix product state $|\Psi^N\rangle_u$ of the form (15), defined on an unfolded Wilson chain of length N (e.g., the converged ground state $|E_G^N\rangle_u$). To *refold* it (subscript r), it is expressed as a state of the following form [same as Eq. (7)],

$$|\Psi^N\rangle_r = \sum_{\{\boldsymbol{\sigma}^N\}} |\boldsymbol{\sigma}^N\rangle (B^{[\sigma_0]}B^{[\sigma_1]}\dots B^{[\sigma_N]})_{11}, \qquad (19)$$

defined on a folded Wilson chain of length N and normalized to unity, $_r\langle\Psi^n|\Psi^n\rangle_r=1$. Graphically speaking, this corresponds to rewriting a state of the form shown in Fig. 2(c) in terms of states of the form of Fig. 2(a). To obtain the matrices needed for Eq. (19), one constructs, for every site n of the refolded chain, a set of d matrices $B^{[\sigma_n]}$ from a combination of the two sets of spin-resolved matrices $B^{[\sigma_{n\downarrow}]}$ and $B^{[\sigma_{n\uparrow}]}$ of the unfolded chain (Appendix A2 gives the details of this construction). This is done in such a way, using singular value decomposition, that (i) the resulting matrices $B^{[\sigma_n]}$ satisfy the orthonormality conditions (6) (with $A \to B$), thereby guaranteeing the unit normalization of the refolded state $|\Psi^N\rangle_r$; and (ii) the $B^{[\sigma_n]}$ matrices have a structure similar to that of the matrices $A^{[\sigma_n]}$ generated by NRG, except that their dimensions, $D_n^r \times D_{n+1}^r$, are governed by

$$D_n^r = \min(d^n, d^{N+1-n}, D'^2) \qquad (20)$$

[instead of Eq. (8)], for reasons explained in Appendix A2. Thus, their dimensions have the maximal value $D^r \times D^r$, with $D^r = D'^2$, in the central part of the refolded chain, while decreasing at its ends toward $1 \times d$ or $d \times 1$ for $n=0$ or N, respectively.

Now, suppose that a converged variational ground state $|E_G^N\rangle_u$ has been obtained and refolded into the form $|\Psi^N\rangle_r$, so that the corresponding orthonormalized matrices $B^{[\sigma_n]}$ for the refolded Wilson chain of length N are the building blocks of the ground state of the system. Then it is possible to extract from them information about the many-body excitation spectrum at energy scale $\Lambda^{-n/2}$ that is analogous to the information produced by NRG. To this end, consider a *subchain* of length n of the full refolded Wilson chain, and use the definition

$$|\Psi_\beta^n\rangle_r = \sum_{\{\boldsymbol{\sigma}^n\}} |\boldsymbol{\sigma}^n\rangle (B^{[\sigma_0]}B^{[\sigma_1]}\dots B^{[\sigma_n]})_{1\beta}, \qquad (21)$$

[as in Eq. (19), but with N replaced by n] to construct a set of states $|\Psi_\beta^n\rangle_r$ on this subchain. These states, shown schematically by sites 0 to n of Fig. 2(a), form an orthonormal set, $_r\langle\Psi_\alpha^n|\Psi_\beta^n\rangle_r=\delta_{\alpha\beta}$, due to the orthonormality [Eq. (6)] of their constituent matrices. They can thus be viewed as a basis for that *subspace* of the many-body Hilbert space for the length-n Wilson chain, i.e., of that subspace of span$\{|\boldsymbol{\sigma}^n\rangle\}$, which VMPS sweeping has singled out to be most relevant for describing the ground state $|E_G^N\rangle_u$ of the full chain of length N. Therefore we shall henceforth call the $|\Psi_\beta^n\rangle_r$ "(re-folded) VMPS basis states" for this subchain.

This basis can be used to define an effective "refolded Hamiltonian" $\mathcal{H}_r^n$ for this subchain, with matrix elements

$$(\mathcal{H}_r^n)_{\alpha\beta} = {}_r\langle\Psi_\alpha^n|\mathcal{H}_n|\Psi_\beta^n\rangle_r. \qquad (22)$$

Its eigenvalues and eigenstates, say $(E_\beta^n)_r$ and $|E_\beta^n\rangle_r$, are the VMPS analogs of the NRG eigenvalues and eigenstates, $(E_\beta^n)_f$ and $|E_\beta^n\rangle_f$, respectively. They differ, in general, because VMPS and NRG use different truncation criteria, but are expected to agree well for sufficiently large choices of D'

and D. This is indeed found to be the case, as will be shown in detail in the next section.

V. COMPARISON OF NUMERICAL RENORMALIZATION GROUP AND VARIATIONAL MATRIX PRODUCT STATE RESULTS

Having outlined the NRG and VMPS strategies in Sec. III and IV, we now turn to a comparison of their results. This will be done, in successive sections, by comparing their ground-state energies and the overlaps of the corresponding ground states; the eigenspectra and density of states obtained from both approaches; and finally, the energy eigenstates used in the two approaches. We will thereby gain more insights into the differences between NRG and VMPS truncation criteria. Before embarking on a detailed comparison, though, some remarks on the choices to be made for D and D' are in order.

Since the structure of the matrix products occurring in Eqs. (7) and (15) differ, the spaces consisting of all states of the type $|E_\beta^n\rangle_f$ or $|E_\beta^n\rangle_r$, to be called the "NRG-subspace" or "VMPS-subspace" for a length-n chain, respectively, constitute nonidentical subspaces of the d^{n+1}-dimensional Hilbert space spanned by the basis states $|\boldsymbol{\sigma}^n\rangle$. The extent to which they describe the energy eigenstates of $\mathcal{H}_N$ with comparable accuracy will depend very strongly on the choices made for D and D'. It turns out (numerical evidence will be presented below) that with the choice

$$D' = d'^m \sqrt{D}, \qquad (23)$$

the VMPS-subspace is sufficiently large to give highly accurate representations of all kept states of NRG (including, in particular, the ground state) for the choice $m=1$, or of all kept *and* discarded states of NRG for the choice $m=2$. The fact that D' should be proportional to $\sqrt{D}$ can be made plausible by considering the following question: Given a folded Wilson subchain of length n (i.e., consisting of sites 0 to n) and its equivalent unfolded version, what are the smallest values for the dimensions D and D' for which both approaches describe the ground state *exactly*, i.e., without any truncation? Answer: On the one hand, the folded subchain has $n+1$ sites of dimension d, and hence a total dimension d^{n+1}; to ensure that the ground state in this space is described exactly, the kept space of the previous iteration must not involve any truncation, implying $D=d^n$. On the other hand, for the equivalent unfolded subchain, the spin $\downarrow$ and $\uparrow$ parts each have $n+1$ sites of dimension d', hence each have a Hilbert space of total dimension $d'^{(n+1)}$; to ensure that this space is described without truncation, its dimension should equal the maximal dimension of the B-matrices at sites 0μ, implying $D'=d'^{(n+1)}$. Using $d'=\sqrt{d}$ we readily find $D'=d'\sqrt{D}$, establishing the proportionality between D' and $\sqrt{D}$ and suggesting the choice $m=1$ to achieve an accurate VMPS-representation of the ground state. Actually, we find numerically that already $m=0$ yields good qualitative agreement between the VMPS and NRG ground states, while $m=1$ yields a *quantitatively* accurate VMPS-representation of the NRG ground state also for larger chain lengths, that do involve truncation. Since such ground states are built from the kept spaces of previous iterations, this implies that for $m=1$, all *kept* states in NRG (not only the ground state) are likewise well represented by VMPS. Indeed, we will find this to be the case. Moreover, it turns out numerically that with $m=2$, it is also possible to achieve an accurate VMPS-representation of all kept *and* discarded NRG-type states, as will be extensively illustrated below.

For the results reported below, we show data only for even iteration number n, to avoid even/odd oscillation effects that are typical and well understood for Wilsonian logarithmic discretization, but not of particular interest here. We set $D'=d'^m\sqrt{D}$ throughout and specify the choices made for m. All VMPS results shown in this section are extracted from randomly initialized, fully converged variational ground states $|E_G^N\rangle_u$ of the form (15).

