Title	Operator-based truncation scheme based on the many-body fermion density matrix(Published version)
Author(s)	Cheong, Siew Ann; Henley, Christopher L.
Citation	Cheong, S. A., & Henley, C. L. (2004). Operator-based truncation scheme based on the many-body fermion density matrix. Physical Review B., 69(7), 1-20.
Date	2004
URL	http://hdl.handle.net/10220/4601
Rights	Physical Review B. @ copyright ©2004 The American Physical Society. The journal's website is located at http://www.aps.org/.

Operator-based truncation scheme based on the many-body fermion density matrix

Siew-Ann Cheong and Christopher L. Henley

Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501, USA

In an earlier work [S. A. Cheong and C. L. Henley, preceding paper], we derived an exact formula for the many-body density matrix ρ_B of a block of B sites cut out from an infinite chain of noninteracting spinless fermions, and found that the many-particle eigenvalues and eigenstates of ρ_B can all be constructed out of the one-particle eigenvalues and one-particle eigenstates, respectively. In this paper we improved upon this understanding, and developed a statistical-mechanical analogy between the density-matrix eigenstates and the many-body states of a system of noninteracting fermions. Each density-matrix eigenstate corresponds to a particular set of occupation of single-particle pseudoenergy levels, and the density-matrix eigenstate with the largest weight, having the structure of a Fermi sea ground state, unambiguously defines a pseudo-Fermi level. Based on this analogy, we outlined the main ideas behind an operator-based truncation of the density-matrix eigenstates, where single-particle pseudoenergy levels far away from the pseudo-Fermi level are removed as degrees of freedom. We report numerical evidence for scaling behaviors in the single-particle pseudoenergy spectrum for different block sizes B and different filling fractions $\bar{n}$. With the aid of these scaling relations, which tell us that the block size B plays the role of an inverse temperature in the statistical-mechanical description of the density-matrix eigenstates and eigenvalues, we looked into the performance of our operator-based truncation scheme in minimizing the discarded density-matrix weight and the error in calculating the dispersion relation for elementary excitations. This performance was compared against that of the traditional density-matrix-based truncation scheme, as well as against an operator-based plane-wave truncation scheme, and found to be very satisfactory.

I. INTRODUCTION

An unsolved problem in many-body physics is to augment numerical exact diagonalization, which is only feasible on comparatively small systems, so as to give the maximum information about the thermodynamic limit. We know from the renormalization group that only a few degrees of freedom, viz., those with low energies and long wavelengths, really matter. Hence it should be possible in principle to discard the irrelevant parts of the Hilbert space, but no method has been developed in real space for higher-dimensional, interacting lattice models. The density-matrix renormalization group (DMRG) (Refs. 1–4) is the only successful method to date, but it is limited to one-dimensional chains, or two-dimensional systems that can be forced into one dimension as strips.[5,6]

To get more mileage out of density-matrix-based renormalization-group methods, surely we must develop a deep understanding of the structure of density matrices of the simplest possible systems, for which analytic results are available to guide us. Therefore, it is appropriate to begin by considering the ground state of a one-dimensional chain of noninteracting spinless fermions described by a nearest-neighbor hopping Hamiltonian

$$H = -\sum_j \left[c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j \right]. \tag{1.1}$$

Here we observe that, in one dimension, one has the same Fermi sea ground state for a variety of translationally invariant Hamiltonians with hoppings to further neighbors, provided their dispersion relation is monotonic so that the Fermi

surface occurs at the same wave vector for a given filling. A "block" is then identified within this one-dimensional system by choosing B sites that need not be contiguous, following which we can define the many-body density matrix ρ_B of the block starting from the zero-temperature many-body wavefunction and tracing out all sites outside the chosen block of B sites. The basis of this paper is the simple factorized nature of ρ_B and its eigenfunctions, which we derived exactly for noninteracting (spinless or spinfull) fermions,[7] by extending a technique introduced by Chung and Peschel.[8]

In Ref. 8, it was shown that many-particle density-matrix eigenstates are built from a set of single-particle creation operators and eigenvalues, quite analogous to the energy eigenstates of a noninteracting system of fermions. For a block of B sites cut out from a larger overall system, there are B such *pseudocreation operators* $f_l^\dagger$, and the block's Hilbert space is therefore spanned by all 2^B products of the $f_l^\dagger$'s. The density matrix of the block can be written as

$$\rho_B = \mathcal{Q}^{-1} \exp(-\tilde{H}), \tag{1.2}$$

which is the exponential of a *pseudo-Hamiltonian* given by

$$\tilde{H} = \sum_{l=1}^{B} \varphi_l f_l^\dagger f_l, \tag{1.3}$$

with *single-particle pseudoenergy spectrum* φ_l.

This paper is devoted entirely to noninteracting fermions, because the analytic results of Ref. 7 permit many calculations that could not be done by numerical brute force in an interacting case, or can be carried out only by quite different methods (such as Monte Carlo). Furthermore, our hypotheti-

cal renormalization or projection algorithm using density-matrix (DM) truncation would first be applied to a well-understood system in which the low-energy excitations behave as noninteracting quasiparticles (as in a Fermi liquid or a d-wave superconducting phase). Our imagined numerical method in such a system would extract renormalized pseudocreation operators $\widetilde{f}_l^{\dagger}$ which are related to the single-particle operators $\{f_l^{\dagger}\}$ in the same way as quasiparticle creation operators are related to bare fermion creation operators. The pseudodispersion relation for the renormalized $\{\widetilde{f}_l^{\dagger}\}$ is thus expected to scale in the same fashion as the pseudodispersion φ_l encoded in Eq. (1.3) of the noninteracting fermions considered in this paper.

Through the calculations in this paper we aim to understand the analytic structure of ρ_B, and to begin to understand the quantitative errors due to truncation. This includes the question: what is the proper measure of error? The most familiar measure, the retained fraction of total density-matrix weight, does not seem to be the best measure of error, as evidenced by the small errors obtained in Sec. VI for the calculations of the dispersion relation. Another question to be investigated is whether—when we are severely truncating the Hilbert space, and attempting only to obtain the low-energy excitations—the density-matrix eigenstates are the optimal basis. As shown by the comparison in Sec. VII, they are certainly superior to the other plausible basis (plane waves).

In Sec. II A, we summarize first our much improved understanding of the analytic structure of the density matrix for noninteracting fermions (and by implication for any Fermi liquid) following from Eqs. (1.2) and (1.3), and elucidate the statistical-mechanical analogy between the density-matrix eigenstates and the many-body states of a system of noninteracting spinless fermions. In Sec. II B, we summarize the important results which we obtained in Ref. 7, giving exact relationships between the block density-matrix ρ_B, the Green-function matrix G restricted to the block, as well as their eigenstates and eigenvalues. Then in Sec. II C, we develop the main ideas behind an operator-based DM truncation scheme based on the statistical-mechanical analogy described in Sec. II A.

The relation between the single-particle density-matrix eigenstates and single-particle energy eigenstates of a system of noninteracting spinless fermions also suggests how the distribution of single-particle pseudoenergies φ_l is expected to scale with the block size B. Numerically, a scaling relationship between φ_l and B was found indeed for the overall chain at various fillings $\bar{n}$. Our analytical results from Ref. 7 shed light on this eigenvalue scaling in two ways. First, as in Ref. 8, they allow numerical study of the density matrix for system sizes so large that they would be inaccessible to any other techniques. Second, the exact connection of the block density matrix ρ_B to the block Green-function matrix G gives hints about the eigenvalue distribution. All of these will be discussed further in Sec. III, and the implications of the scaling behavior of the single-particle pseudoenergies are discussed in Sec. IV, where we derived the asymptotic behavior, in the limit of infinite block sizes, of the largest density-matrix weight and the truncated weight W_t, which is the sum of weights of the density-matrix eigenstates retained in the operator-based DM truncation scheme.

Compared to the traditional density-matrix truncation scheme used in the DMRG, our operator-based density-matrix truncation scheme gives for the same number of density-matrix eigenstates retained a slightly larger discarded weight $\epsilon = 1 - W_t$ (see Fig. 8 in Sec. IV D). This quantity gives an $O(\epsilon)$ estimate as an upper bound—a worst case scenario—for the error incurred when computing the expectation of a most general observable. As with examples in numerical integration, the performance of an algorithm in integrating some classes of functions may be much better than that expected from the straightforward error analysis. In any case, we are not really interested in arbitrary observables, but rather in the dispersion relation of elementary excitations, which we calculate in Sec. VI. The results are highly encouraging: the dispersion relation calculated in the operator-based DM truncation scheme differs from the true dispersion relation by an amount much smaller than what is suggested by the discarded weight.

Besides quasiparticle dispersion relations, real-space correlation functions are also interesting quantities to calculate, and these invariably depend on the real-space structure of the many-body ground-state wave function. Since this wave function is to be written in terms of the one-particle density-matrix eigenfunctions, it is important to understand the real space structure of these as well. The one-particle density-matrix eigenfunctions kept in our operator-based DM truncation scheme have spatial structures that are very similar to each other. In Sec. V we look into a representative one-particle density-matrix eigenfunction, the pseudo-Fermi eigenfunction, for each B, and find that they also obey a universal scaling relation. Then in Sec. VI C, we check how well such a truncated basis of one-particle density-matrix eigenfunctions can approximate the true single-particle wave function at the Fermi level, which is a plane wave with wave vector k_F. We find the approximation to be good even when less than one quarter of the one-particle density-matrix eigenfunctions are kept. This is impressive, considering the fact that in the operator-based DM truncation scheme, the number of many-particle density-matrix eigenstates thus represented by the one-particle density-matrix eigenstates retained constitutes a miniscule fraction of the total number of density-matrix eigenstates.

Finally, for systems where we know that the true single-particle wave functions are plane waves, it seems *a priori* plausible that a plane-wave-based operator-based truncation scheme might outperform the operator-based density-matrix truncation scheme. We look into this possibility in Sec. VII, and find that while there are a few aspects in calculating the dispersion relation where the operator-based plane-wave (PW) truncation scheme outperforms the operator-based DM truncation scheme, the overall performance of the PW scheme is inferior to the DM scheme. We then conclude in Sec. VIII by summarizing the important results obtained in this paper, and discuss where all these might fit into the numerical analysis of an interacting system.

II. OPERATOR-BASED DENSITY-MATRIX TRUNCATION

A. Structure of density-matrix eigenvalues and eigenstates

Because the Hamiltonian in Eq. (1.1) conserves particle number, the eigenstates of ρ_B have definite particle number, and may be grouped into various P-particle sectors, where $P = 0, 1, \ldots, B$. A consequence of our fundamental formulas (1.2) and (1.3) is that every eigenstate of ρ_B has the form

$$|\chi_L\rangle = \prod_{p=1}^{P} f_{l_p}^{\dagger}|0\rangle. \tag{2.1}$$

Each eigenstate is specified by a list of *pseudo-occupation numbers* $\{n_l\}$, where $n_{l_p} = 1$ for the factors contained in Eq. (2.1) and is zero otherwise. Furthermore, the corresponding eigenvalue, the *density-matrix weight*, is simply given by

$$w_L = \mathcal{Q}^{-1} \exp\left[-\sum_{l=1}^{B} n_l \varphi_l\right], \tag{2.2}$$

where the quantity

$$\Phi = \sum_l n_l \varphi_l \tag{2.3}$$

appearing in the exponent is the *total pseudoenergy*. In terms of the single-particle pseudoenergies $\{\varphi_l\}$, the normalization constant of the density matrix in Eq. (1.2) can be written as

$$\mathcal{Q} \equiv \sum_{\{n_l\}} \exp\left[-\sum_l n_l \varphi_l\right], \tag{2.4}$$

where the summation is over all 2^B combinations of occupations.

It is immediately clear from Eq. (2.2) that the density-matrix eigenstate of maximum weight corresponds to the minimum total pseudoenergy. This is obtained by setting $n_l = 1$ for $\varphi_l < 0$ and $n_l = 0$ for $\varphi_l > 0$. In complete analogy to the real energy of a noninteracting system of fermions, we simply fill up the single-particle pseudoenergy levels (PEL's) from the lowest up to a *pseudo-Fermi level*. The maximum density-matrix weight always turns out to occur with the block fractional filling that is closest to the bulk filling of the Fermi sea ground state. More generally, the maximum-weight state in the P-particle sector is obtained by filling the states with the P lowest single-particle pseudoenergies. Finally, it is clear that the next-highest weights, or equivalently the next-lowest total pseudoenergies in the P-particle sector, are obtained by making particle-hole excitations involving the PEL's near to the last one filled.

The above analogy may be extended to note that Eq. (1.2) is exactly the density matrix that would be obtained (see, for example, Ref. 9) at temperature $T = 1$ if $\tilde{H}$ were the Hamiltonian. The reciprocal of the normalization constant $\mathcal{Q}$ of ρ_B in Eq. (2.4) just corresponds to the grand partition function for the block of B sites. Among other things, this implies that $\langle n_l\rangle$, the average particle number in a particular PEL, has the functional form of the Fermi-Dirac distribution.

We will actually apply this idea in a slightly different way, so as to relate the single-particle pseudoenergies φ_l for different block sizes. If we were dealing with an actual Hamiltonian, the dispersion relation would imply a density of states which would be multiplied by the system size to obtain the actual distribution of states. Our numerical scaling results in Sec. III A confirm that pseudoenergies behave similarly to real energies.

B. Relation of pseudoenergies to block Green-function matrix

In our earlier work[7] we obtain an exact formula

$$\tilde{H} = -\sum_{ij} [\log_e G(\mathbb{1} - G)^{-1}]_{ij} c_i^{\dagger} c_j, \tag{2.5}$$

which, with Eq. (1.2), relates ρ_B to the block Green-function matrix G, whose matrix elements are $G_{ij} = \langle c_i^{\dagger} c_j\rangle$ with i and j restricted to sites within the block. Clearly Eq. (2.5) becomes Eq. (1.2) when the pseudo-Hamiltonian is diagonalized. Also, Eq. (2.5) tells us that the quadratic form of $\tilde{H}$ in Eq. (2.5) and G are simultaneously diagonalizable. If we denote by

$$|\chi_l\rangle = f_l^{\dagger}|0\rangle = \sum_{j=1}^{B} \chi_l(j) c_j^{\dagger}|0\rangle, \tag{2.6}$$

the single-particle eigenstate of $\tilde{H}$ with eigenvalue φ_l, then $\{\chi_l(j)\}$ is the eigenvector of G with eigenvalue λ_l. The single-particle pseudoenergies are related to the eigenvalues of G by

$$\varphi_l = -\log_e \frac{\lambda_l}{1 - \lambda_l}, \tag{2.7}$$

or, equivalently,

$$\lambda_l = \frac{1}{\exp \varphi_l + 1}, \tag{2.8}$$

i.e., the eigenvalues of G are the average pseudo-occupation numbers $\langle n_l\rangle$. Note that we sometimes write $\varphi_l \to \varphi(l, B)$ to make explicit the dependence on block size B. We will assume that φ_l are ordered from the most negative to the most positive values.