A. Ground-state energies and overlaps

Figures 5(a) and 5(b) compare the NRG and VMPS *ground-state energies*, $(E_G^N)_f$ and $(E_G^N)_u$, for three values of Λ and, in Fig. 5(a), two values of m. They illustrate three points. First, for a given Λ the VMPS ground-state energies are smaller than those of NRG, $(E_G^N)_u < (E_G^N)_f$, as expected, since VMPS is a variational method and NRG is not. Second, Fig. 5(a) shows that larger values of m yield lower VMPS ground-state energies, as expected, since their variational space is larger. Third, the improvement of VMPS over NRG, as measured by the energy difference $(E_G^N)_f - (E_G^N)_u$ shown in Fig. 5(b), becomes more significant for smaller Λ, as expected, since the truncation scheme of NRG relies heavily on energy scale separation, and hence becomes less efficient for smaller Λ.

Figure 5(c) compares the overlap between NRG and VMPS ground states, characterized by the deviation from 1 of the overlap $\left|_f\langle E_G^N|E_G^N\rangle_u\right|$. The latter can be calculated straightforwardly from

$$_f\langle E_G^N|E_G^N\rangle_u = \sum_{\{\boldsymbol{\sigma}^N\}} (A^{[\sigma_N]\dagger} \dots A^{[\sigma_0]\dagger})_{G1}$$
$$\times (B^{[\sigma_{N\downarrow}]} \dots B^{[\sigma_{0\downarrow}]}B^{[\sigma_{0\uparrow}]} \dots B^{[\sigma_{N\uparrow}]})_{11}, \quad (24)$$

where the index contractions associated with the summation over repeated indices are illustrated in Fig. 6(a). Figure 5(c) shows that the deviation of the overlap from 1 becomes larger the smaller Λ, again illustrating that then the NRG truncation scheme becomes less reliable.

B. Comparison of eigenspectra and density of states

Figure 7 compares the energy *flow diagrams* obtained from NRG and refolded VMPS data, the latter obtained by diagonalizing the effective Hamiltonian of Eq. (22). It shows the rescaled energies $(\varepsilon_\beta^n)_{f,r}$ of Eq. (9) as functions of n, for four combinations of m and Λ, and illustrates the same trends as found in the previous subsection: First, the NRG and VMPS flow diagrams clearly agree not only for the ground state but also for a significant number of excited states. Evidently, the variational space searched by VMPS is large enough to capture considerable information about excited states, too, although the VMPS method was designed to

FIG. 6. (Color online) Contraction patterns used to calculate (a) the overlap $_f\langle E_G^N|E_G^N\rangle_u$ [Eqs. (24)] between folded NRG and unfolded VMPS ground states, and (b) the overlap matrix $\tilde{S}_{\alpha\beta}^n = {}_r\langle\Psi_\alpha^n|E_\beta^n\rangle_f$ [Eq. (27a) and (27b)] between refolded VMPS basis states and folded NRG eigenstates. The boxes represent A or B matrices in the graphical representation of Figs. 2(a) and 2(c), respectively (including the labeling conventions used there), and links connecting them represent indices that are being summed over.

As a complementary way of analyzing spectral information we also consider the "density of states," for a given iteration number n,

$$\rho_n(\varepsilon) = \sum_{\alpha=1}^{D_{\max}} \delta_\sigma(\varepsilon - \varepsilon_\alpha^n), \qquad (25)$$

using the rescaled eigenenergies ε_β^n of Eq. (9). Here $\delta_\sigma(\varepsilon) = e^{-\varepsilon^2/\sigma^2}/(\sigma\sqrt{\pi})$ is a Gaussian peak of width σ and unit weight, used to broaden the discrete spectrum in order to be able to plot it, and the number of states included in the sum is taken as $D_{\max}=dD$ or d^mD for NRG or VMPS results, respectively. Figure 8 shows such a density of states for several choices of m and iteration number n. It illustrates three points:

First, although for small energies $\rho_n(\varepsilon)$ grows rapidly with ε, as expected for a many-body density of states, it does not continue to do so for larger ε (the exact density of states would), due to the truncation inherent in both NRG and VMPS strategies. For NRG, $\rho_n(\varepsilon)$ drops to 0 very abruptly, because by construction Wilsonian truncation is sharp in energy space (at each iteration only the lowest dD eigenstates are calculated). In contrast, for VMPS $\rho_n(\varepsilon)$ decreases more gradually for large ε, because VMPS truncation for states at site n is based not on their energy, but on the variationally determined weight of their contribution to the ground state of the full Wilson chain of length N. Evidently, these weights

FIG. 5. (Color online) Comparison of NRG and VMPS results for (a,b) the ground-state energies and (c) the ground-state overlaps, plotted as functions of D with $D'=d'^m\sqrt{D}$, for three values of Λ and, in (a), for two values of m. In (a) the reference energies E_{ref}^N for each Λ were obtained by extrapolating the VMPS data points for $m=2$ to $D'\to\infty$, which represents the best estimate of the true ground-state energy available within the present set of methods. The power law fits to the numerical data in (b) and (c), shown as dashed lines, were made for the three data points with largest D, for which the dimensions are large enough to have reliable NRG data.

optimize only the ground state. Moreover, for a given choice of Λ, NRG and VMPS eigenenergies coincide for a larger number of states for $m=2$ than for $m=0$ [compare (b) to (a) and (d) to (c)], because the variational space is larger. Second, for a given choice of m, NRG and VMPS eigenenergies agree better for $\Lambda=2.5$ than for $\Lambda=1.5$ [compare (c) to (a) and (d) to (b)], as expected, because larger Λ leads to better energy scale separation and reduces the inaccuracies inherent in NRG's Wilsonian truncation scheme.

FIG. 7. (Color) Comparison of energy flow diagrams from NRG (dashed red lines) and refolded VMPS data (solid black lines), showing the rescaled energies $(\varepsilon_\beta^n)_{f,r}$ [Eq. (9)] versus n, calculated for even iteration numbers and four combinations of m (=0 or 2) and Λ (=1.5 or 2.5). The number of NRG states shown (kept *and* discarded) is Dd; the number of refolded VMPS states shown is $D^r = D'^2 = d^m D$, this being the maximal dimension of refolded matrices $B^{[\sigma_n]}$. For $m=2$ and $\Lambda = 2.5$, the NRG and DMRG flow diagrams agree very well, see (d).

decrease with increasing ε less rapidly than assumed by NRG.

Second, the agreement of the VMPS curve for $\rho_n(\varepsilon)$ with that of NRG is rather poor for $m=0$ (disagreement sets in already within the range of kept states of NRG, indicated by the shaded region), better for $m=1$ (the range of kept states is fully reproduced), and very good for $m=2$ (disagreement sets in only close to the upper end of range of discarded states).

Third, for large n, $\rho_n(\varepsilon)$ becomes increasingly spiky. This reflects the fact that the spectrum approaches a fixed point with regularly-spaced eigenenergies, as is evident in the energy flow diagrams of Fig. 7.

C. Comparison of energy eigenstates

To compare the energy *eigenstates* produced by NRG and refolded VMPS for a chain of length n, we analyze the overlap matrix

$$S_{\alpha\beta}^n = {}_r\langle E_\alpha^n | E_\beta^n \rangle_f. \qquad (26)$$

It can be conveniently calculated from $S^n = U^n \tilde{S}^n$, where $U_{\alpha\beta}^n = {}_r\langle E_\alpha^n | \Psi_\beta^n \rangle_r$ is the matrix that diagonalizes the effective Hamiltonian matrix $\mathcal{H}_{\alpha\beta}^n$ of Eq. (22), and the matrix

$$\tilde{S}_{\alpha\beta}^n = {}_r\langle \Psi_\alpha^n | E_\beta^n \rangle_f, \qquad (27a)$$

$$= \sum_{\{\boldsymbol{\sigma}^N\}} (B^{[\sigma_n]\dagger} \dots B^{[\sigma_0]\dagger})_{\alpha 1} (A^{[\sigma_0]} \dots A^{[\sigma_n]})_{1\beta} \qquad (27b)$$

characterizes how much weight the NRG eigenstates have in the space spanned by the refolded VMPS basis states, and vice versa. The contractions implicit in Eq. (27b) are illustrated in Fig. 6(b).

Figure 9 shows the overlap matrix $S_{\alpha\beta}^n$ on a color scale ranging from 0 to 1, for $m=1$ and several values of n. For the region of low excitation energies (about the first hundred or so states) its structure is evidently close to block-diagonal, indicating that both sets of states from which it is built are reasonably good energy eigenstates. Had both sets been perfect energy eigenstates, as would be the case for D' and D large enough to avoid all truncation, the blocks would be completely sharp, with sizes determined by the degeneracies of the corresponding energies. Sharp blocks are indeed observed for $n=2$ [Fig. 9(a)], because no truncation has oc-
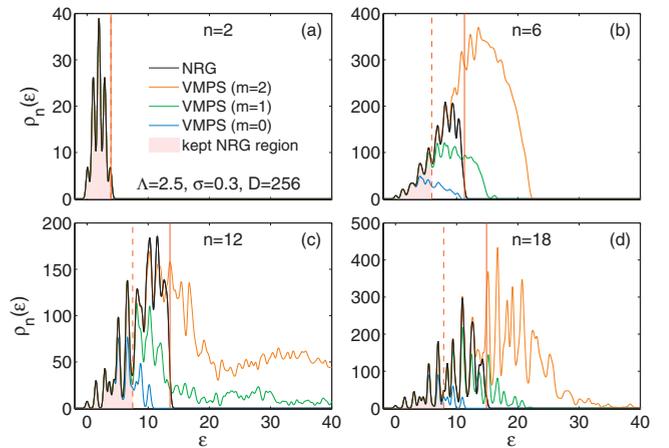
FIG. 8. (Color) Results for the density of states, $\rho_n(\varepsilon)$ [Eq. (25)], broadened with a Gaussian broadening function. In each panel, the red vertical dashed and solid lines [which coincide in (a)] indicate the energies of the highest-lying kept and discarded NRG states of that iteration, while the shaded area indicates the range of kept NRG states.
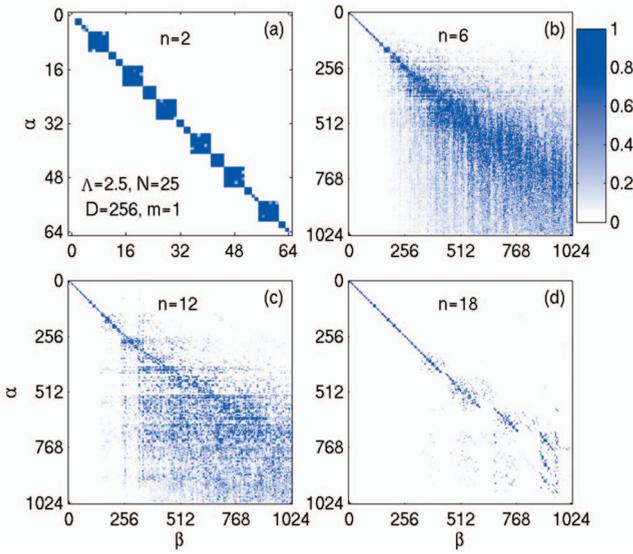
FIG. 9. (Color) Plot of the overlap matrix $S^n_{\alpha\beta} = {}_r\langle E^n_\alpha | E^n_\beta \rangle_f$ [Eq. (26)] between refolded VMPS and NRG energy eigenstates, with a color scale ranging between 0 and 1. In (a), with $n=2$, no truncation occurs at all, and both state labels α and β run from 1 to $d^{n+1} = 64$. In (b) to (d), truncation does occur: For the folded NRG eigenstates $|E^n_\beta\rangle_f$, the label α runs from 1 to $Dd = 1024$, i.e., it includes all kept and discarded NRG states, while for the refolded VMPS eigenstates $|E^n_\beta\rangle_r$, the label β runs from 1 to $D^r_n = D'^2 = 1024$ [Eq. (20)].

curred yet. The "fuzziness" shown by the blocks in Figs. 9(b)–9(d) for larger n implies that truncation is beginning to make itself felt, causing NRG and VMPS to increasingly disagree on how to construct the eigenstates corresponding to a given range of eigenenergies. Note that the fuzziness becomes markedly more pronounced for $\alpha, \beta > 256$. The reason is that whenever $S^n_{\alpha\beta}$ is nonzero for $\beta > D$, the associated VMPS states have weight among the discarded states of NRG, implying that NRG discards some states relevant for building the VMPS ground state. Thus, $S^n_{\alpha\beta}$ quite literally measures to what extent the truncation criteria of NRG and VMPS are compatible. Near the end of the chain, for $n=18$ [Fig. 9(d)], the off-diagonal spread is significantly reduced compared to the middle of the chain ($n=6,12$) [Figs. 9(b) and 9(c)], for two reasons. First, the dimensions of the refolded B-matrices become small for n near N [see Eq. (20)] so that the amount of truncation is much less severe near the end of the chain than in its middle. Second, the eigenspectra have converged to their fixed-point values, so that the number of different eigenenergies in a given energy interval is reduced, thus reducing the fuzziness in Fig. 9(d).

Next consider the total weight which a given NRG state $|E^n_\beta\rangle_f$ has within the refolded VMPS-subspace for that n,

$$w^{(n)}_\beta = \sum_{\alpha=1}^{D^r_n} |S_{\alpha\beta}|^2 = \sum_{\alpha=1}^{D^r_n} |\widetilde{S}_{\alpha\beta}|^2. \quad (28)$$

It satisfies $0 \leq w^{(n)}_\beta \leq 1$. Weights less than 1 imply that the VMPS-subspace is too small to adequately represent the corresponding NRG state. The second equality in Eq. (28), which follows from the unitarity of U, is useful since it im-
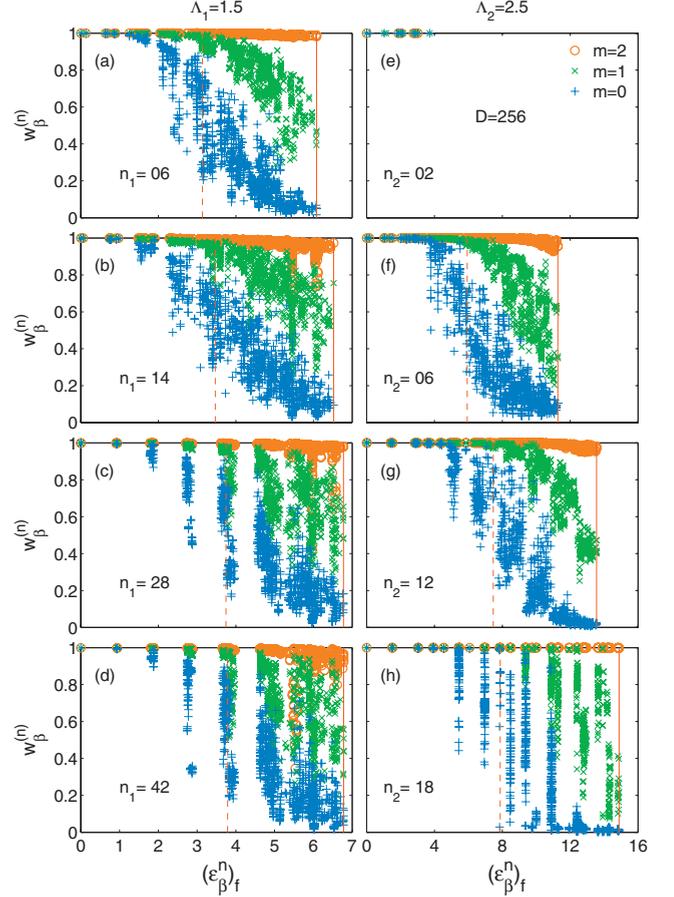
FIG. 10. (Color) For several NRG iteration numbers n and two values of Λ (different panels), this figure shows the weights $w^{(n)}_\beta$ [Eq. (28)] with which NRG states $|E^n_\beta\rangle_f$ with rescaled NRG eigenenergies $(\varepsilon^n_\beta)_f$ [Eq. (9)] are found to lie in the VMPS-subspace of dimension $D' = d'^m \sqrt{D}$, with $m=0$, 1 or 2 (indicated by +, × or ○, respectively). In each panel, the red vertical dashed and solid lines indicate the energies of the highest-lying kept and discarded NRG states of that iteration. For $n=3$, both of these lines are missing, since truncation has not yet set in. The choices for n in the left and right panels of each row are related by $\Lambda_1^{-n_1/2} = \Lambda_2^{-n_2/2}$, to ensure that both panels show a comparable energy scale.

plies that these weights can also be calculated directly from the refolded states $|\Psi^n_\beta\rangle_r$ [Eq. (21)], without the need for diagonalizing the large ($D'^2 \times D'^2$-dimensional) effective refolded Hamiltonian $\mathcal{H}^n_r$ [Eq. (22)].