Another notable result that was derived in Ref. 7 is that

$$\mathcal{Q}^{-1} = \det(\mathbb{1} - G) = \prod_{l=1}^{B} (1 - \lambda_l). \tag{2.9}$$

Along with Eqs. (2.5) and (2.7), Eq. (2.9) comprises the final ingredients that allow numerical computation of the density-matrix even in very large blocks. Aside from the possibilities of truncation, Eqs. (2.1)–(2.9) have completely reduced a $2^B \times 2^B$ diagonalization problem into a $B \times B$ problem, a computational shortcut which allows numerical studies of large blocks.

C. Recipe for operator-based truncation

The analytical structure of Eq. (1.3) hints at the proper design of a truncation scheme. The retained Hilbert space of a block *should not* be the span of those density-matrix eigenstates whose weight exceeds a cutoff. Instead, we should implement a "consistent" truncation, such that the truncated Hilbert space consists of exactly $2^{l_{\max}}$ states, built from all combinations of $l_{\max}$ pseudocreation operators $f_1^\dagger, \ldots, f_{l_{\max}}^\dagger$, acting on a block "vacuum state" $|0\rangle_B$ and satisfying fermion anticommutation relations.

In the traditional density-matrix-based truncation scheme used in DMRG, the recipe for truncation is to sort all density-matrix weights in descending order, and then retain only the eigenstates associated with the weights that exceed a certain cutoff. Let us refer to this as the *weight-ranked* DM truncation scheme. In light of our understanding of the structure of the many-body density-matrix presented in this paper, we can see that the weight-ranked DM truncation scheme will certainly retain the eigenstate with maximum weight, the pseudo-Fermi sea described in Sec. II A, along with eigenstates that are "particle excitations," "hole excitations" and "particle-hole excitations" from the pseudo-Fermi sea. If we arrange the entire collection of many-particle density-matrix eigenstates into a state graph, then the state graph looks like a B-dimensional hypercubic lattice near the pseudo-Fermi sea. What the weight-ranked DM truncation does in this state graph picture is to remove nodes, and in effect cut bonds, out from this hypercubic lattice, producing a subgraph that is much less connected and containing tenuous links. Because of this, when the Hamiltonian is projected onto the weight-ranked DM truncated basis, spurious interactions are introduced.

We can apply the pseudoenergy analogy in choosing how to truncate, given the form of the density matrix. It is familiar, in the truncation used in Fourier-space-based quantum renormalization groups, to discard all single-particle degrees of freedom except for a shell around the Fermi surface. In the same way, let us discard all operators f_l as degrees of freedom, except those for which $|\varphi_l|$ is less than some threshold φ_*. For all other f_l, we "freeze" n_l by setting n_l to its ground state value,

$$n_l = \begin{cases} 1 & \text{for } \varphi_l < -\varphi_* \\ 0 & \text{for } \varphi_l > \varphi_*. \end{cases} \quad (2.10)$$

This choice gives the maximum density-matrix weight, among the eigenstates having any particular set of n_l for the retained single-particle pseudoenergy levels. The spirit of this truncation scheme is similar to that used in quantum chemistry,[10–12] except that the notion of a Fermi surface is more fuzzy in atoms and molecules. The idea that truncation consists of decreasing the thickness of a shell of wave vectors around the Fermi surface appeared in the original renormalization group (RG) for a quantum-mechanical solid-state problem (Anderson's poor man's RG for the Kondo problem[13]). This obvious notion—that the action is around the Fermi surface—necessarily appears in every effective form of truncation intended for a metal (see, for example,

Refs. 14 and 15, among others). However, to our knowledge all such schemes used a basis of plane waves or of energy eigenstates. Our variation uses PEL's in analogy to the use of energy eigenstates in these previous problems. Deriving from the density matrix, it makes sense only in procedures that involve cutting a real-space block out of a larger system.

Within this operator-based DM truncation scheme, we can define an effective Hamiltonian for the truncated Hilbert space, just by taking the matrix elements of the true Hamiltonian between all the retained states. Using an operator-based truncation, this will have a particularly clean form: first replace each creation operator $c_j^\dagger$ by the equivalent combination of all $\{f_l^\dagger\}$; then replace $f_l^\dagger f_l \to 1$ for each single-particle pseudoenergy $\varphi_l < -\varphi_*$ (these are frozen to be always occupied), and otherwise remove all terms involving the operators that are truncated. Thus, if the original Hamiltonian has at most $l_{\max}$-fermion terms, the same will be true for the truncated Hamiltonian. This prescription shows that such a truncation is possible for general models, once one knows the appropriate density matrix, but in this paper we have applied it only to noninteracting models.

III. SCALING BEHAVIOR OF EIGENVALUE DISTRIBUTION

Since the many-particle density-matrix eigenvalues are built, according to Eq. (2.2), from the single-particle pseudoenergies, the latter are the focus of our numerical investigations. Now if our entire system consisted of the block in a pure state at $T=0$, then every eigenvalue λ_l of the block Green-function matrix G, being the average pseudo-occupation number $\langle n_l \rangle$ of a PEL, would either be one or zero. At $T>0$, λ_l follow a Fermi-Dirac distribution. We will see later that cutting out a finite block from a $T=0$ system, by tracing over the environment of the block, has a similar effect on the eigenvalues of G as would taking $T>0$ when the block is the whole system.

In a translationally invariant system with filling $\bar{n}$ (at $T=0$), a fraction $\bar{n}$ of the eigenvalues of G are one, while the rest are zero. Cutting out a block of length B must smooth out this step (much as having $T>0$ makes it into a Fermi-Dirac distribution), and we expect the transition from one to zero to occur over a fraction $\sim 1/B$. This guess was inspired by the analogy of pseudoenergy φ_l in Eq. (1.3) to the real energy, which near the Fermi level scales linearly with wave vector $\sim 1/B$. This B^{-1} scaling suggests the conjecture of a scaling form for the single-particle pseudoenergy like $\varphi_l \approx B f(l/B)$, and indeed we find below just such a scaling form.

A. Pseudoenergies and pseudo-occupation numbers

In this section we calculate numerically the eigenvalues λ_l of the block Green-function matrix G, and use Eq. (2.7) to compute the single-particle pseudoenergies φ_l. For a chain of free spinless fermions in its ground state, the matrix elements of the block Green-function matrix G are
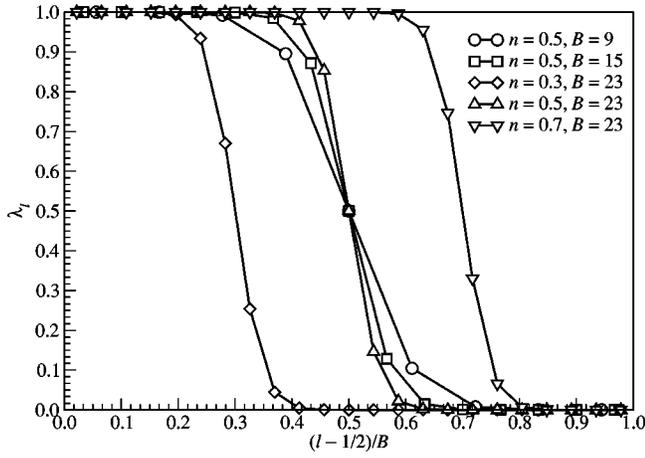
FIG. 1. Distribution of λ_l for different filling fractions and block sizes. In order to compare λ_l for different block sizes, the interval $l \in [1,B]$ is rescaled such that $(l-\frac{1}{2})/B \in (0,1)$. With this rescaling, $\lambda = \frac{1}{2}$ always occurs at $[(l-\frac{1}{2})/B] = \bar{n}$.

$$G_{ij} = \frac{\sin \pi \bar{n} |i-j|}{\pi |i-j|}, \tag{3.1}$$

where $\bar{n}$ is the filling fraction. Figure 1 shows how λ_l, the eigenvalues of G, are distributed for different filling fractions $\bar{n}$ and different block sizes B.

For $\bar{n} = \frac{1}{2}$, our numerical studies suggest a scaling relationship of the form

$$\varphi(l,B) \cong Bf(x), \tag{3.2}$$

where $x \equiv [(l-\frac{1}{2})/B] - \frac{1}{2}$, as shown in Fig. 2. There are two points on Fig. 2 we would like to note. First of all, with our choice of the scaling variable x, the scaling function $f(x)$ always passes through the origin, i.e.,
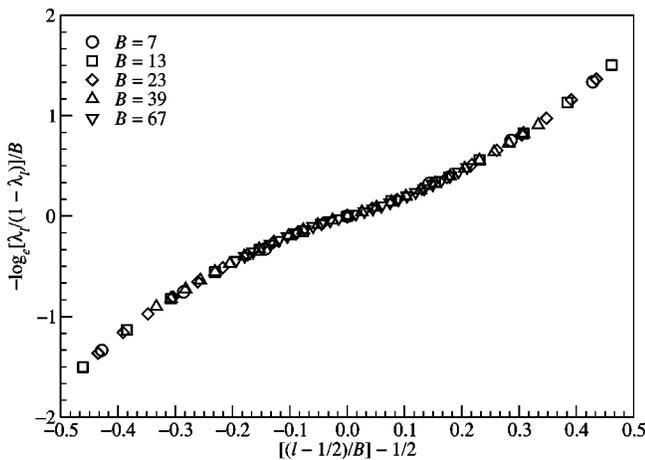

FIG. 2. Plot of $-B^{-1}\log_e[\lambda_l/(1-\lambda_l)]$ as a function of the scaling variable $x \equiv [(l-\frac{1}{2})/B] - 1/2$ for various block sizes at half filling, showing a scaling collapse onto the scaling function $f(x)$. For $B > 23$, the largest and smallest pseudoenergies are severely affected by numerical truncation errors in the diagonalization routines, and thus not shown.
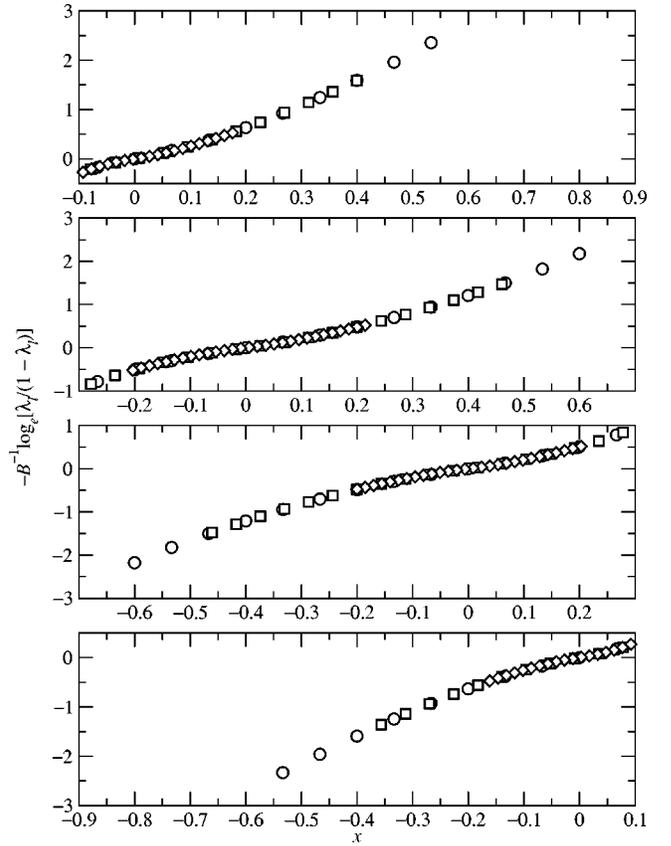

FIG. 3. Scaling collapses for (from top to bottom) $\bar{n} = 0.1$, 0.3, 0.7, and 0.9, plotted against the scaling variable $x \equiv (l-\frac{1}{2})/B - \bar{n}$, for three different block sizes: $B = 15$ ($\bigcirc$), $B = 23$ ($\square$), and $B = 67$ ($\diamond$).

$$f(0) = 0. \tag{3.3}$$

Second, from Fig. 2, we see that $f(x)$ is an odd function, i.e.,

$$f(-x) = -f(x), \tag{3.4}$$

which is what we would expect from particle-hole symmetry when the overall system is at half filling, and $f(x)$ has a finite positive slope at $x = 0$, i.e.,

$$f'(0) > 0. \tag{3.5}$$

Similar scaling behaviors of the form

$$\varphi(l,B,\bar{n}) \cong Bf(\bar{n},x) \tag{3.6}$$

are found for all $\bar{n}$, with the generic scaling variable

$$x \equiv (l-l_F)/B, \tag{3.7}$$

where

$$l_F = \bar{n}B + \frac{1}{2} \tag{3.8}$$

plays the role of a Fermi wave vector, and a filling fraction-dependent scaling function $f(\bar{n},x)$, as shown in Fig. 3. The scaling functions continue to satisfy

$$f(\bar{n},0)=0, \tag{3.9}$$

and

$$f'(\bar{n},0)>0, \tag{3.10}$$

but $f(\bar{n},x)$ is no longer an odd function of x for $\bar{n} \neq \frac{1}{2}$. Instead, the particle-hole symmetry inherent in our model is manifested as

$$f(\bar{n},-x)=-f(1-\bar{n},x). \tag{3.11}$$

From Eqs. (2.8) and (3.6), we can write λ_l as

$$\lambda_l = \frac{1}{\exp(Bf)+1}, \tag{3.12}$$

which tells us that $f(\bar{n},x)$ plays the role of a dispersion relation $\epsilon(k)$, while B plays the role of the inverse temperature β. This confirms our suspicion that the effect of cutting a block out of an overall system in its ground state at $T=0$ is to ascribe to the block an effective temperature. As expected, this effective temperature approaches zero as the block size is increased, since we are keeping more and more information about the overall system, which we know to be at $T=0$.