Figure 10 shows such weights $w^{(n)}_\beta$ for various choices of n, Λ, and m. Their dependence on m reinforces the conclusions of the previous subsection: For $m=0$ (blue+symbols), the weights are equal to 1 for the lowest state of each iteration, but less than 1 for many of the kept states. This shows that the VMPS-subspace is large enough to accurately represent the NRG ground state, but significantly too small to accurately represent all kept states. For $m=1$ (green ×symbols), the weights are close to 1 only for the kept states, while smoothly decreasing toward 0 for higher-lying discarded states. Finally, for $m=2$ (orange○symbols), the weights of both kept and discarded NRG states are all close to 1, implying that the VMPS-subspace is large enough to accurately represent essentially *all* states kept track of by
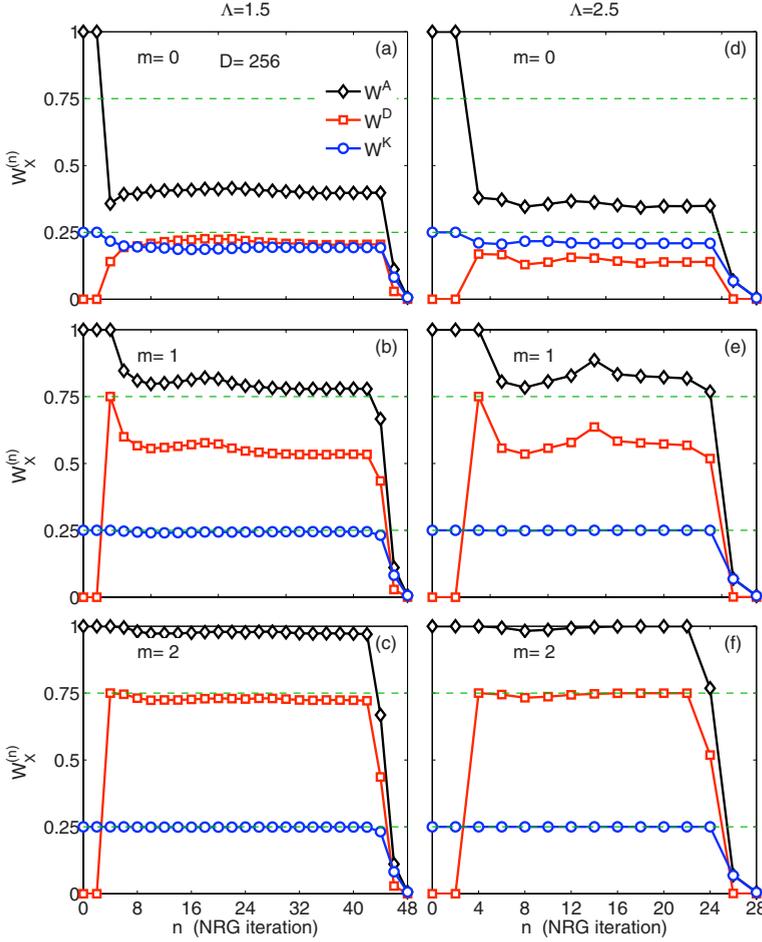
FIG. 11. (Color online) Integrated weights $W_X^{(n)}$ [see Eq. (29)] for two different Λ and three values of m. The dashed lines depict the maximum possible values of the kept and discarded weights, $\frac{1}{4}$ and $\frac{3}{4}$ for W_K^n and W_D^n, respectively.

NRG. Note that for $m=0$ and 1, the decrease of the weights $w_\beta^{(n)}$ with increasing energy occurs in a smooth and gradual fashion, illustrating yet again the smooth nature of VMPS truncation when viewed in energy space. When a smaller value of Λ is used [compare panels (a)—(d) to (e)-(h)] the weights of the higher-lying states of a given iteration tend to spread out over a larger range of values, since NRG has a weaker energy scale separation for smaller Λ. Finally, the increasing spikiness of the eigenspectrum with increasing n [see Figs. 10(d) and 10(h)] is due to the approach to a fixed-point spectrum with regularly-spaced eigenenergies, as mentioned above.

The results just discussed may be represented more compactly by considering, for a given iteration n, the integrated weights obtained by summing up the weights of *all* NRG states of type X,

$$W_X^{(n)} = \frac{1}{dD} \sum_{\beta \in X} w_\beta^{(n)}, \qquad (29)$$

where $X = K, D, A$ stands for *kept*, *discarded* or *all*, respectively. All three types of integrated weights are normalized to the total number dD of all NRG states calculated at a given iteration (with $d=4$ here), and reach their maximal values ($\frac{1}{4}$, $\frac{3}{4}$ and 1, respectively) when all the individual weights for that iteration equal 1. Figure 11 shows such integrated weights for several values of m and Λ. Upon increasing m

from 0 to 2, the integrated weights tend toward their maximal values, doing so more rapidly for larger Λ. For $m=2$, they essentially saturate their maximal values, indicating yet again that the VMPS variational space is now large enough to fully retain all information kept track of by NRG.

To summarize the result of this section: The VMPS approach reproduces NRG ground-state properties much more cheaply, requiring only $D' = \sqrt{D}$ for qualitative agreement, and $D' = d'\sqrt{D}$ for quantitative agreement. Moreover, it can also reproduce all kept and discarded NRG eigenstates if $D' = d'^2\sqrt{D}$ is used. However, to obtain excited energy eigenstates, we have to refold, requiring the diagonalization of matrices of dimension $D'^2 \times D'^2$. The numerical cost of doing so is comparable to that of NRG.

The fact that VMPS gives access to the same information on eigenstates and eigenvalues as NRG has a very significant and reassuring consequence: *all* physical properties of the model that can be calculated by NRG can also be calculated by VMPS, in combination with refolding.

VI. CLONING AND VARIATIONAL IMPROVEMENT OF NUMERICAL RENORMALIZATION GROUP GROUND STATE

Viewed in MPS language, the NRG method constructs the ground state in a single sweep along the chain: Each A is calculated only once, without allowing for possible feedback

of information from A's describing lower energies to those of higher energies calculated earlier. Thus, the resulting NRG ground state $|E_G^N\rangle_f$, to be denoted simply by $|G\rangle_f$ below, is not optimal in a variational sense. In this section we investigate to what extent the ground-state energy can be lowered further by performing variational energy optimization sweeps on $|G\rangle_f$ that serve to account for feedback of information from low to high energy scales. This feedback turns out to be small in practice, as will be seen below, but it is not strictly zero and its importance increases as the logarithmic discretization is refined by taking $\Lambda \to 1$.

A. Mapping folded to unfolded states by cloning

Our first step is to rewrite a given NRG ground state $|G\rangle_f$ in a form amenable to subsequent energy minimization sweeps. To this end, we use a variational *cloning* procedure (subscript c),

$$|G\rangle_f \overset{\text{cloning}}{\longrightarrow} |G\rangle_c \in \{|\Psi^N\rangle_u\}, \tag{30}$$

which maps $|G\rangle_f$ of the form of Eq. (7) [Fig. 2(a)] onto an unfolded state $|G\rangle_c$ of the form $|\Psi^N\rangle_u$ of Eq. (15) [Fig. 2(c)]. Since their matrix product structures differ, this mapping will, for general values of D and D', not be exact, though its accuracy should improve systematically with increasing D' and hence increasing dimensions of the variational space. To be explicit, we seek the best possible approximation to $|G\rangle_f$ in the space of all unfolded states of the form (15), by solving the minimization problem

$$\min_{|G\rangle_c \in \{|\Psi^N\rangle_u\}} \left[\| |G\rangle_f - |G\rangle_c \|^2 + \lambda(\| |G\rangle_c \|^2 - 1) \right], \tag{31}$$

which minimizes the "distance" between $|G\rangle_c$ and $|G\rangle_f$ under the constraint, implemented using a Lagrange multiplier λ, that the norm $_c\langle G|G\rangle_c = 1$ remains constant. Varying Eq. (31) with respect to the matrices defining $|G\rangle_c$ leads to a set of equations, one for each $k\mu$, of the form

$$\frac{\partial}{\partial B^{[\sigma_{k\mu}]}} \left[(1+\lambda)_c\langle G|G\rangle_c - 2\,\mathrm{Re}(_f\langle G|G\rangle_c) \right] = 0, \tag{32}$$

which determine the B-matrices of the desired "cloned" state $|G\rangle_c$. These equations can be solved in a fashion entirely analogous to energy optimization: Pick a particular site of the unfolded chain, say $k\mu$, and solve the corresponding Eq. (32) for the matrix $B^{[\sigma_{k\mu}]}$ while regarding the matrices of all other sites as fixed. Then move on to the neighboring site and in this fashion sweep back and forth through the chain until convergence is achieved. Appendix A3 describes some details of this procedure.

A figure of merit for the success of cloning is the deviation of the overlap $|_c\langle G|G\rangle_f|$ from 1. This deviation decreases monotonically with successive cloning sweeps and converges to a small but finite (D'-dependent) value when the cloning process converges, as illustrated in the inset of Fig. 12. The main part of Fig. 12 shows that when the number D' of VMPS states is increased, the converged value of the overlap deviation approaches 0 as a power law in D' (red circles). It also shows that the corresponding VMPS trunca-
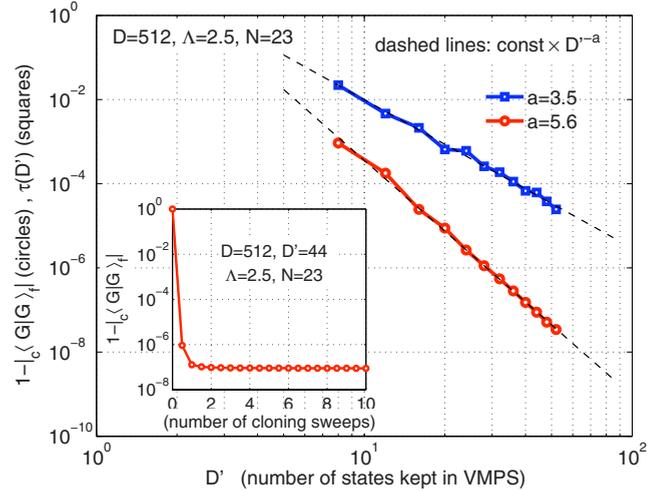
FIG. 12. (Color online) The deviation of the overlap $|_c\langle G|G\rangle_f|$ from 1 (red circles) and the cloning truncation error $\tau(D')$ (blue squares), as functions of the number D' of kept states in the cloning procedure. Both approach 0 in power-law fashion, as indicated by the dashed line fits. The inset shows how the overlap deviation from 1 decreases and converges to a small but finite constant in the course of sequential cloning sweeps.

tion error $\tau(D')$ incurred during cloning (blue squares), calculated according to Eq. (18), likewise decreases in power-law fashion with D'. All in all, Fig. 12 confirms that cloning works very well if D' is sufficiently large.

B. Variational energy minimization of $|G\rangle_c$

Having used cloning to find the optimal unfolded representation $|G\rangle_c$ of the NRG ground state $|G\rangle_f$, we now variationally minimize its energy by sweeping. We thereby obtain a sequence of states $|G\rangle_c^{k_{em}}$ of ever lower energy, $E_{k_{em}}$, where the index $k_{em} = 0, 1, 2, \dots$ gives the number of energy minimization sweeps that have been performed. The procedure is precisely analogous to that described in Sec. IV A, the only difference being that the random initial state used there is here replaced by the cloned state $|G\rangle_c^0 = |G\rangle_c$ as initial state.

Figure 13(a) shows the evolution of the ground-state energy $E_{k_{em}}$ as function of the number k_{em} of energy minimization sweeps, for both random (squares) and cloned (circles) initial states. $E_{k_{em}}$ is displayed with respect to the energy E_{ref} of a reference state $|G\rangle_{ref}$, defined in the figure caption, which represents our best approximation to the true ground state. Figure 13(b) shows how $1 - |_{ref}\langle G|G\rangle_c|$ decreases as sweeping proceeds, converging to a small but finite value. For a given value of Λ (1.5, shown in green, open symbols connected by dashed lines, or 2.5, shown in blue, filled symbols connected by solid lines), the energies for random and cloned initialization shown in Fig. 13(a) converge to the same value within just a few sweeps. However, the convergence is quicker for the cloned (circles) than the random (squares) input state, since the former represents an already rather good initial approximation (namely that of NRG) to the true ground state, whereas the latter is simply a random state. Nevertheless, the circles show strikingly that the NRG
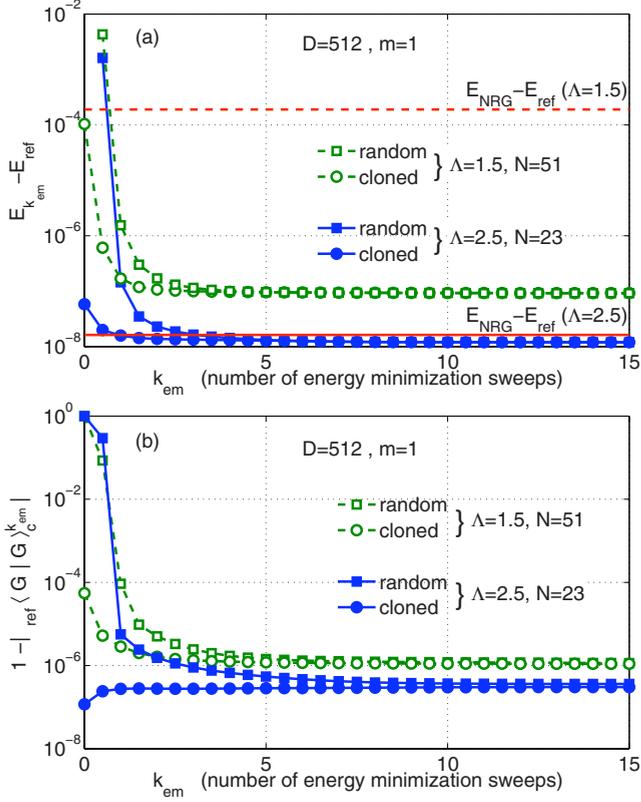
FIG. 13. (Color online) Comparison of (a) energies and (b) wave function overlaps for random initialization (squares) vs NRG-cloned initialization (circles), as functions of the number k_{em} of variational energy minimization sweeps. Results are shown for $\Lambda = 1.5$ (green, open symbols, dashed lines) and $\Lambda = 2.5$, (blue, filled symbols, solid lines). The energies in (a) and overlaps in (b) are calculated with respect to a reference ground state $|G\rangle_{ref}$ with $D' = 64$, obtained by performing 50 energy minimization sweeps starting from random initialization. The red horizontal straight lines in (a) (dashed or solid for $\Lambda = 1.5$ or 2.5, respectively), show the energy difference $E_{NRG} - E_{ref}$, where E_{NRG} is the energy of the NRG ground state $|G\rangle_f$ used as input into cloning. The fact that E_{NRG} does not completely coincide with the energy $E_c = E_{k_{em}=0}$ of the cloned state (horizontal straight lines do not meet circles at $k_{em} = 0$) is due to the fact that the deviation of the overlap $|_c\langle G|G\rangle_f|$ from 1 is not strictly equal to 0 (see Fig. 12).

ground energy is *not* optimal, in that the energy can be lowered still further by sweeping. Moreover, this improvement is more significant for small than large Λ (for circled data points, compare dashed green to solid blue lines for $\Lambda = 1.5$ or 2.5, respectively. The reason is that the NRG truncation scheme becomes less accurate the smaller Λ is, implying that the NRG result can be improved more significantly by further sweeping. This is again a reminder that the systematic error of NRG increases as Λ approaches 1, as already observed in Fig. 5.

VII. CONCLUSIONS

In this paper we presented a systematic comparison between NRG and DMRG, which we mainly referred to as VMPS, for the single-impurity Anderson model within the framework of matrix product states. We first reformulated both NRG and DMRG in the language of MPS, using a folded Wilson chain for NRG and an unfolded one for DMRG. Then we quantitatively compared the results of NRG and the VMPS approach for energy eigenvalues and eigenstates and explicitly analyzed the difference in their truncation criteria, which are sharp or smooth in energy space, respectively.