B. Normalization constant

Having understood that λ_l is related to the scaling function $f(\bar{n},x)$ as in Eq. (3.12), we are now ready to investigate the scaling behavior of the normalization constant Q^{-1}, which is related to λ_l by Eq. (2.9). We do this first at half filling. As can be seen from Fig. 1, at half filling roughly half of the λ_l are approximately one, whereas the other half are approximately zero. The product $\Pi_l(1-\lambda_l)$ is therefore determined primarily by the $\sim B/2$ λ_l's that are nearly one. For these eigenvalues, $\exp[Bf(x)] \ll 1$ and thus $1-\lambda_l \approx \exp[Bf(x)]$ [when it is clear what the filling fraction is, we will drop the argument $\bar{n}$ in $f(\bar{n},x)$ to keep the notations neat]. With this, we find that

$$Q^{-1}=\prod_{l=1}^{B}(1-\lambda_l) \approx \prod_{l=1}^{B/2}\exp[Bf(x)]$$

$$\approx \exp\left\{B\int_{-1/2}^{0}f(x)\,dx\right\}$$

$$=\exp\left\{-B\int_{0}^{1/2}f(x)\,dx\right\}, \tag{3.13}$$

where we made use the observed odd symmetry of the scaling function in Eq. (3.4) when the overall system is at half filling, so that the integral within the exponent is positive. From Eq. (3.13), we see that Q^{-1} decreases exponentially with block size B.

In general, for $\bar{n}$ not too close to zero, where the argument that those λ_l's that are near one make the dominant contribution holds, we find that

$$Q^{-1} \approx \exp\left\{B\int_{-\bar{n}}^{0}f(\bar{n},x)\,dx\right\}=\exp\left\{-B\int_{0}^{\bar{n}}f(1-\bar{n},x)\,dx\right\}, \tag{3.14}$$

where we made use of Eq. (3.11). For $\bar{n}$ very close to zero, there are a handful of λ_l's of $O(1)$, and the rest are all nearly zero, behaving like $\lambda_l \approx \exp[-Bf(\bar{n},x)]$. Ignoring these handful of $O(1)$ λ_l's, we find that the contribution to Q^{-1} from those $\lambda_l \ll 1$ is proportional to the product $\Pi_l(1-\lambda_l) \approx 1 - \Sigma_l \lambda_l$, and so

$$Q^{-1} \propto \left[1-\int_{0}^{1-\bar{n}}e^{-Bf(\bar{n},x)}\,dx\right]. \tag{3.15}$$

The integral can be evaluated as a cumulant expansion, but we can already see that for large B, the integral will not be important, and thus Q^{-1} derives most of its value from the few $O(1)$ λ_l's. In contrast, when $\bar{n}$ is very close to 1, then most of the eigenvalues λ_l of G are close to 1, and these continue to dominate the product $\Pi_l(1-\lambda_l)$, and the asymptotic formula derived in Eq. (3.13) continues to be valid.

IV. DENSITY-MATRIX WEIGHTS: IMPLICATIONS OF EIGENVALUE SCALING

With our understanding of the structure of the many-particle density-matrix eigenvalues and eigenstates developed in Sec. II A, and on the scaling behavior of the single-particle found in Sec. III, we want to now address the question of how much of the Hilbert space we can truncate. Clearly, the answer to this question depends on what measure of error we intend to use as our criteria for judging how well the truncated Hilbert space describes the physics associated with the parent model. In this section we look at the most common measure of error, used in the DMRG (Refs. 1 and 2) and quantum chemistry calculations:[10–12] for a properly normalized density matrix, the density-matrix weights w_L satisfy the sum rule

$$\sum_L w_L = 1. \tag{4.1}$$

If the ordinal numbers L are chosen such that w_L is ranked in decreasing order, and a total of L_{max} density-matrix eigenstates are retained, then the truncated weight

$$W_t=\sum_{L \leqslant L_{max}}w_L \leqslant 1, \tag{4.2}$$

and the discarded weight

$$\epsilon=1-W_t \tag{4.3}$$

are frequently used as figures of merit for the truncation scheme, since for a bounded operator A, the truncation error in $\langle A \rangle$ is $O(\epsilon)$.
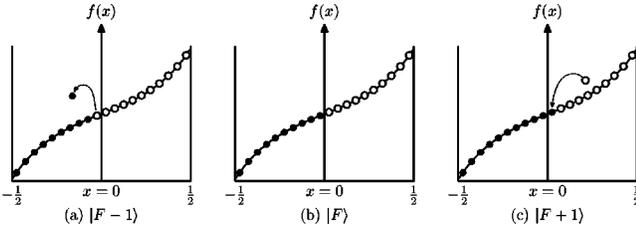
FIG. 4. Schematic diagram showing the three many-particle density-matrix eigenstates, (a) $|F-1\rangle$, (b) $|F\rangle$, and (c) $|F+1\rangle$, with the largest weights, for a block of B (B even) sites within an overall system that is half filled.

A. The gapless chain of noninteracting spinless fermions

Instead of diving in to look at W_t or ϵ, let us consider first a related question: how large is the maximum weight for a block of B sites embedded within an overall system of gapless noninteracting spinless fermions? For our discussions, we will consider the half-filled case; it will be straightforward to extend the arguments presented below to $\bar{n} \neq \frac{1}{2}$. For convenience, let us take B to be even.[19] Let us denote by $|F\rangle$, where $F = B/2$, the many-particle density-matrix eigenstate having the largest weight. This state is always kept in the operator-based truncation scheme. Recall from Sec. II that, as shown in Fig. 4(b), this is the analog among density-matrix eigenstates of the Fermi sea ground state among energy eigenstates. In this $B/2$-particle state, the single-particle pseudoenergy is filled up to just before $x=0$, which we shall call the pseudo-Fermi level. The many-particle density-matrix eigenstates of the block, of which there are two, having the next largest weights will be called $|F-1\rangle$ [Fig. 4(a)] and $|F+1\rangle$ [Fig. 4(c)], having, respectively, one less or one more particle.

We can understand the weight ratio w_{F-1}/w_F or w_{F+1}/w_F as follows. By Eq. (2.2), $|F\rangle$ is the state with $n_l = 1$ for $l=1,\ldots,B/2$ and $n_l=0$ for $l=B/2+1,\ldots,B$. The state $|F+1\rangle$ differs only in having $n_{B/2+1}=1$, while $|F-1\rangle$ differs only in having $n_{B/2}=0$. Near the pseudo-Fermi level, the scaling function has a slope of $f'(0)$, while the spacing between adjacent pseudoenergies on the rescaled l/B axis is $1/B$. Thus $\varphi_{B/2+1} - \varphi_{B/2} \approx f'(0)$. But when the actual filling is $\bar{n} = \frac{1}{2}$, we know by particle-hole symmetry that $\varphi_{B/2+1} = -\varphi_{B/2}$, so $\varphi_{B/2} \approx -f'(0)/2$ and $\varphi_{B/2+1} \approx f'(0)/2$. It follows from Eq. (2.2), that

$$\frac{w_{F+1}}{w_F} = \frac{w_{F-1}}{w_F} \approx \exp[-f'(0)]. \quad (4.4)$$

For $\bar{n} \neq \frac{1}{2}$, Eq. (3.6) would tell us that these ratios are approximately $\exp[-f'(\bar{n},0)]$.

This is quite different from what would happen when the block contains half of the entire system, as considered in the standard DMRG algorithm, or in Ref. 8. If the fraction (B/N) in the block approached one, the state $|F\rangle$ must become the Fermi sea ground state of the overall system, and consequently contain all the weight. If the block is merely a finite fraction of the system, we still expect a much larger ratio than Eq. (4.4). It would be interesting to check the

behavior of the ratios in Eq. (4.4) for the case $B,N \rightarrow \infty$ with $B/N=1/2$, but we have not investigated this.

B. The gapped chain of noninteracting spinless fermions

For the purpose of understanding the pseudoenergy spectrum of noninteracting systems better, we also considered the dimerized tight-binding Hamiltonian

$$H = -t \sum_{j=1}^{N} [1 + (-1)^j \delta](c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j), \quad (4.5)$$

where the hopping integral t is modulated by the $(-1)^j\delta$ term to produce an energy gap. Henceforth we choose the scale of energy to be such that $t=1$. This system was solved analytically by Gebhard et al.,[16] wherein the Hamiltonian can be written as

$$H = \sum_{|k|<\pi/2a} E(k)(a_{k,+}^\dagger a_{k,+} - a_{k,-}^\dagger a_{k,-}), \quad (4.6)$$

with $E(k) = \sqrt{\epsilon^2(k) + \Delta^2(k)}$, where $\epsilon(k) = -2\cos k$ and $\Delta(k) = 2\delta \sin k$. In terms of $\Delta(k)$ and $\epsilon(k)$, we can define an angle ϕ_k such that $\tan 2\phi_k = \Delta(k)/\epsilon(k)$, and whose sine and cosine we denote as $\alpha_k = \cos \phi_k$ and $\beta_k = \sin \phi_k$. In terms of these, the operators $a_{k,+}$ and $a_{k,-}$ for the upper and lower bands, respectively, are given by

$$a_{k,-} = \alpha_k \tilde{c}_k + i\beta_k \tilde{c}_{k+\pi},$$

$$a_{k,+} = -\beta_k \tilde{c}_k + i\alpha_k \tilde{c}_{k+\pi}. \quad (4.7)$$

At half filling, the lower band is completely filled while the upper band is completely empty, and the ground state can be written simply as

$$|\Psi\rangle = \prod_{|k|<\pi/2} a_{k,-}^\dagger |0\rangle. \quad (4.8)$$

For this ground state, the two-point function is given by

$$G_{ij} = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} dk \, e^{ik(i-j)} \frac{\cos k - i(-1)^i \delta \sin k}{\sqrt{\cos^2 k + \delta^2 \sin^2 k}}, \quad (4.9)$$

using which we can construct the block Green-function matrix G, and hence using Eq. (2.7) the pseudoenergies which correspond to the density-matrix eigenvalues. For a fixed block size of $B=23$, the pseudoenergy spectra for different hoping modulations δ are shown in Fig. 5, compared to that of the gapless case. Scaling behavior of the single-particle pseudoenergies was found for all δ, each governed by a different scaling function $f(\bar{n},\delta,x)$. The scaling collapse plot for $\delta=0.5$ is shown in Fig. 6, compared to the scaling function $f(\bar{n},\delta=0,x)$ for the gapless case.

Repeating our analysis for the three density-matrix eigenstates with the largest weights, we find again that the ratios of density-matrix weights w_{F+1}/w_F and w_{F-1}/w_F to be independent of block size B when the overall system is at half filling. However, these ratios depend strongly on the hopping modulation δ. As we can see from Fig. 5, the slope of the
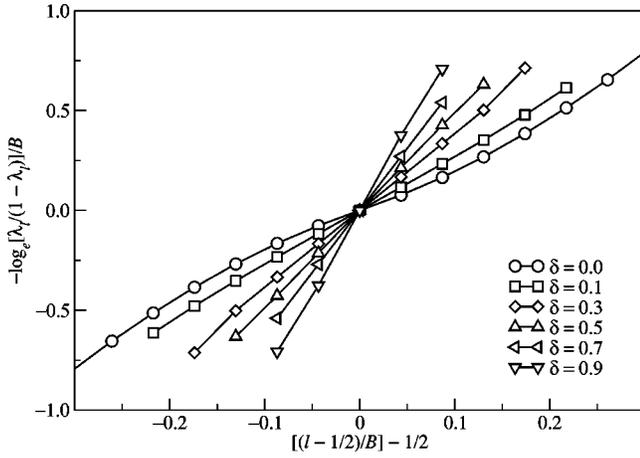
FIG. 5. Plot of $-B^{-1}\log_e[\lambda_l/(1-\lambda_l)]$ as a function of the scaling variable $x=(l-\frac{1}{2})/B-\frac{1}{2}$, for a block of size $B=23$, with different hoping modulations $\delta=0.1,0.3,0.5,0.7,0.9$. Pseudoenergies for $|x|>0.2$ for $\delta>0$ are not shown because these are severely affected by numerical errors incurred in the numerical integration and diagonalization routines. The various sets of straight line segments are intended to guide the eye in visualizing the data.

scaling curve at $x=0$ is steeper for larger δ. This indicates that—everything else being equal for finite B—a smaller fraction of density-matrix states is needed to capture the same total weight if the system is gapped.

We have not investigated the case $B/N\rightarrow1/2$, as in the standard DMRG algorithm, but we naturally expect the ratio to increase in a gapped system. Thus $|F\rangle$ would be a better approximation to the ground state in a gapped system than in a gapless system, which is known as an empirical fact in the DMRG context. Our approach, if extended to the case $B/N>0$, would give an analytic justification for this common observation.
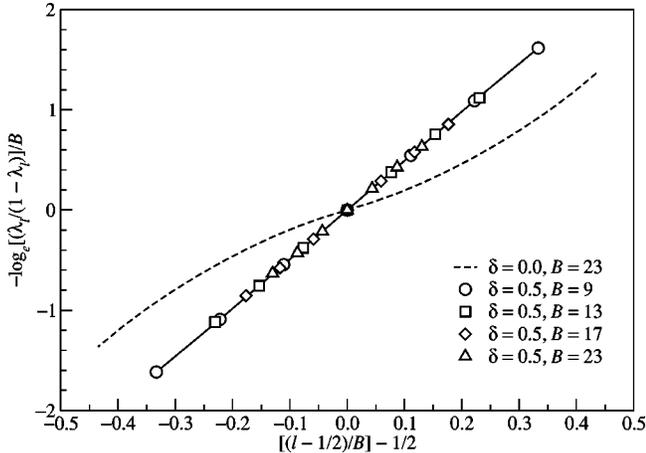

FIG. 6. Plot of $-B^{-1}\log_e[\lambda_l/(1-\lambda_l)]$ as a function of the scaling variable $x=(l-\frac{1}{2})/B-\frac{1}{2}$, for hoping modulation $\delta=0.5$ and different block sizes $B=9,13,17,23$. The pseudoenergies for $|x|>0.3$ are not shown as these are severely affected by numerical errors in the numerical integration and diagonalization routines. Also shown is an approximate dashed curve for the scaling function $f(x)$ for $\bar{n}=\frac{1}{2}$ and $\delta=0$, obtained from the data for $B=23$. The straight line segments are intended to guide the eye in visualizing the data.
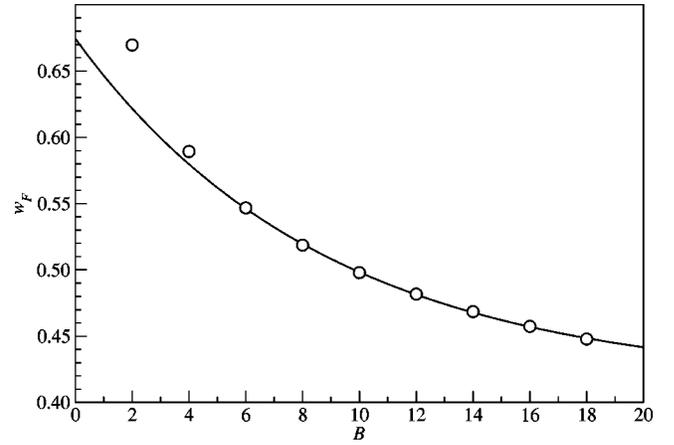

FIG. 7. Plot of the largest density-matrix weight w_F as a function of the block size B for B even. The solid curve shown is a fit of the form $w_F=w_{F,\infty}+\Delta w_F\exp(-B/B_0)$. The best fit to this small data set is obtained by neglecting the data points for $B=2$ and $B=4$, for which we get $w_{F,\infty}=0.41$, $\Delta w_F=0.26$, and $B_0=0.11$.

C. Largest density-matrix weight

For even B blocks on a gapless chain of noninteracting spinless fermions described the Hamiltonian (1.1), the largest density-matrix weight w_F can be numerically computed reliably till $B\approx20$, and its dependence on B is shown in Fig. 7. Also shown in Fig. 7 is a fit of the numerical data to

$$w_F(B)=w_{F,\infty}+\Delta w_F\exp(-B/B_0), \qquad (4.10)$$

where $w_{F,\infty}$, Δw_F, and B_0 are curve-fitting parameters. Here the exponentially decaying term is merely chosen to produce a good curve fit—we believe the B dependence may be more complex—but what is interesting is the fact that w_F tends to a constant $w_{F,\infty}$ in the limit of $B\rightarrow\infty$. We find that we can understand this in terms of the scaling formulas developed so far.

From Sec. II we saw that the largest many-particle density-matrix weight w_F corresponds to the situation for which all PEL's below the pseudoFermi level $\varphi_F=0$ are occupied, and all those above are empty. This means that

$$w_F=\mathcal{Q}^{-1}\prod_{l<l_F}e^{-\varphi_l}. \qquad (4.11)$$

Using the fact that $\mathcal{Q}$ can be written explicitly as

$$\mathcal{Q}=\prod_l(1+e^{-\varphi_l}), \qquad (4.12)$$

we then find that

$$w_F=\prod_{l<l_F}\frac{e^{-\varphi_l}}{1+e^{-\varphi_l}}\prod_{l>l_F}\frac{1}{1+e^{-\varphi_l}}=\prod_l\frac{1}{1+e^{-|\varphi_l|}}. \qquad (4.13)$$

To evaluate w_F, we evaluate first its logarithm, which is

$$-\log_e w_F=\sum_l\log_e(1+e^{-|\varphi_l|}). \qquad (4.14)$$

Here we make two approximations. First, because of Eq. (3.6), we know that $\varphi_l \propto B$, and so except for a handful of single-particle pseudoenergies φ_l very near to φ_F, all the exponentials are very small numbers. Using the approximation $\log_e(1+x) \approx x$ for $x \ll 1$, we write Eq. (4.14) as

$$-\log_e w_F \approx \sum_l e^{-|\varphi_l|}. \tag{4.15}$$

Second, we note that because of Eq. (3.6), single-particle pseudoenergies far away from $\varphi_F = 0$ will contribute negligibly to the above sum. For B sufficiently large, those single-particle pseudoenergies making significant contribution in Eq. (4.14) will lie within a small interval about l_F where a linear approximation of the form

$$\varphi_l \approx B f'(0) \frac{l - l_F}{B} = f'(0)(l - l_F) \tag{4.16}$$

adequately describes the pseudodispersion relation. Substituting Eq. (4.16) into Eq. (4.15), we find then that

$$-\log_e w_F \approx \sum_{l=1}^{B} e^{-f'(0)|l - l_F|} \approx 2\sum_{l>l_F}^{\infty} e^{-f'(0)(l - l_F)}. \tag{4.17}$$

This is a geometric series which we can readily sum to give

$$-\log_e w_F = \frac{2}{1 - \exp(-f'(0))}, \tag{4.18}$$

i.e., the largest density-matrix weight w_F is found to approach a constant value of

$$w_F = \exp\left[-\frac{2}{1 - \exp[-f'(0)]}\right] \tag{4.19}$$

as $B \to \infty$. From Figs. 2 and 3, we see that $f'(0) \approx 5$, and so this asymptotic value of w_F is ≈ 0.13. This is smaller than the $w_{F,\infty} = 0.41$ found numerically.

D. Discarded weight

Now that we understand more about the scaling behavior of the largest density-matrix weight w_F, let us analyze the discarded weight incurred by the operator-based DM truncation scheme. We compute numerically the discarded weight incurred by the operator-based DM truncation scheme and that incurred by the weight-ranked DM truncation scheme, and show them in Fig. 8 as a function of the number of many-body states kept as a comparison. As we can see, the discarded weight incurred by the operator-based DM truncation scheme is larger compared to the weight-ranked DM truncation scheme, for the same number of many-particle eigenstates kept. This is expected, since the weight-ranked DM truncation scheme is by definition the most efficient scheme in exhausting the sum rule given in Eq. (4.1). In spite of this seemingly poorer "convergence" property, we believe that the operator-based truncation scheme has advantages

that cannot be reproduced by the weight-ranked truncation scheme, to be argued in detail in Sec. VI.

Writing the total density-matrix weight explicitly as

$$W = \mathcal{Q}^{-1} \prod_{\text{kept}} (1 + e^{-\varphi_l}) \prod_{\text{below}} (1 + e^{-\varphi_l}) \prod_{\text{above}} (1 + e^{-\varphi_l}), \tag{4.20}$$

where the subscript "kept," "below," and "above" refer to PEL's retained, approximated as always occupied, and approximated as always empty in the operator-based truncation scheme respectively. The truncated weight W_t calculated from the operator-based truncation scheme is

$$W_t = \mathcal{Q}^{-1} \prod_{\text{kept}} (1 + e^{-\varphi_l}) \prod_{\text{below}} e^{-\varphi_l} \prod_{\text{above}} 1. \tag{4.21}$$

Since $W = 1$, the ratio $W_t/W = W_t$ can be written as

$$W_t = \prod_{\text{below}} \frac{e^{-\varphi_l}}{1 + e^{-\varphi_l}} \prod_{\text{above}} \frac{1}{1 + e^{-\varphi_l}}$$

$$= \prod_{\text{below}} \frac{1}{1 + e^{\varphi_l}} \prod_{\text{above}} \frac{1}{1 + e^{-\varphi_l}}. \tag{4.22}$$

This expression has a simple interpretation in terms of the pseudo-occupation numbers $\{\lambda_l\}$. Using Eq. (2.8), we find that we can write W_t as

$$W_t = \prod_{\text{below}} \lambda_l \prod_{\text{above}} (1 - \lambda_l), \tag{4.23}$$

i.e., the truncated weight W_t is given by the product of pseudo-occupation numbers λ_l of those PEL's we insist are always occupied, together with the product of the single-hole pseudo-occupation numbers $(1 - \lambda_l)$ of those PELs we insist are always empty. From Fig. 1 we see that λ_l changes fairly rapidly from $\lambda_l \lesssim 1$ to $\lambda_l \gtrsim 0$, over a small range of PEL's. Therefore, it appears that there is a fairly large range of l's for which λ_l is very close to one or very close to zero. However, this does not mean that we should perform an
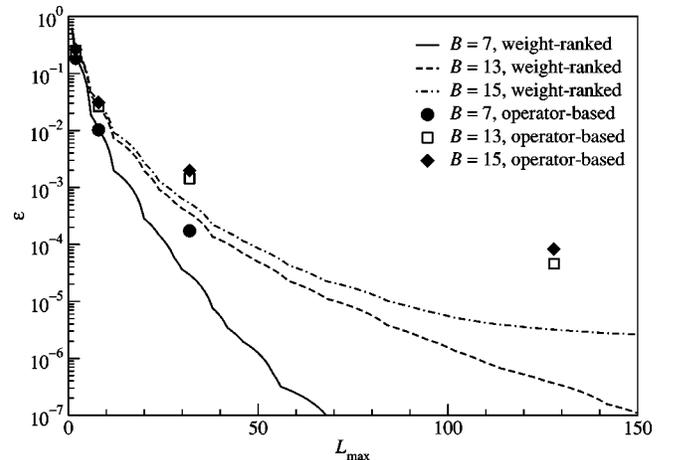

FIG. 8. Discarded weight ϵ as a function of the number of states kept: weight ranked (as done in the DMRG) and operator based.

operator-based truncation scheme keeping only the small number of PEL's whose λ_l's are significantly different from one or zero. This is because W_t is bounded from above by

$$W_t \leq \prod_{\text{below}} \lambda_{\max} \prod_{\text{above}} (1 - \lambda_{\min}) \leq (\lambda^*)^{(1-\gamma)B}, \quad (4.24)$$

where γ is the fraction of PELs retained in the operator-based truncation scheme, and

$$\lambda^* = \max(\lambda_{\max}, 1 - \lambda_{\min}). \quad (4.25)$$

Because the exponent is $O(B)$, this number can still be very small.

This brings us to the question we posed in the beginning of this section: how much of the Hilbert space do we truncate? If W_t is the only criterion then we see that a compromise is necessary. For a small block, the number of PEL's with λ_l significantly different from one or zero is a sizeable fraction of the total number of PEL's, but this number is manageable. For a large block, the number of PEL's with λ_l significantly different from one or zero is a tiny fraction of the total number of PEL's, but we still need γ to be reasonably large for W_t to be appreciable in magnitude. This of course means that an unmanageably large number of PEL's has to be retained.

To make the above discussions more watertight, let us make use of the scaling relations obtained thus far to find a formula relating the truncated weight W_t to both the block size B and the fraction γ of PEL's retained. Taking the logarithm of Eq. (4.22) we find, using the fact that $\exp(-|\varphi_l|) \ll 1$ for l far below l_F, and $\exp(-\varphi_l) \ll 1$ for l far above l_F, that

$$-\log_e W_t = \sum_{\text{below}} \log_e(1 + e^{\varphi_l}) + \sum_{\text{above}} \log_e(1 + e^{-\varphi_l})$$

$$\approx \sum_{\text{below}} e^{\varphi_l} + \sum_{\text{above}} e^{-\varphi_l}$$

$$= \sum_{l < l_F} e^{\varphi_l} - \sum_{l = l_F - \gamma B/2}^{l_F} e^{\varphi_l} + \sum_{l > l_F} e^{-\varphi_l}$$

$$- \sum_{l_F}^{l_F + \gamma B/2} e^{-\varphi_l}$$

$$= \Phi^* - \sum_{l = l_F - \gamma B/2}^{l_F} e^{\varphi_l} - \sum_{l_F}^{l_F + \gamma B/2} e^{-\varphi_l}, \quad (4.26)$$

where Φ^* is a constant.

If B is large and γ small, then the linear approximation (4.16) for φ_l is valid, in which case the two sums in Eq. (4.26) are equal and given by

$$\sum_{l_F}^{l_F + \gamma B/2} e^{-\varphi_l} \approx \sum_{l_F}^{l_F + \gamma B/2} e^{-f'(0)(l - l_F)} = \frac{1 - \exp[-\gamma B f'(0)/2]}{1 - \exp[-f'(0)]}. \quad (4.27)$$

With this, we can write W_t as

$$W_t \approx W^* \exp\left[\frac{2}{1 - e^{-f'(0)}} (1 - e^{-\gamma B f'(0)/2}) \right], \quad (4.28)$$

where $W^* = \exp(-\Phi^*)$. We can find W^* by taking the limit $\gamma \to 0$, in which case we retain no degree of freedom in the PEL's. Within the operator-based truncation scheme, this means that we insist all PEL's below φ_F to be always occupied and all those above φ_F to be always empty, i.e., only the density-matrix eigenstate with the largest weight is retained, and we should have

$$W_t = w_F = W^*, \quad (4.29)$$

and so

$$W_t \approx w_F \exp\left[\frac{2}{1 - e^{-f'(0)}} (1 - e^{-\gamma B f'(0)/2}) \right]. \quad (4.30)$$

This can be simplified further, using Eq. (4.19) to get

$$W_t \approx \exp\left[-\frac{2}{1 - e^{-f'(0)}} e^{-\gamma B f'(0)/2} \right]. \quad (4.31)$$

In the limit of $\gamma \to 1$, we see from the above expression that W_t does not tend to one, but we understand that this is because the linear approximation (4.16) is only valid for a small range of PEL's about φ_F, i.e., only for small γ. In this regime, we may further approximate W_t as

$$W_t \approx \exp\left[-\frac{2}{1 - e^{-f'(0)}} \left(1 - \frac{f'(0) \gamma B}{2} \right) \right]$$

$$\approx w_F \exp\left(\frac{f'(0)}{1 - e^{-f'(0)}} l_{\max} \right), \quad (4.32)$$

where

$$l_{\max} = \gamma B \quad (4.33)$$

is the number of PEL's retained. As we can see, for small γ, the truncated weight W_t increases exponentially with $l_{\max}$. Also, whenever Eq. (4.32) is valid, we get approximately the same truncated weight W_t whether we use $B = 100$ and $\gamma = 0.2$ or $B = 200$ and $\gamma = 0.1$. We will see in Sec. VI that whenever the retained γB PEL's lie within the regime where the pseudodispersion relation is linear, the truncation errors are essentially determined by $l_{\max} = \gamma B$.