The most important conclusion of our study is this: For the purpose of obtaining the ground state of this model, the VMPS approach applied to the unfolded Wilson chain yields a very significant increase in numerical efficiency compared to NRG ($D' = d'\sqrt{D}$), essentially without loss of relevant information. The physical reason is that the spin-down and -up chains are only weakly entangled for this model, so that the NRG matrices $A^{[\sigma_n]}$ of dimension D that describe site n of the Wilson chain, can, in effect, be factorized as a direct product $\underline{B}^{[\sigma_n\downarrow]} \otimes \underline{B}^{[\sigma_n\uparrow]}$ of two matrices, each having dimension $d'\sqrt{D}$. It should be emphasized, though, that this property relies on the physics of the model, namely the weak entanglement of the spin-down and -up chains. To what extent this property survives for other impurity models should be a subject for further research, the two-channel Kondo model being a particularly interesting candidate in this respect.

Nevertheless, the possibility of using unfolded Wilson chains to reduce numerical costs for ground-state calculations is very attractive for possible applications of the VMPS method to more complicated models involving more than one conduction band.[22] For example, the conductance through a quantum dot coupled to two leads can under certain conditions (linear response, zero temperature, Fermi liquid behavior, etc.) be expressed in terms of a set of phase shifts that are uniquely determined by the ground-state occupation of the dot energy levels.[23] Thus, in such situations reliable knowledge of the ground state is sufficient to calculate transport properties.

Going beyond ground-state properties, we showed that the entire excited-state eigenspectrum of both kept and discarded NRG states can be recovered within the VMPS approach with at least the same accuracy as NRG, by using $D' = d'^2 \times \sqrt{D}$ and refolding. However, the latter step requires a subsequent additional diagonalization of matrices of dimensions D'^2, giving rise to a significant increase in numerical resources compared to the case that only ground-state information is required. A quantitative comparison between NRG and VMPS for the eigenspectrum's energies and eigenstates showed better agreement for $\Lambda = 2.5$ than 1.5, due to the fact that the NRG truncation scheme becomes increasingly less accurate the closer Λ approaches 1.

Finally, we used a cloning procedure to recast a given folded NRG ground state into an unfolded form, and showed that its energy could be lowered further by subsequent energy minimization sweeps. As expected, we found that sweeping improves the relative accuracy with which the ground-state energy can be determined, the more so the smaller the value of Λ. For example, for $\Lambda = 1.5$ the accuracy changed from $\mathcal{O}(10^{-4})$ before sweeping to $\mathcal{O}(10^{-7})$ thereafter [see Fig. 13(a)]. The fact that such a further variational im-

provement of the NRG ground state is possible, however, is of significance mainly as a matter of principle, not of practice: For the numerous situations where NRG works well (in particular, for Λ not too close to 1), we expect that such further variational improvement of the NRG ground state will not noticeably affect any physical observables.

Let us conclude with some comments about the pros and cons of NRG and VMPS. For quantum impurity models with a comparatively low degree of complexity, such as the single-lead Anderson and Kondo models, NRG works exceedingly well and for practical purposes nothing is to be gained from switching to the VMPS approach. The latter is a potentially attractive alternative to NRG only for two types of situations, namely (i) more complex quantum impurity models, and (ii) nonlogarithmic discretization of the leads. We briefly discuss these in turn.

(i) For complex quantum impurity models, in particular ones involving several leads, VMPS achieves a very significant reduction in memory cost, relative to NRG, for describing ground-state properties via unfolding the Wilson chain. There are several caveats, though. First, this reduction in memory cost applies only when *only* ground-state properties are of interest. To obtain excited-state eigenspectra, the memory costs of NRG and VMPS are comparable. Second, unfolding is expected to work well only for models for which the subchains that are being unfolded are only weakly entangled, which will not be the case for all impurity models. For example, the two-channel model might be an example where unfolding works less well. In general, one needs to check the extent to which degrees of freedom on different subchains are entangled with each other, by calculating the mutual information of two sites on different subchains. If this does not decrease rather rapidly with their separation from the impurity site, then unfolding will be a poor strategy. Appealingly, though, such a check can be done purely using NRG data, as illustrated in Sec. III C. Third, the fact that VMPS relies on variationally optimizing the ground state might cause convergence problems for models which have degenerate ground states. Conceivably this problem can be reduced by systematically exploiting all relevant symmetries of the Hamiltonian, including nonAbelian symmetries.[11,24,25] However, if states in the local state space of a folded Wilson chain are related by a nonAbelian symmetry, then this symmetry would not be manifest in the unfolded representation. Thus, the two possible strategies for achieving significant memory reduction, namely unfolding and exploitation of symmetries, might not always be mutually compatible; which one is more favorable will depend on the details of the model, and is an interesting subject for further study.

(ii) The VMPS approach offers clear advantages over NRG in situations where Wilson's logarithmic discretization of the conduction band cannot be applied. In the present paper, we found clear indications for this fact in the observation that the improvement of VMPS relative to NRG becomes more significant as Λ is chosen closer to 1. More importantly, VMPS offers the possibility, inaccessible to NRG, to improve the frequency resolution of spectral functions at high frequencies, by using a flexible (nonlogarithmic) discretization scheme which reduces the level spacing of effective lead states in the energy regimes where higher frequency resolution is desired. For such a discretization scheme Wilsonian energy scale separation is lost and NRG truncation cannot be applied. However, the ground state can still be found variationally, and spectral functions can be computed using projection operator techniques. In this fashion, it has recently been possible to calculate the spectral function for the Anderson model at large magnetic fields, $B > T_K$, and to resolve the split Kondo resonance with sufficient accuracy to reproduce the widths expected from perturbation theory in this regime. These developments, though, go beyond the scope of the present paper and will be published separately.[16]

APPENDIX: TECHNICAL DETAILS

In this Appendix, we collect some technical details on various manipulations involving matrix product states.

1. Orthonormalization of *B*-matrices of unfolded Wilson chain

To keep the notation simple, in this section we shall imagine the sites of the unfolded Wilson chain to be stretched along a line running from left to right, enumerated by an index k running from 1 for site $N\downarrow$ to $K=2(N+1)$ for site $N\uparrow$. Correspondingly, matrix product states will generically be written as $|\Psi\rangle = \sum_{\{\boldsymbol{\sigma}^K\}} |\boldsymbol{\sigma}^K\rangle (\prod_{k=1}^{K} B^{[\sigma_k]})$, with matrix elements $B^{[\sigma_k]}_{\nu\eta}$.

It is convenient to ensure that every B-matrix in a matrix product state satisfies one of the following two orthonormality conditions:

$$\sum_{\sigma_k} B^{[\sigma_k]\dagger} B^{[\sigma_k]} = \mathbb{1}, \tag{A1a}$$

$$\sum_{\sigma_k} B^{[\sigma_k]} B^{[\sigma_k]\dagger} = \mathbb{1}. \tag{A1b}$$

In particular, if *all* B-matrices satisfy either the first or the second of these conditions, the corresponding matrix product state is automatically normalized:

$$\langle \Psi | \Psi \rangle = \sum_{\{\boldsymbol{\sigma}^K\}} (B^{[\sigma_K]\dagger}_{1\nu'} \cdots B^{[\sigma_1]\dagger}_{\eta'1})(B^{[\sigma_1]}_{1\eta} \cdots B^{[\sigma_K]}_{\nu1}) = 1. \tag{A2}$$

This follows by iteratively applying Eq. (A1). To start the iteration, note that for matrices at the beginning or end of the chain, where one of the matrix indices is a dummy index with only a single value, Eq. (A1a) or Eq. (A1b) imply

$\Sigma_{\sigma_1} B^{[\sigma_1]\dagger}_{\eta'1} B^{[\sigma_1]}_{1\eta} = \delta_{\eta'\eta}$ or $\Sigma_{\sigma_K} B^{[\sigma_K]}_{\nu 1} B^{[\sigma_K]\dagger}_{1\nu'} = \delta_{\nu\nu'}$, respectively. In the NRG approach, all A-matrices naturally satisfy Eq. (A1a) [cf. Eq. (6)].

In the VMPS approach, it is convenient to ensure that during variational optimization sweeps, Eq. (A1a) holds for all matrices to the left of the site, say k, currently being updated, and Eq. (A1b) for all matrices to its right. Thus, after optimizing the set of matrices $B^{[\sigma_k]}$ at site k, this set should be orthonormalized before moving on to the next site, such that it satisfies Eq. (A1a) when sweeping from left to right [or Eq. (A1b) when sweeping from right to left]. This can be achieved using singular value decomposition [cf. Eq. (17)]: Arrange the matrix elements of the set of matrices $B^{[\sigma_k]}$ into a rectangular matrix carrying only two labels, with matrix elements $\mathcal{B}_{\bar\nu\eta} = B^{[\sigma_k]}_{\nu\eta}$ [or $\mathcal{B}_{\nu\bar\eta} = B^{[\sigma_k]}_{\nu\eta}$], by introducing a composite index $\bar\nu = (\sigma_k, \nu)$ (or $\bar\eta = (\sigma_k, \eta)$). Using singular value decomposition [Eq. (17)], write this new matrix as $\mathcal{B} = \mathcal{U}\mathcal{S}\mathcal{V}^\dagger$. Then rewrite the matrix product of two neighboring B-matrices as $B^{[\sigma_k]}B^{[\sigma_{k+1}]} = \tilde{B}^{[\sigma_k]}\tilde{B}^{[\sigma_{k+1}]}$ (or $B^{[\sigma_{k-1}]}B^{[\sigma_k]} = \tilde{B}^{[\sigma_{k-1}]}\tilde{B}^{[\sigma_k]}$), where the new matrices $\tilde{B}$ are defined by

$$\tilde{B}^{[\sigma_k]}_{\nu\gamma} = \mathcal{U}_{\bar\nu\gamma}, \quad \tilde{B}^{[\sigma_{k+1}]}_{\gamma\delta} = (\mathcal{S}\mathcal{V}^\dagger B^{[\sigma_{k+1}]})_{\gamma\delta}, \quad (A3)$$

$$[\text{or } \tilde{B}^{[\sigma_k]}_{\delta\eta} = \mathcal{V}^\dagger_{\delta\bar\eta}, \tilde{B}^{[\sigma_{k-1}]}_{\gamma\delta} = (B^{[\sigma_{k-1}]}\mathcal{U}\mathcal{S})_{\gamma\delta}]. \quad (A4)$$

The property $\mathcal{U}^\dagger\mathcal{U} = \mathbb{1}$ (or $\mathcal{V}^\dagger\mathcal{V} = \mathbb{1}$) ensures that the new set of matrices $\tilde{B}^{[\sigma_k]}$ at site k is orthonormal according to Eq. (A1a) [or Eq. (A1b)], as desired. Now proceed to the next site to the right (or left) and orthonormalize $\tilde{B}^{[\sigma_{k+1}]}$ (or $\tilde{B}^{[\sigma_{k-1}]}$) in the same manner, etc.

The above procedure can be used to orthonormalize the matrices of a randomly generated matrix product state before starting VMPS sweeping. Likewise, during VMPS sweeping, each newly optimized matrix can be orthonormalized in the above fashion before moving on to optimize the matrix of the next site.

2. Refolding

This section describes how to refold an unfolded matrix product state of the form

$$|\Psi^n_{\nu\eta}\rangle_u = \sum_{\{\boldsymbol{\sigma}^N\}} |\boldsymbol{\sigma}^n\rangle (B^{[\sigma_{n\downarrow}]} \dots B^{[\sigma_{0\downarrow}]} B^{[\sigma_{0\uparrow}]} \dots B^{[\sigma_{n\uparrow}]})_{\nu\eta},$$

$$(A5)$$

shown schematically by sites $n\downarrow$ to $n\uparrow$ of Fig. 2(c). Its two indices will be treated as a composite index $\beta = (\nu, \eta)$ below. The variational matrix product state $|\Psi^N\rangle_u$ of Eq. (15) discussed in the main text is a special case of Eq. (A5), with $n = N$ and $\nu = \eta = 1$. The goal is to express Eq. (A5) as a linear combination,

$$|\Psi^n_{\nu\eta}\rangle_u = \sum_\alpha |\Psi^n_\alpha\rangle_r C^n_{\alpha\beta}, \quad (A6)$$

[$\beta = (\nu, \eta)$ is a composite index] of an orthonormal set of "refolded basis states" of the form of Eq. (21),

$$|\Psi^n_\alpha\rangle_r = \sum_{\{\boldsymbol{\sigma}^n\}} |\boldsymbol{\sigma}^n\rangle (B^{[\sigma_0]} B^{[\sigma_1]} \dots B^{[\sigma_n]})_{1\alpha}, \quad (A7)$$

shown schematically by sites 0 to n of Fig. 2(a). To this end, we proceed iteratively in n. We use singular value decomposition to iteratively merge, for every pair of sites $n\downarrow$ and $n\uparrow$ of the unfolded chain, the matrices $B^{[\sigma_{n\downarrow}]}_{\nu\nu'}$ and $B^{[\sigma_{n\uparrow}]}_{\eta'\eta}$ into a new set of matrices $B^{[\sigma_n]}_{\alpha'\alpha}$ for site n of the refolded chain, thereby trading the indices $\sigma_{n\downarrow}, \sigma_{n\uparrow}$ and $\nu\eta$ of Fig. 2(c) for the indices σ_n and α of Fig. 2(a). This is to be done in such a way that the matrices $B^{[\sigma_n]}$ are orthonormal in the sense of Eq. (6), and that for the first few sites their dimensions increase in a way analogous to those of the $A^{[\sigma_n]}$ matrices of NRG, starting from $1 \times d$ at site $n = 0$.

For the first iteration step, start with $n = 0$, make a singular value decomposition of the matrix product

$$(B^{[\sigma_{0\downarrow}]} B^{[\sigma_{0\uparrow}]})_{\nu'\eta'} = (\mathcal{U}^0 \mathcal{S}^0 \mathcal{V}^{0\dagger})_{\sigma_0\beta'}, \quad (A8)$$

with $\beta' = (\nu', \eta')$, and use $\mathcal{U}^0$ to define a new set of d matrices $B^{[\sigma_0]}$ for site 0 of the refolded chain, with matrix elements $B^{[\sigma_0]}_{1\alpha'} = \mathcal{U}^0_{\sigma_0\alpha'}$. The $B^{[\sigma_0]}$ have dimensions $1 \times d$ (the dummy first index has just one value), and are by construction orthonormal in the sense of Eq. (6), since $\mathcal{U}^{0\dagger}\mathcal{U}^0 = \mathbb{1}$. Upon inserting Eq. (A8) into Eq. (A5), the factor $\mathcal{U}^0$ produces the first matrix factor $B^{[\sigma_0]}$ in the refolded state (A7), thus completing the first iteration step. For the second iteration step, contract the factors $\mathcal{S}^0\mathcal{V}^{0\dagger}$ with the factors $B^{[\sigma_{1\downarrow}]}$ and $B^{[\sigma_{1\uparrow}]}$ in Eq. (A5), factorize the result as $\mathcal{U}^1\mathcal{S}^1\mathcal{V}^{1\dagger}$ and use $\mathcal{U}^1$ to construct new matrices $B^{[\sigma_1]}$ for site 1 of the refolded chain, etc. To be explicit, for general n, make a singular value decomposition of the matrix product

$$\sum_{\nu'\eta'} B^{[\sigma_{n\downarrow}]}_{\nu\nu'} (\mathcal{S}^{(n-1)}\mathcal{V}^{(n-1)\dagger})_{\alpha'\beta'} B^{[\sigma_{n\uparrow}]}_{\eta'\eta} = (\mathcal{U}^n\mathcal{S}^n\mathcal{V}^{n\dagger})_{\bar\alpha\beta}, \quad (A9)$$

with composite indices $\bar\alpha = (\sigma_n, \alpha')$, $\sigma_n = (\sigma_{n\downarrow}, \sigma_{n\uparrow})$, $\beta = (\nu, \eta)$ and $\beta' = (\nu', \eta')$. Then use $\mathcal{U}^n$ to define a new set of orthonormal matrices $B^{[\sigma_n]}$ for site n of the refolded chain, with matrix elements $B^{[\sigma_n]}_{\alpha'\alpha} = \mathcal{U}^n_{\bar\alpha\alpha}$. In this way one readily establishes that $|\Psi^n_{\nu\eta}\rangle_u$ can be written in the form of Eq. (A6), with $C^n_{\alpha\beta} = (\mathcal{S}^n\mathcal{V}^{n\dagger})_{\alpha\beta}$.

The dimensions of the matrices $B^{[\sigma_n]}$ initially grow by a factor of d with each iteration step, until their dimensions are restricted by the number of possible values of the composite index β, namely D'^2_n, with D'_n given by Eq. (16). Thus, the $B^{[\sigma_n]}$ have dimensions $D^r_n \times D^r_{n+1}$, with $D^r_n = \min(d^n, D'^2_{n-1})$, which leads to Eq. (20).