V. SINGLE-PARTICLE DENSITY-MATRIX EIGENFUNCTIONS

A. *A Priori* expectations

As noted already in Sec. II C, in the many-body eigenstates with largest weights, all the very negative PEL's will be occupied and all the very positive PEL's will be empty. The only PEL's with significant varying occupancy are those near the pseudo-Fermi level.

By construction, the many-body density-matrix eigenstates with large weights constitute the likely configurations of the block. The difference between the large-weight eigen-
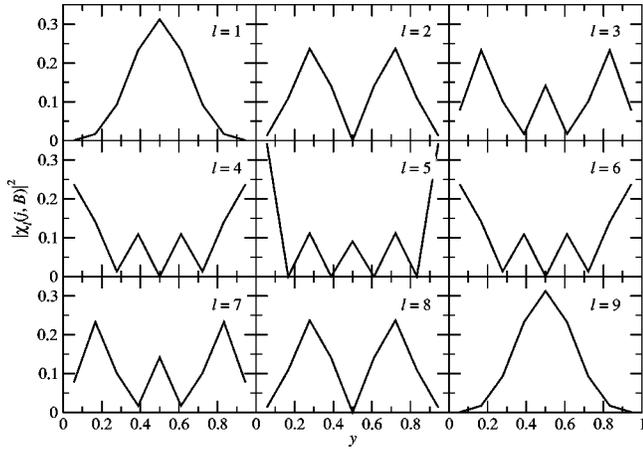
FIG. 9. Probability density of the normalized one-particle eigenfunctions $\chi_l(j)$ [plotted against the scaling variable $y=(j-\frac{1}{2})/B$] of ρ_B on a block of $B=9$ sites at half filling, showing the particle-hole symmetry of the overall system. The subplots are arranged in order of increasing pseudoenergy.
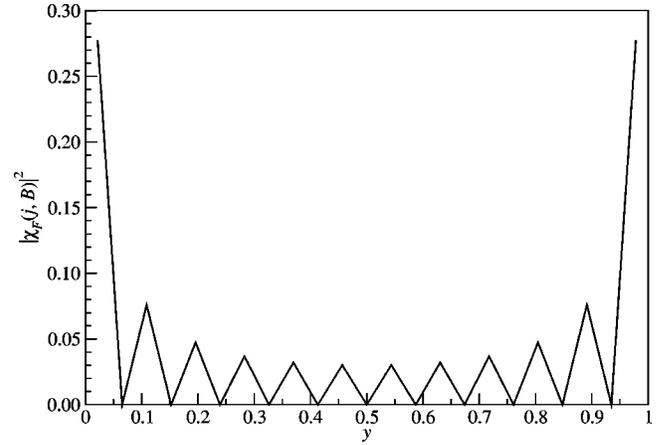

FIG. 10. Probability density function $|\chi_F(j,B)|^2$ of the pseudo-Fermi eigenfunction for $B=23$, plotted against the scaled variable $(j-\frac{1}{2})/B$.

states of the P-particle and $(P+1)$-particle sectors of the density-matrix is in the application of a creation operator $f_l^\dagger$ such that the pseudoenergy $|\varphi_l|$ is close to $\varphi_F=0$. In real space, it is likeliest that we can add a particle near the ends of the B-site block, for one can imagine that, in the first configuration, this particle was just past the end in an adjacent block, and we merely hopped it a short distance across the boundary to create the state of $(P+1)$ particles on the block in question. It follows that the single-particle eigenfunctions with single-particle pseudoenergies near the pseudo-Fermi level have their greatest amplitude near the block's boundaries. In other words, it is the sites near the end that are most correlated with the environmental information that we discarded.

B. General features

As noted earlier, the eigenstates of ρ_B are all built up from the eigenstates $|\chi_l\rangle$ of G, which are simultaneously the one-particle eigenstates of ρ_B. As such, the effects of basis truncation, particularly in obtaining a truncated expansion of the target state $|\Psi\rangle$, must be understood in terms of the features of these one-particle eigenstates. The real-space features of $|\chi_l\rangle$ can most easily be understood in terms of the corresponding eigenfunctions $\chi_l(j)$, where $j=1,\ldots,B$ are sites on the block. At half filling, the probability densities $|\chi_l(j)|^2$ exhibit particle-hole symmetry, as is shown in Fig. 9 for the example of $B=9$. In general, by node counting, we see that the sequence of B single-particle eigenfunctions is in one-to-one correspondence with the sequence of B plane-wave states on the block, where the ordinal number l of the single-particle eigenfunctions is related to the wave vector k of the plane-wave states on the block by

$$k=\frac{2\pi(l-1)}{Ba}, \quad l=1,\ldots,B. \tag{5.1}$$

C. Scaling behavior

For odd B, the pseudoenergy $\varphi_{(B+1)/2}$ sits at the pseudo-Fermi level, and we may call the corresponding eigenfunction $\chi_F(j)\equiv\chi_{(B+1)/2}(j)$ the *pseudo-Fermi eigenfunction*. The probability density associated with $\chi_F(j)$ has nodes at every even j, as shown in Fig. 10 for the case of $B=23$. The most prominent feature of the pseudo-Fermi eigenfunction, i.e., the amplitude being strongest near the boundaries of the block, was first observed by White.[2] This appears to be a generic feature that occurs in both integrable and nonintegrable one-dimensional systems. Using the example of a chain of coupled harmonic oscillators, Gaite explained this "concentration of resolution of quantum states near the boundaries" as a simple consequence of angular quantization of the density-matrix.[17]

To analyze $|\chi_F(j,B)|^2$ [where we write the B dependence of $\chi_F(j)$ more explicitly] more carefully, we first rescale the eigenvectors obtained from OCTAVE[18] such that

$$|\chi_F((B+1)/2,B)|^2=1 \tag{5.2}$$

for $B=4p+1$, $p=1,2,\ldots$. For $B=4p+3$, $|\chi_F((B+1)/2,B)|^2=0$ and the rescaling cannot be carried out as unambiguously as for the $B=4p+1$ series. This rescaling is harmless, since eigenvectors are only defined up to an arbitrary normalization. After this trivial rescaling, we find that the pseudo-Fermi probability density can be put in a scaling form

$$|\chi_F(j,B)|^2\cong N(B)g(y)\frac{\frac{1}{2}[1-(-1)^j]}{\sin^2\pi y}, \tag{5.3}$$

where $y\equiv(j-\frac{1}{2})/B$ and $g(y)$ is the scaling function shown in Fig. 11.

In Eq. (5.3), $N(B)$ is a B-dependent normalization factor, chosen to ensure that the pseudo-Fermi wave function given is properly normalized in the limit of $B\to\infty$, i.e.,
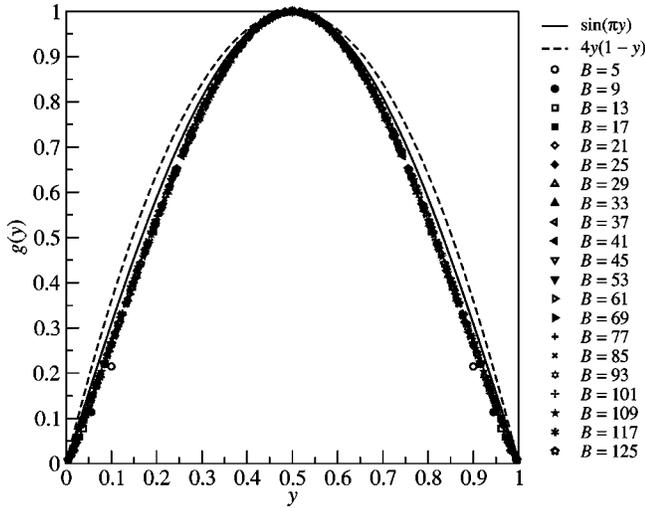
FIG. 11. Plot of the rescaled envelope function $g(y)$ for various block sizes $B = 4p + 1$, $p = 1,2, \ldots$, compared against $\sin \pi y$ and $4y(1-y)$, where $y = (j - \frac{1}{2})/B$ is the rescaled coordinate on the block.

$$\lim_{B \to \infty} \sum_{j=1}^{B} |\chi_F(j,B)|^2 = 1. \quad (5.4)$$

Although we cannot compute $N(B)$ analytically, we venture a guess to its B dependence by noting that the functions $g(y)$ and $\sin \pi y$ are not very different from the function $4y(1-y)$, and so we expect

$$\sum_{j=1}^{B} g(y) \frac{\frac{1}{2}[1-(-1)^j]}{\sin^2 \pi y} \sim \sum_{j=1}^{B} 4y(1-y) \frac{\frac{1}{2}[1-(-1)^j]}{[4y(1-y)]^2}$$

$$= \frac{1}{4} \sum_{j \text{ odd}}^{B} \left[\frac{1}{y} + \frac{1}{1-y} \right], \quad (5.5)$$

which we can easily work out to have the form

$$\sum_{j=1}^{B} g(y) \frac{\frac{1}{2}[1-(-1)^j]}{\sin^2 \pi y} \sim B(\log_e B + C), \quad (5.6)$$

where C is a constant. Numerically, the best fit for $N(B)$ in the range of block sizes from $B = 33$ to $B = 125$ is obtained with

$$N^{-1}(B) = 0.249B \log_e B + 0.668B. \quad (5.7)$$

Because of the enhanced amplitude near the edge of the block exhibited in the real-space structure of density-matrix eigenfunctions with single-particle pseudoenergies close to the pseudo-Fermi level, and conversely, enhanced amplitude near the center of the block exhibited in the real-space structure of density-matrix eigenfunctions with single-particle pseudoenergies far away from the pseudo-Fermi level, we worry that these eigenfunctions might not be a good basis to use for expanding spatially uniform plane waves, which are the true single-particle energy eigenstates in our model. We address this concern in Sec. VI C.
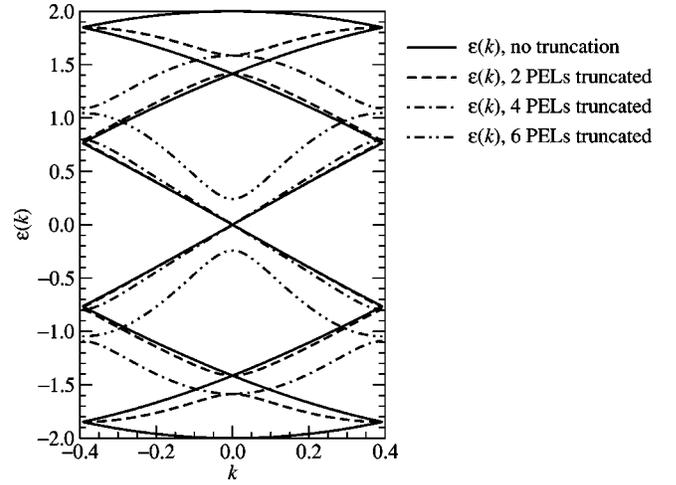

FIG. 12. Dispersion relation $\epsilon(k)$ for a block of $B = 8$ sites, where the effects of truncating 2, 4, and 6 PEL's are shown. For this block size, truncating 2, 4 and 6 PEL's corresponds to fractions $\gamma = 0.75$, 0.50, and 0.25 of PEL's retained. For $\gamma = 0.75$, the energy bands (dashed curves) just below and above the Fermi level $\epsilon_F = 0$ agree with the true dispersion relation (solid curve) so well that the difference is not discernible at the scales presented in the figure.

VI. OPERATOR-BASED DENSITY-MATRIX TRUNCATION APPLIED TO DISPERSION RELATION

In a gapless system, we conjecture that low-lying excitations *above* the ground state are built from the same operators as the long-wavelength fluctuations *within* the ground state. This supposition is certainly validated if the system has a continuous symmetry and the long-wavelength modes are Goldstone modes. In general it is justified by the relationships between correlation functions (for the ground-state fluctuations) and response functions (for low-energy excitations).

Despite its poor convergence properties as far as exhausting the sum rule (4.1) is concerned, the operator-based truncation would still get the salient features of the physics right. We check this by projecting the Hamiltonian in Eq. (1.1) onto the truncated set of fermion operators f_l, and calculate the dispersion relation therefrom. There are two physical quantities of interest here: (a) for odd number of sites B, the middle band crosses the Fermi level, and we can ask how the Fermi velocity, given by the slope of the dispersion relation at the Fermi level, scales with B and the fraction γ of fermion operators kept; and (b) for even B, a band gap develops as a result of truncation at the Fermi level, and we can ask how the size of this band gap depends on B and γ.

A. Energy gap at Fermi-level

In Fig. 12, we show the general features of the dispersion relation $\epsilon(k)$ calculated within the operator-based truncation scheme, using the example of a block of $B = 8$ sites. Apart from the energy gap ΔE that opens up at the Fermi level $\epsilon_F = 0$, we see that there is a one-to-one correspondence between the PEL truncated and the energy band absent from the dispersion relation. More precisely, if we order the en-
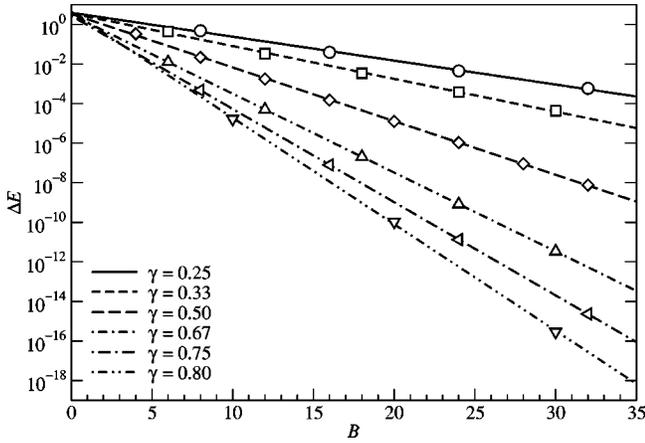
FIG. 13. Plot of the gap ΔE as a function of block size B for various constant fractions $\gamma = 1/4$, $1/3$, $1/2$, $2/3$, $3/4$, $4/5$ of PEL's retained. Also shown are fits to the data points for various γ to $\Delta E(\gamma, B) = \Delta E_0 \exp[-\kappa(\gamma)B]$.

ergy bands and the PEL's from the lowest to the highest as $\{\epsilon_1(k), \ldots, \epsilon_B(k)\}$ and $\{\varphi_1, \ldots, \varphi_B\}$, then if we truncate PEL φ_l, the energy band $\epsilon_l(k)$ will also be removed from the numerically calculated dispersion relation. For fixed γ, the gap ΔE decays exponentially with block size B, as is shown in Fig. 13, i.e., we have

$$\Delta E = \Delta E_0 \exp[-\kappa(\gamma)B], \tag{6.1}$$

where $\kappa(\gamma)$ is an attenuation coefficient whose γ dependence is shown in Fig. 14. Here we see also that $\Delta E(B, \gamma)$ for different γ appears to converge onto $\Delta E_0 \equiv \Delta E(B = 0)$. Of course, there is no physical sense in talking about a block of zero size, but it is nevertheless a useful number to keep in mind when studying the scaling behavior of $\Delta E(B, \gamma)$ as γ varies. ΔE_0 is ≈ 4, which is the bandwidth of the exact dispersion relation, in all cases.
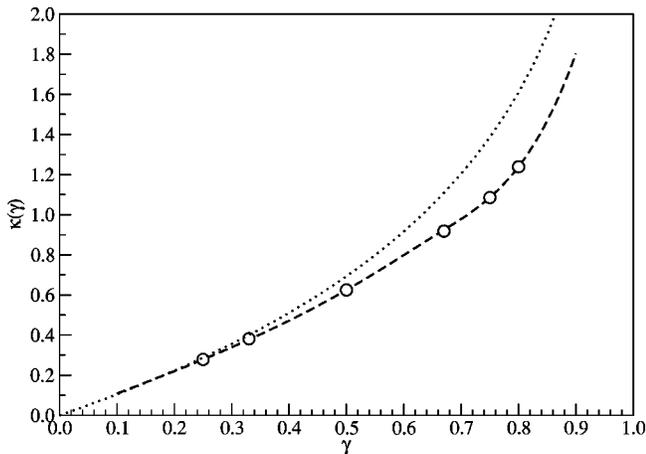

FIG. 14. Plot of the attenuation coefficient $\kappa(\gamma)$ with which the spurious gap ΔE decays, as a function of the fraction γ of PEL's retained. A cubic spline curve (dashed curve) is superimposed on the data points to aid visualization. Also shown (dotted curve) is $-\log_e(1-\gamma)$.

In particular, in the limit of $\gamma \rightarrow 1$, where all PEL's are retained, the gap is exactly zero for all nonzero block sizes B. In this limit, if we start out at a "gap" of ΔE_0 at a "block size" of $B = 0$, then to have $\Delta E = 0$ at $B = 1$, we need the attenuation coefficient κ to be infinite, i.e., we expect the limiting behavior $\lim_{\gamma \rightarrow 1} \kappa(\gamma) = \infty$. On the other hand, in the limit of $\gamma \rightarrow 0$, where we retain none of the PEL's, it is again physically meaningless to talk of a dispersion relation. Nevertheless, if we pretend that we are able to calculate a "dispersion relation" in this limit, then it is reasonable, following the trend observed in Fig. 13, that the gap never closes, i.e., we expect the limiting behavior $\lim_{\gamma \rightarrow 0} \kappa(\gamma) = 0$. These limiting behaviors appear to be borne out in the trend observed in Fig. 14.

Another notable feature in Fig. 14 is the fact that $\kappa(\gamma) \approx \gamma$ for $\gamma \ll 1$, which is the regime we are most interested to apply the operator-based truncation scheme in. To appreciate the relevance of this observation, let us first note from Fig. 12 the general feature that the smaller the gap ΔE, the better the truncated dispersion relation matches the true dispersion relation about the Fermi level. From Eq. (4.32) we saw that the truncated weight W_t depends only on the combination $l_{max} = \gamma B$ in this regime, and as far as W_t is concerned, there is no difference whether we choose to keep 10 out of 100 PEL's ($\gamma = 0.1$) or 10 out of 200 PEL's ($\gamma = 0.05$). Here we see a similar exponential dependence on l_{max} for the spurious gap ΔE that arises due to truncation: if we write $\kappa \approx \gamma$ in this regime, then $\Delta E \approx \Delta E_0 \exp(-B\gamma) = \Delta E_0 \exp(-l_{max})$. Such an exponential behavior implies that we have very good control over the numerical accuracy of the dispersion relations—in particular near the Fermi level—calculated in the operator-based DM truncation scheme.

B. Fermi velocity

When the block size B is odd, the central energy band crosses the Fermi level, and the quantity of interest becomes the Fermi velocity v_F. This can be determined from the truncated dispersion relation by taking the numerical central derivative of the central energy band at $k = \pm \bar{n}\pi/B$. At half filling, $\epsilon(k = \pm \pi/2B) = 0$ exactly because of particle-hole symmetry. This feature of the dispersion relation was found to be preserved in the numerical dispersion relations computed within the operator-based DM truncation scheme. On the global scale, we find numerically that the shifts in the central energy band at the Brillouin zone center and Brillouin zone edge are such that $v_F > 2$ always. However, when B is large, the numerical diagonalization routine introduces artifacts on the energy scale of 10^{-13}, resulting in the locally evaluated v_F coming out to be very slightly less than 2. As such, we analyze the behavior of v_F as a function of B and γ only for $B < 150$, as shown in Fig. 15.

As can be seen from Fig. 15, the difference $(v_F - 2)$ decays more or less exponentially with B for various γ, i.e.,

$$(v_F - 2) \cong \exp[-\xi(\gamma)B], \tag{6.2}$$

where $\xi(\gamma)$ is the γ-dependent attenuation coefficient for the average exponential decay. The γ dependence of $\xi(\gamma)$ is shown in Fig. 16.
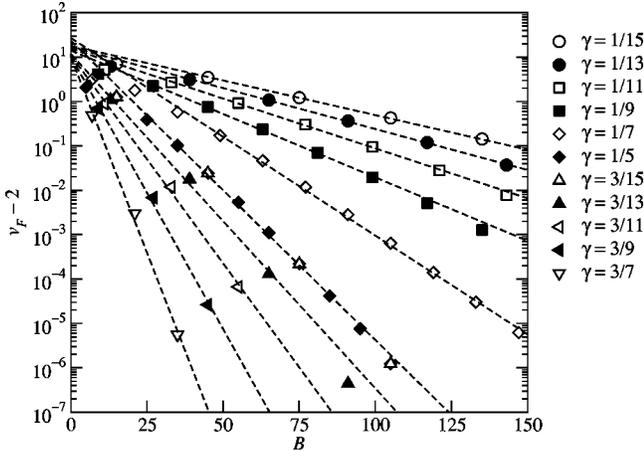
FIG. 15. Plot of Fermi velocity deviation (v_F-2) calculated from the truncated dispersion relation, as a function of the block size B, for various fractions $\gamma=1/15,1/13,1/11,1/9,1/7,1/5,3/15,3/13,3/11,3/9,3/7$ of PEL's retained. Fits to average exponential decays are also shown.

C. Real-space structure of eigenfunctions of truncated Hamiltonian

The eigenfunctions of the untruncated Hamiltonian (1.1) are spatially uniform plane waves, with amplitude $\exp(ikj)/\sqrt{B}$ on site j of the block of B sites. These can be expanded in terms of the density-matrix eigenfunctions $\chi_l(j,B)$. Naively, we expect that if we drop those $\chi_l(j,B)$ associated with pseudoenergies $\varphi(l,B)$ far from the pseudo-Fermi level φ_F, as we would in our operator-based truncation scheme which removes these single-particle pseudoenergy levels as degrees of freedom, the remaining terms, all having enhanced amplitudes at the edge of the block, would sum to a function with enhanced amplitude at the edge of the block. It would therefore seem like we are attempting to approximate a spatially uniform plane wave with a function with the wrong real-space structure.

However, the key insight we gain from our study of block density matrices is that while the system-wide density matrix
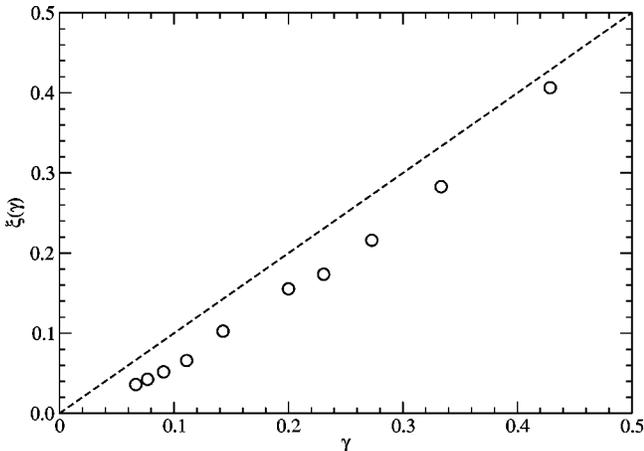

FIG. 16. Plot of the attenuation coefficient $\xi(\gamma)$ as a function of the fraction γ of PEL's retained. Also shown (dashed line) is the expected behavior $\xi(\gamma)=\gamma$.
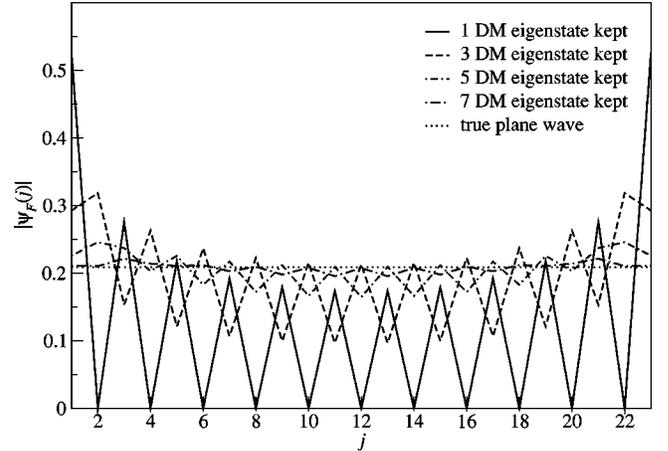

FIG. 17. Amplitudes of eigenfunctions of the truncated Hamiltonian $\tilde{H}(k_F)$ for different numbers of PEL's retained, for a block of $B=23$ sites at half filling.

ρ_0 commutes with H in Eq. (1.1), the block density matrix ρ_B obtained by tracing down ρ_0 does not commute with $H(k)$, for all k. Therefore, after operator-based truncation $H(k)\to\tilde{H}(k)$, we would need to diagonalize $\tilde{H}(k)$ to find the truncated dispersion relation $\epsilon_l(k)$. Thus, the function that would approximate the plane wave is not the latter's truncated expansion in terms of the eigenfunction of the one-particle block density matrix, but rather, a particular eigenfunction of $\tilde{H}(k)$, which is an appropriate linear combination of the $\chi_l(j,B)$ retained in the operator-based truncation scheme. We show in Fig. 17 the spatial structure of such a function, for various numbers of density-matrix eigenfunctions retained. As we can see, for a block of $B=23$ sites, keeping seven density-matrix eigenfunctions with pseudoenergies around φ_F would produce a decent approximation to the plane wave with $k=\pi/2B$.

D. Discussion

We speculate that the fact that the operator-based density-matrix truncation scheme succeeds so well suggests that appropriate linear combinations of the density-matrix eigenfunctions can closely approximate a plane wave with wave vector $\pm k_F$. This is only possible by taking the difference of two eigenfunctions so as to cancel the enhancements of the envelope function seen at the block ends (see Sec. V). Indeed, the fact that we get the correct slope v_F of the dispersion near k_F suggests that by taking different weights, a continuously varying effective wave vector can be approximated.

The fact that the goodness of approximation depends only on the number of eigenfunctions kept means that we approximate the wave function about as well in two successive blocks of B sites, as we do in one big block of $2B$ sites. One could speculate that there might exist some sort of approximate composition formula, analogous to Clebsch-Gordan formulas for combining angular momenta that provides the $2B$-site eigenfunctions in terms of the direct product of the B-site eigenfunctions.
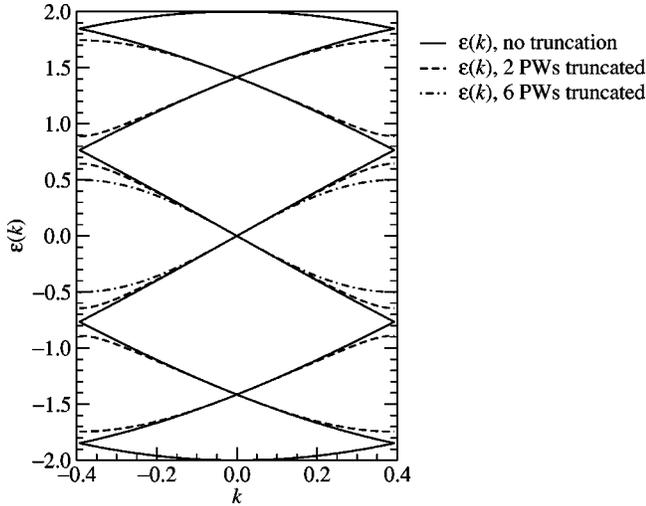
FIG. 18. Dispersion relation $\epsilon(k)$ for a block of $B=8$ sites, where the effects of truncating 2 and 6 plane waves (PWs) are shown. For $B=4n$, the Fermi level is located at the zone center, and $\epsilon(k)$ is always gapless here regardless of the number of plane waves truncated.

VII. OPERATOR-BASED PLANE-WAVE TRUNCATION SCHEME

As we saw in Sec. V, eigenstates of the density matrix ρ_B are approximately plane waves (with wave vector Q determined by the boundary conditions on the block of B sites) modified by some envelope function. Apart from the effects of the envelope functions, the operator-based truncation scheme described above is likened to truncating wave vectors Q far away from the Fermi wave vector k_F. It is therefore natural to investigate how an operator-based truncation scheme based on plane waves would fare against that based on the density-matrix eigenstates.