3. Cloning

This section gives some details of the cloning procedure of Sec. VIA. The goal is to solve the variational Eq. (32), which determines the B-matrices of the cloned state $|G\rangle_c$. As described in the main text, this can be done by sweeping back and forth along the unfolded Wilson chain, and updating one matrix at a time.

Let $k\mu$ label the site to be updated and write the cloned state, which is assumed to be of the form (15), as

$$|G\rangle_c = (X_l^{k\mu})_{1\nu} B_{\nu\nu'}^{[\sigma_{k\mu}]} (X_r^{k\mu})_{\nu'1}. \tag{A10}$$

Here we introduced the shorthands

$$(X_l^{k\mu})_{1\nu} = (B^{[\sigma_{N\downarrow}]} \dots B^{[\sigma_{k_l\mu_l}]})_{1\nu}, \tag{A11a}$$

$$(X_r^{k\mu})_{\nu1} = (B^{[\sigma_{k_r\mu_r}]} \dots B^{[\sigma_{N\uparrow}]})_{\nu1}, \tag{A11b}$$

for the products of matrices standing before or after the one of present interest in the unfolded Wilson chain, and the labels $k_l\mu_l$ or $k_r\mu_r$ label the sites just before or after this site. Moreover, assume that all the B-matrices in X_l and X_r have been orthonormalized according to Eq. (A1a) or Eq. (A1b), respectively. (This can always be ensured by suitably orthonormalizing each B-matrix after updating it, see below.) These orthonormality relations immediately imply similar ones for the matrix products just introduced:

$$\sum_{\sigma_{N\downarrow},\dots,\sigma_{k_l\mu_l}} (X_l^{k\mu})_{\nu1}^\dagger (X_l^{k\mu})_{1\nu'} = \delta_{\nu\nu'}, \tag{A12a}$$

$$\sum_{\sigma_{k_r\mu_r},\dots,\sigma_{N\uparrow}} (X_r^{k\mu})_{\nu1} (X_r^{k\mu})_{1\nu'}^\dagger = \delta_{\nu\nu'}. \tag{A12b}$$

Thus, the norm of $|G\rangle_c$ can be written as

$$_c\langle G|G\rangle_c = \frac{1}{\mathcal{N}} \sum_{\nu\nu'} B_{\nu'\nu}^{[\sigma_{k\mu}]\dagger} B_{\nu\nu'}^{[\sigma_{k\mu}]}, \tag{A13}$$

where $\mathcal{N}$ is a normalization constant ensuring that the norm equals unity.

Using Eq. (A13), the variational Eq. (32) readily reduces to

$$B_{\nu\nu'}^{[\sigma_{k\mu}]} = \sum_{\{\boldsymbol{\sigma}'^N\}} \frac{(A^{[\sigma_N]\dagger} \dots A^{[\sigma_0]\dagger})_{G1}}{1+\lambda} (X_l^{k\mu})_{1\nu} (X_r^{k\mu})_{\nu'1}, \tag{A14}$$

where $\{\boldsymbol{\sigma}'^N\}$ denotes the local indices of all sites except the index $\sigma_{k\mu}$ of site $k\mu$, and we have assumed all A- and B-matrices to be purely real (exploiting the time-reversal invariance of the present model). This equation (graphically represented in Fig. 14) completely determines the new matrix $B^{[\sigma_{k\mu}]}$ in terms of the A-matrices specifying the NRG input state $|G\rangle_f$ and the B-matrices of sites other than the present one, which had been kept fixed during this variational step.
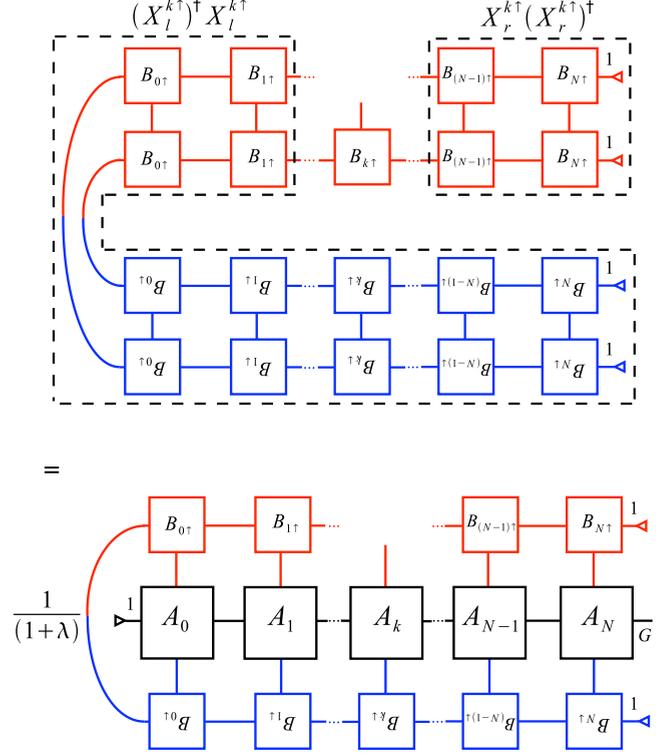
FIG. 14. (Color online) Graphical representation of the variational equation used for cloning, Eqs. (32) and (A14), drawn for the case $\mu=\uparrow$, and assuming all matrix elements to be real, using the labeling conventions of Figs. 2(a) and 2(c). The upper part of the figure represents $\frac{1}{2}\frac{\partial}{\partial B^{[k\uparrow]}}\langle G|G\rangle_c$; it simplifies to $B^{[k\uparrow]}$ [left hand side of Eq. (A14)] upon realizing that the parts in dashed boxes represent the left-hand sides of Eqs. (A12a) and (A12b), and hence reduce to unity.

Having calculated $B^{[\sigma_{k\mu}]}$, it should be properly orthonormalized, following the procedure of Eq. (A3) or Eq. (A4), depending on whether we are sweeping from left to right or vice versa. In other words, use the singular value decomposition $\mathcal{U}\mathcal{S}\mathcal{V}^\dagger$ of the newfound matrix $B^{[\sigma_{k\mu}]}$, to transfer a factor $\mathcal{S}\mathcal{V}^\dagger$ or $\mathcal{U}\mathcal{S}$ onto its right or left neighbor, respectively, and rescale this neighbor by an overall constant to ensure that the new state $|G\rangle_c$ is still normalized to unity. This concludes the update of site $k\mu$. Now move on to its neighbor, etc., and thus sweep back and forth through the unfolded Wilson chain, until convergence is reached.

[1] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[2] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980).

[3] R. Bulla, T. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

[4] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005); Phys. Rev. B **74**, 245113 (2006).

[5] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006).

[6] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[7] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[8] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[9] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[10] S. Ostlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).

[11] J. Dukelsky, M. A. Martín-Delgado, T. Nishino, and G. Sierra,

Europhys. Lett. **43**, 457 (1998).

[12] A. Honecker and I. Peschel, J. Stat. Phys. **88**, 319 (1997).

[13] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).

[14] H. Takasaki, T. Hikihara, and T. Nishino, J. Phys. Soc. Jpn. **68**, 1537 (1999).

[15] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004).

[16] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, arXiv:cond-mat/0504305v2 (unpublished).

[17] C. Raas, G. S. Uhrig, and F. B. Anders, Phys. Rev. B **69**, 041102(R) (2004).

[18] P. Jordan and E. P. Wigner, Z. Phys. **47**, 631 (1928).

[19] It has recently been shown (Ref. 4) that the discarded states can be used to construct a complete basis of many-body states spanning the full d^{N+1}-dimensional Hilbert space of $\mathcal{H}_N$, and that this basis can be used to greatly improve the accuracy of the NRG calculations of spectral functions (Refs. 5 and 6).

[20] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[21] The technical details of this procedure will be published separately.

[22] A. Holzner, A. Weichselbaum, and J. von Delft, arXiv:0804.0550 (unpublished).

[23] M. Pustilnik and L. Glazman, J. Phys.: Condens. Matter **16**, R513 (2004).

[24] A. I. Tóth, C. P. Moca, O. Legeza, and G. Zaránd, arXiv:0802.4332 (unpublished).

[25] I. P. McCulloch, J. Stat. Mech.: Theory Exp. **2007**, P10014.