A. Exact dispersion at zone center

Compared to the operator-based DM truncation scheme developed above, the most striking feature of the operator-based PW truncation scheme is that it gets the dispersion exactly right at the zone center, as shown for the case of $B=8$ in Fig. 18, and for the case of $B=10$ in Fig. 19. We understand this as follows.

To evaluate the dispersion relation in a blocked description, we start by defining the direct Bloch basis states

$$|j,k\rangle = \frac{1}{\sqrt{N/B}} \sum_J e^{ikJB}|j,J\rangle, \quad j=1,\ldots,B, \quad (7.1)$$

where $|j,J\rangle = c^{\dagger}_{j+JB}|0\rangle$ is the single-particle occupation number basis state at site $j+JB$ along the chain. In this basis, the Hamiltonian (1.1) for a chain of N noninteracting spinless fermions take on a block diagonal form. Diagonalizing the $B \times B$ diagonal block
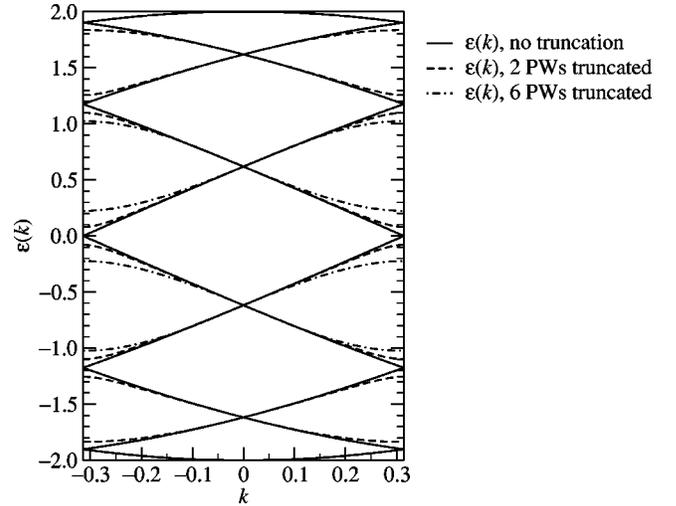

FIG. 19. Dispersion relation $\epsilon(k)$ for a block of $B=10$ sites, where the effects of truncating 2 and 6 plane waves (PWs) are shown. For $B=4n+2$, the Fermi level is located at the zone boundary, and the operator-based plane wave truncation scheme introduces an energy gap ΔE here.

$$H(k) = \begin{bmatrix} 0 & -1 & 0 & \cdots & -e^{-ikJB} \\ -1 & 0 & -1 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ -e^{ikJB} & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$(7.2)$$

for $-\pi/B \leq k < \pi/B$ then gives the dispersion relation within the reduced zone scheme.

For the operator-based PW truncation scheme, we need to work with the plane-wave states $|Q,J\rangle$ on each block of B sites, where the wave vector Q is determined by periodic boundary condition, i.e., $\exp(iQB)=1$. These plane-wave states are related to the single-particle occupation number basis states by

$$|Q,J\rangle = \frac{1}{\sqrt{B}} \sum_{j=1}^{B} e^{iQj}|j,J\rangle. \quad (7.3)$$

A Bloch basis state parallel to Eq. (7.1) can be defined as

$$|Q,k\rangle = \frac{1}{\sqrt{N/B}} \sum_J e^{ikJB}|Q,J\rangle, \quad (7.4)$$

where $QB/2\pi = 0, \ldots, B-1$. From Eqs. (7.3) and (7.4), it is easy to see that

$$|Q,k\rangle = \frac{1}{\sqrt{B}} \sum_{j=1}^{B} e^{iQj}|j,k\rangle. \quad (7.5)$$

At the zone center, $k=0$, and the $B \times B$ Hamiltonian matrix in the $|j,k\rangle$ basis that we need to diagonalize becomes

$$H(0) = \begin{bmatrix} 0 & -1 & 0 & \cdots & -1 \\ -1 & 0 & -1 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \qquad (7.6)$$

It is trivial to check that the eigenstates of this Hamiltonian matrix are precisely the plane waves $|Q,0\rangle$ on the block. Therefore, in the $|Q,k\rangle$ basis, $H(k)$ is diagonal at $k=0$, and so truncating some plane waves from the Hilbert space produces no effect on the dispersion here.

To be more precise, in performing truncation, a linear subspace of the Hilbert space is chosen, and the Hamiltonian projected onto this subspace. If $|\psi\rangle$ is an eigenstate of the full Hamiltonian, and if $|\psi\rangle$ is retained in the truncated Hilbert space, then it will continue to be an eigenstate of the truncated Hamiltonian, with the same eigenvalue.

B. Energy gap at zone boundary

For even block sizes with $B=4n$, the Fermi level is located at the zone center in the reduced zone scheme, and so there is no energy gap to speak of. On the other hand, for even block sizes with $B=4n+2$, the Fermi level is located at the zone boundary. At the zone boundary, operator-based PW truncation introduces an energy gap ΔE at the Fermi level, as shown in Fig. 19 for $B=10$.

As in the case for the operator-based DM truncation scheme, we investigate the behavior of the energy gap ΔE as a function of the block size B for a fixed fraction $(1-\gamma)$ of block states truncated. However, for the operator-based PW truncation scheme, the number of plane-wave states that can be truncated, if B is even, is $4m+2$, $m=0, 1, 2, \ldots$. Thus the only realizable series of block sizes B on which we can perform fixed $(1-\gamma)$ truncation is of the form $B=q(4m+2)$, $q=2, 3, \ldots$. The fraction γ of block plane-wave states retained is related to the series index q by

$$\gamma = 1 - \frac{1}{q}. \qquad (7.7)$$

Half of these realizable series have block sizes that are multiples of four, for which the Fermi level is at the Brillouin zone center where the dispersion relation we have shown in the previous section to be gapless. In this section we are interested in those block sizes for which q is an odd integer, since for these block sizes the Fermi level is at the Brillouin zone boundary, where a gap develops in the dispersion relation as a result of truncation. The behavior of ΔE as a function of B for three series of q is shown in Fig. 20.

As can be seen from Fig. 20, the gap depends on block size as an inverse power law

$$\Delta E(B,\gamma) = \frac{\Delta E_1(\gamma)}{B}, \qquad (7.8)$$

where $\Delta E_1(\gamma)$ is a γ-dependent prefactor. This is in stark contrast to the exponential dependence (6.1) found for the case of the operator-based density-matrix truncation scheme.
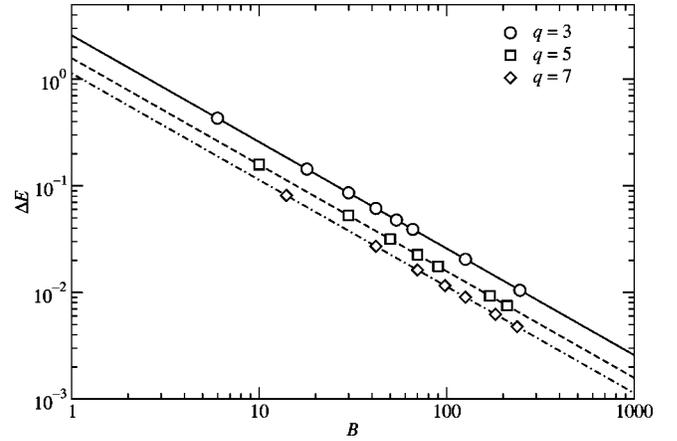

FIG. 20. Plot of ΔE as a function of block size B for $q=3$, 5, and 9, corresponding to the fractions $\gamma=2/3,4/5,6/7$ of block plane-wave states retained. Also shown are the fits of the data points to the formula $\Delta E(B,\gamma) = \Delta E_1(\gamma)/B$. From the fits, we have $\Delta E_1 = 2.582\,63$ for $\gamma=2/3$, $\Delta E_1 = 1.579\,91$ for $\gamma=4/5$, and $\Delta E_1 = 1.135\,32$ for $\gamma=6/7$.

C. Fermi velocity

For odd B, we again investigate the behavior of v_F as a function of B for the operator-based plane-wave truncation scheme. The number of block plane waves that can be truncated is $4m+3$, $m=0, 1, 2, \ldots$ and the series of realizable block sizes is $B=q(4m+3)$, $q=3, 5, \ldots$. Unlike in the operator-based DM truncation scheme, there appears to be two different systematic behaviors for $v_F(B,\gamma)$, one for $q=4p-1$ and another for $q=4p+1$ ($p=1,2,\ldots$). We find that the Fermi velocity can be fitted very well to the formula

$$v_F = \begin{cases} \bar{v}_F(\gamma) + c_+(\gamma)/B, & q=4p-1 \\ \bar{v}_F(\gamma) - c_-(\gamma)/B, & q=4p+1. \end{cases} \qquad (7.9)$$

The plots of $\bar{v}_F(\gamma)$ and $c_\pm(\gamma)$ as a function of γ are shown in Figs. 21 and 22, respectively.
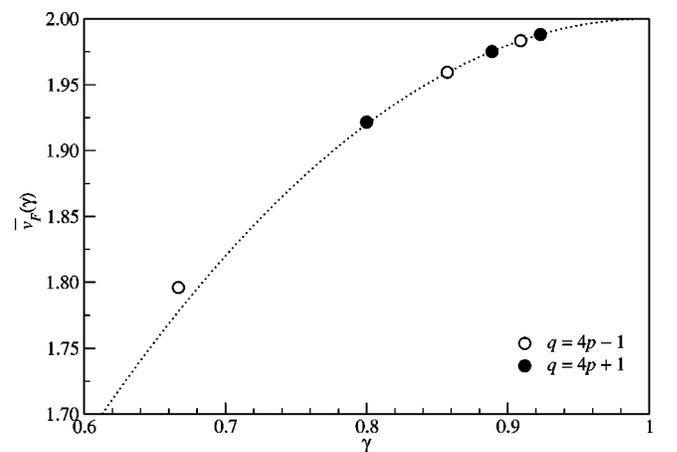

FIG. 21. Plot of $\bar{v}_F$ as a function of γ in the operator-based PW truncation scheme for both the $q=4p-1$ series and $q=4p+1$ series. Also shown as the dashed curve is $2[1-(1-\gamma)^2]$, which appears to fit the data points well near $\gamma=1$.
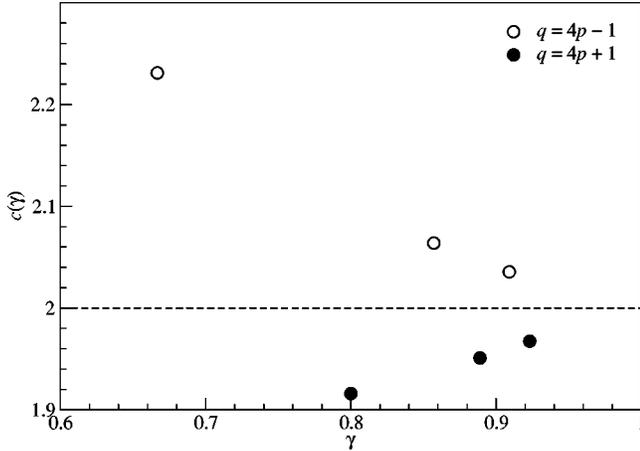
FIG. 22. Plot of $c_{\pm}$ as a function of γ in the operator-based PW truncation scheme for both the $q=4p-1$ series $[c_+(\gamma)]$ and $q=4p+1$ $[c_-(\gamma)]$ series. Both $c_+(\gamma)$ and $c_-(\gamma)$ appear to be converging towards 2. In fact, from the graph we find that $|c_{\pm}(\gamma)-2|\approx 0.43(1-\gamma)$.

As can be seen from Fig. 21, the exact value of the Fermi velocity is obtained only in the double limit of $B\rightarrow\infty$ and $\gamma\rightarrow 1$. Compared to the operator-based DM truncation scheme, where we manage to achieve the exact Fermi velocity for any γ, this is clearly undesirable. Furthermore, even very close to $\gamma=1$, the computed Fermi velocity approaches the limiting value $\bar{v}_F(\gamma)$ as B^{-1}. This is much slower than the exponential convergence of $\exp[-\xi(\gamma)B]$ found for the operator-based DM truncation scheme.

VIII. SUMMARY AND DISCUSSIONS

A. Statistical-mechanical analogy and scaling

To summarize, we have in this paper developed an in-depth understanding of the structure of the eigenvalues and eigenstates of the density matrix ρ_B of a block of B sites embedded within an infinite one-dimensional chain of non-interacting spinless fermions described by the set of fermion operators $\{\tilde{c}_k\}$ with dispersion relation $\epsilon(k)$.

From Ref. 7 we know that the block density matrix ρ_B can be written in Eq. (1.2) as the exponential of a pseudo-Hamiltonian $\tilde{H}$ given by Eq. (1.3) acting only within the block, which describes another system of noninteracting spinless fermions with fermion operators $\{f_l\}_{l=1}^B$ and dispersion relation φ_l. We use the prefix *pseudo-* when talking about operators $\{f_l\}_{l=1}^B$ and energies φ_l to distinguish them from the operators $\{\tilde{c}_k\}$ and energies $\epsilon(k)$ of the system we started out with.

The many-particle eigenstates of ρ_B can then be interpreted simply as the many-body energy eigenstates (2.1) of the system of noninteracting spinless fermions described by $\tilde{H}$, and their associated density-matrix weights (2.2) are then their statistical weights in the grand canonical ensemble, *as though* the block is at a finite temperature. Within this statistical-mechanical picture, we can apply intuitions learnt from the statistical mechanics of real fermionic systems, and

talk about the filling of single-particle PEL's as dictated by the Fermi-Dirac distribution, the pseudo-Fermi sea with its pseudo-Fermi level φ_F, and particle-hole excitations of the pseudo-Fermi sea.

Using results obtained in Ref. 7, we make this statistical-mechanical picture more precise, by identifying the single-particle eigenstates $|\chi_l\rangle$ in Eq. (2.6) of ρ_B as those built up from the eigenvectors $\{\chi_l(j)\}$, $j=1,\ldots,B$, of the Green-function matrix G restricted to the block. The eigenvalue λ_l of G, related to the single-particle pseudoenergies φ_l by 2.7 and 2.8, are the average pseudo-occupation numbers of the lth PEL.

The statistical mechanics of real fermionic systems tells us that, at finite temperature, the physically important many-body states are those low-energy particle-hole excitations involving single-particle energy levels within k_BT of the Fermi level. Single-particle energy levels far away from the Fermi level contribute negligibly to the thermodynamic properties of the fermionic system, and are precisely the degrees of freedom to be truncated in a renormalization-group analysis.

We capitalize on this insight, and described a recipe for operator-based truncation of the density-matrix eigenstates, where out of the B pseudofermion operators $\{f_l\}_{l=1}^B$, we truncate those f_l's associated with single-particle pseudoenergies φ_l far away from φ_F and retain $l_{\max}$ of them with $\varphi_l\approx\varphi_F$. Within this operator-based truncation scheme, the effective pseudo-Hamiltonian acting on the truncated Hilbert space can be made to have the same form as the original pseudo-Hamiltonian, in the same spirit of renormalization-group transformations in statistical mechanics or quantum field theory.

Having laid out the basic principles behind our operator-based DM truncation scheme, we proceeded to look more closely into the distribution of single-particle pseudoenergies φ_l, and how these scale with the block size B. There are two related questions that provide the motivation for doing this.

(1) The statistical mechanics of real fermionic systems suggests that single-particle energy levels within $\beta^{-1}=k_BT$ of the Fermi level are the physically relevant degrees of freedom—what then is the effective temperature $\tilde{\beta}^{-1}$ that we should use as the natural cutoff when performing operator-based truncation on $\{f_l\}_{l=1}^B$?

(2) Although we have associated the pseudodispersion relation φ_l with the dispersion relation $\epsilon(k)$ of a real fermionic system of noninteracting spinless fermions, the wave vector k enumerating $\epsilon(k)$ is an *intensive* quantity whereas the ordinal number l enumerating φ_l is an *extensive* quantity—what would the intensive analog of l that more closely parallels the wave vector k, and how would the pseudo-dispersion relation look like in terms of this intensive label?

The natural answer to the second question would be to write the pseudodispersion relation φ_l as a function of l/B, totally analogous with how the wave vector k is enumerated as $2\pi m/N$ for a chain of N sites satisfying the Born-von Karman boundary condition. In fact, we find strong numerical evidence that suggests that the single-particle pseudoenergies, for various block sizes and fillings, satisfy a scaling relation of the form given in Eq. (3.6), where the scaling

function $f(\bar{n},x)$ is the proper analog of the dispersion relation $\epsilon(k)$ and the scaling variable x given in Eq. (3.7) is the proper analog of the wave vector k. From Eq. (3.12), we see that the block size B plays the role of inverse temperature.

Our scaling results in Sec. III indicate that the density-matrix eigenstates and eigenvalues behave, as block size B is increased, very much as energy eigenstates and eigenvalues do when the system size is increased. In the latter case, we have a dispersion relation and are more or less sampling it at different wave vectors. It is not quite that simple in the density-matrix case, since the scaling function (3.6)—our analog of the energy dispersion relation—depends on the filling $\bar{n}$. We only note that this analogy still lacks an analytical foundation. A more penetrating analysis is called for of the relation of G to ρ_B (or equivalently, the effect on its eigenvalues of restricting G to sites on a block).

Incidentally, we noted that our equation relating G to ρ_B was valid at any temperature, but we assumed zero temperature throughout this paper. We expect nonzero temperature T would become a second scaling variable. Since $T>0$ has similar effects on the Green function $G(r)$ as does the gap introduced in Eq. (4.5), we expect the scaling also behaves similarly and we did not investigate it.

B. Scaling as a guide to truncation in practice

In a real application, it appears quite unlikely that Hamiltonians will be projected directly onto large blocks (meaning blocks of more than 16 sites). What then is the practical value of extracting scaling forms, if they are unambiguously seen only in blocks of 100 or more sites? One answer is that, even though it is an oversimplified cartoon for the non-asymptotic situations in which it usually gets applied, a scaling law is easier to grasp than brute numerical or graphical facts.

The scaling relation (3.6) is a powerful tool that we can use to derive deeper understanding concerning the structure of the block density matrix, as well as various aspects of truncation. But in itself, the scaling relation provides only a partial answer to our first question, which is about how much of the Hilbert space of many-body states on the block of B sites we can truncate. To answer this question more completely, we looked at the three density-matrix eigenstates $|F\rangle$, $|F-1\rangle$, and $|F+1\rangle$ with the largest weights. Using the scaling relation (3.6), we find that the ratios $w_{F\pm1}/w_F$ of their weights (w_F being the largest density-matrix weight) approach a constant, in Eq. (4.4) for a gapless system of noninteracting spinless fermions, as $B\rightarrow\infty$.

The same result was also found for a gapped system of noninteracting spinless fermions, for which we find scaling relations governed by gap-dependent scaling functions. Furthermore, the scaling relation (3.6) allowed us to conclude that as $B\rightarrow\infty$, the largest density-matrix weight w_F approaches a constant given in Eq. (4.19), and derive approximately the dependence on the number l_{max} of PEL's retained and the truncated weight W_t in the operator-based DM truncation scheme.

C. DMRG and operator-based truncation

It is difficult to compare our results with those obtained in the context of the DMRG, because that is an incremental method. Rather than obtaining the density-matrix of a large block all at once, each iteration of the DMRG takes an approximate density-matrix for a block of B sites and produces an approximate density-matrix for a block of $(B+1)$ sites. The fraction of weight kept, which is taken as the figure of merit, refers to the small truncation in each iteration. The cumulative DMRG truncation might be more appropriate to compare with our results for rather large blocks.

Nevertheless, let us note that operator-based truncation can be applied independent of whether we use an incremental or one-shot method. In particular, operator-based truncation could be used in a test run to apply DMRG to a noninteracting Fermi chain. One is given a truncated list of t operators $\{f_l\}$ where $l=(B-t+1)/2,\ldots,(B+a-1)/2$ for the original block and a Hamiltonian projected onto it. One augments this list with the bare creation operators $c^\dagger_{B+1}$ on a new site that will be added, and defines the new Hamiltonian by adding the hopping to the new site.

In light of the derivation of Ref. 7, we anticipate that the density matrix of the augmented system's ground state must have the same quadratic form, with new operators $\{f'_l\}$, which could presumably be obtained merely by diagonalizing the single-particle sector. One simply deletes the least important member of this list to obtain a new truncated list, no longer than the original one.

This difficulty notwithstanding, we still carried out a naive comparison of the performance of the operator-based DM truncation scheme against the traditional weight-ranked DM truncation scheme used in the DMRG, using the ability to exhaust the sum rule (4.1) for a given total number L_{max} of density-matrix eigenstates retained as a criterion. The conclusion: while the operator-based DM truncation scheme does not exhaust the sum rule (4.1) as rapidly as the weight-ranked DM truncation scheme, W_t is still of $O(1)$, i.e., the significant parts of the total density-matrix weight are "captured" by the operator-based DM truncation scheme.

1. Truncation and dispersion relations

However, we believe that it is more important to check how well a truncation scheme do by calculating physical quantities, rather than rely solely on the truncated weight W_t as a performance indicator. To this end, we calculated the dispersion relation of elementary excitations within the operator-based DM truncation scheme (an easy thing to do), and found that the error incurred decays exponentially as l_{max}, the number of PEL's retained when $l_{max}\ll B$. This error is much smaller than $O(\epsilon)$, which is expected from a naive analysis based on the discarded weight $\epsilon=1-W_t$. Here, there is a subtle worry that it may be that truncation works especially well for our chosen hopping Hamiltonian (1.1) is so local. A Hamiltonian with longer range hopping would have the same Fermi sea and hence the same density-matrix, but the truncation errors might be worse.

It would be desirable to also calculate the dispersion relation within the weight-ranked DM truncation scheme, and compare the results to those obtained within the operator-based DM truncation scheme. However, in the latter case it is

problematic even to define the question, since each retained density-matrix eigenstate would be a many-particle state. The new truncated Hamiltonian might be conveniently expressed in terms of the pseudocreation operators $\{f_l\}$, but many combinations of occupations would not exist.

The situation would be somewhat analogous to taking a simple, noninteracting hopping Hamiltonian for spinfull fermions, and imposing a Gutzwiller projection (no doubly occupied sites). In effect, the projection made a noninteracting model into an interacting one. Similarly a weight-ranked truncation must introduce spurious interactions. Hence, even if a system containing several blocks were to be exactly diagonalized (using the truncated basis) we could not immediately identify the elementary excitations. One would need, for example, to numerically compute a spectral function $S(q,\omega)$, where $(\hbar q, \hbar \omega)$ are momentum and energy, and then locate peaks as a function of $\hbar \omega$. On the other hand, a system which is truncated using the operator-based truncation scheme can still be represented by a set of creation and annihilation operators.

2. Inadequacy of plane-wave truncation scheme

Following this, we argued, based on the real-space structure of the one-particle density-matrix eigenfunctions shown in Sec. V, that an operator-based truncation scheme can also be defined naturally using the basis formed by single-particle PW states on the block of B sites. The dispersion relation was calculated within this operator-based PW truncation scheme, and compared to that calculated within the operator-based DM truncation scheme. We find that, other than getting the dispersion exactly right at the zone center of the reduced Brillouin zone, the operator-based PW truncation scheme is generally inferior to the operator-based DM truncation. Instead of decaying exponentially as l_{max}, the error in the dispersion relation calculated within the operator-based PW truncation scheme decays as a power law l_{max}^{-1}, which means that more single-particle basis states need to be retained in the operator-based PW truncation scheme as compared to the operator-based DM truncation scheme.

D. Towards interacting systems in dimensions $d \geq 2$

It is not much harder in principle to analyze the proposed operator-based density-matrix truncation scheme for *noninteracting* fermions in two dimensions. However, it will be harder to understand the scaling since we cannot simply rank the eigenvalues. One dimension was special because we know that each successive state has one more node than the previous one. A further very important difference is that in $d=1$ there are just two Fermi points, whereas in $d \geq 2$ there is a Fermi surface. Thus, whereas in $d=1$ the density-matrix eigenstates near the pseudo-Fermi level are related to the energy eigenstates at $\pm k_F$, in $d \geq 2$ these eigenstates will correspond to mixtures of wave vectors from every piece of the Fermi surface.

Obviously, noninteracting systems do not require numerical studies, so we must clarify how our results are relevant to the problem of interacting systems. First, many (gapless) systems of interest are in a phase—Fermi liquid, d-wave

superconductor—which are noninteracting in the low-energy, long-wavelength limit. When applied to a Fermi liquid system, we expect (to the extent that the truncation has separated out the low-energy modes) that any iterative renormalization scheme will converge to the noninteracting limit, and it must behave properly in that limit to have even the hope of success. Hence, for a density-matrix-based scheme, the first order of business is to study the density-matrix for a free fermion ground state (as in this paper) or for a BCS state,[8] i.e., the density-matrix will actually have the simple operator-based form, and hence its many-particle eigenstates of the density-matrix really are built from the single-particle eigenstates.

For an interacting system, the truncated basis should of course be constructed using the many-body density matrix for that system (not the noninteracting system). We expect that the operators generating this many-particle truncated basis will *not* just be those that create the one-particle density-matrix eigenstates (they were in the noninteracting case studied in this paper). More thought will be needed as to discover the best recipe to optimize the truncation rule so as to balance the needs of sectors with different particle numbers in a strongly interacting system. The simple algebraic structure of the noninteracting density-matrix eigenstates, considered in this paper, has motivated investigations (in progress) of the relationships among the many-particle density-matrix eigenstates for an interacting system.

A separate reason why our results for the noninteracting fermions are relevant to the study of interacting systems is that the scaling behavior of the noninteracting density matrix should be a good guide to that of interacting systems, although details may differ. This is in the same sense that mean-field theory is a good guide to the overall pattern of critical phenomena.

However, other interacting models of interest sit at quantum critical points that are not described by quasiparticle interactions, or possess fractionalized excitations. Since we do not presently understand the proper way to write their wave functions in terms of a spatially blocked basis, nor the proper renormalization step to capture the interblock correlations in the fractionalized systems, we do not know if a plain block density matrix gives the proper basis for truncation of the states. Furthermore, in the absence of an analytic construction of the block density matrix, for example, for Laughlin's quantum Hall wave function or the one-dimensional Su-Schrieffer state, we cannot proceed to scaling studies of large blocks like those found in the present paper, but numerical studies of such density matrices might be an illuminating subject for future research.

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

[2] S.R. White, Phys. Rev. B **48**, 10 345 (1993).

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

[4] S.R. White, Phys. Rep. **301**, 187 (1998).

[5] R.M. Noack, S.R. White, and D.J. Scalapino, in *Proceedings of the Seventh Workshop, Athens, GA, USA, 1994*, edited by D.P. Landau, K.-K. Mon, and H.-B. Schüttler (Athens, Georgia, USA, 1994), pp. 85–98.

[6] S. Liang and H. Pang, Phys. Rev. B **49**, 9214 (1994).

[7] S.-A. Cheong and C.L. Henley, preceding paper, Phys. Rev. B **69**, 075111 (2004).

[8] M.C. Chung and I. Peschel, Phys. Rev. B **64**, 064412 (2001).

[9] R.K. Pathria, *Statistical Mechanics*, 2nd ed. (Butterworth-Heinemann, Oxford, 1996).

[10] E.R. Davidson and L.L. Jones, J. Chem. Phys. **37**, 2966 (1962).

[11] C.F. Bender and E.R. Davidson, J. Phys. Chem. **70**, 2675 (1966).

[12] F. Jensen, *Introduction to Computational Chemistry* (Wiley, Chichester, 1999).

[13] P.W. Anderson, J. Phys. C **3**, 2436 (1970).

[14] T. Xiang, Phys. Rev. B **53**, R10445 (1996).

[15] J. Dukelsky and G. Sierra, Phys. Rev. Lett. **83**, 172 (1999).

[16] F. Gebhard, K. Bott, M. Scheidler, P. Thomas, and S.W. Koch, Philos. Mag. B **75**, 1 (1997).

[17] J. Gaite, Mod. Phys. Lett. A **16**, 1109 (2001).

[18] OCTAVE, URL http://www.octave.org/

[19] The same conclusion holds for the case of B odd, apart from the technical annoyance that there are *two* many-particle density-matrix eigenstates with the largest weight.