Density matrix renormalization group algorithms with a single center site

Steven R. White

Department of Physics and Astronomy, University of California, Irvine, CA 92697

(Dated: March 23, 2022)

We develop a correction to the density matrix used in density matrix renormalization group calculations to take into account the incompleteness of the environment block. The correction allows successful calculations using only a single site in the center of the system, rather than the standard two sites, improving typical computation times by a factor of two to four. In addition, in many cases where ordinary DMRG can get stuck in metastable configurations, the correction eliminates the sticking. We test the new method on the Heisenberg $S = 1$ chain.

Since the density matrix renormalization group (DMRG) was developed\cite{1,2}, it has gradually been applied to more and more difficult systems, such as wide ladders and 2D clusters, and systems with long-range interactions. One of the problems arising in these systems is the possibility that the simulation gets stuck far from the ground state\cite{3}. Several approaches have been developed to overcome this problem, such as controlling the starting wavefunction through potentials or quantum numbers, with the controls later removed. Nevertheless, there has remained much room for improvement. In 1D short-range systems, the standard DMRG finite system algorithm avoids convergence problems remarkably well because of the presence of the second center site in the block configuration. However, the extra site increases the computation time and memory requirements. An alternative to utilizing the extra site, which works better in the more difficult cases, has not been available. In this paper we describe such an alternative method, which relies on a correction to the reduced density matrix in order to retain a broader variety of states.

In the top panel of Fig. 1 we show the “superblock” configuration for the standard finite-system algorithm, where the lattice is divided into two large blocks, the system and the environment blocks, both with truncated bases, with two sites between them. The algorithm for a single DMRG step consists of finding the ground state for this “superblock”; obtaining the density matrix for the system block plus site; diagonalizing this density matrix; and then changing basis to the most probable eigenvec-

dition of states between various quantum numbers in the environment also translates directly to the renormalized system block. The distribution of states between various quantum numbers in the environment also translates directly to the renormalized system block. Note that if the environment block has $m$ states, then the maximum number of nonzero eigenvalues of the density matrix is also $m$, and the number of states never increases unless states are added “artificially” despite having density matrix eigenvalues of zero. Simple fixes, such as adding extra random states with a larger range of quantum numbers, improve but do not fix the very poor convergence of the single site algorithm.

The essential problem here arises when a particular fluctuation between the system and environment which should be present is not because the environment block

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure1.png}
\caption{Standard two-site DMRG method (top) and the single site method.}
\end{figure}
does not have the relevant states. Hence, the fluctuation is not represented in the density matrix and the new system block will not possess its relevant states for that fluctuation. Later, when the roles of system and environment are reversed, the relevant states again do not appear. In a 1D system with short range interactions the extra environment site does a very good job of ensuring that the most relevant fluctuations are at least approximately present the environment, so that subsequent sweeps can build in the fluctuations to high accuracy.

In wide ladders or systems with longer range interactions, the addition of a single site to the environment is not always adequate. There may be missing fluctuations which are far from the extra site, and so are never built in. Even in these cases the extra site allows $m$ to increase sensibly and as one lets $m \to \infty$ one obtains exact results. However, for practical values of $m$ one may find unacceptably slow convergence.

In this paper we describe an approximate correction to the density matrix to describe the key states which have been left out because the environment block is inadequate. With this correction, the single site superblock configuration converges well. In addition, convergence in more difficult systems is dramatically improved, in either the single site or two site configurations. We present two different derivations of the correction, and give examples using the $S = 1$ Heisenberg chain.

We first give a simple, rough argument. Consider the power method for finding the ground state: iterate $\psi_{n+1} = (1 - \varepsilon H)\psi_n$, where $\varepsilon$ is a small constant. As long as $\psi_0$ is not orthogonal to the exact ground state, and $\varepsilon$ is small enough, the power method is guaranteed to converge to the ground state. Consequently, if the basis represents both $\psi$ and $H\psi$ exactly, and we minimize the energy within this basis, we expect exact convergence. The crucial point is the need to enlarge the basis to represent $H\psi$. Within the standard DMRG basis obtained from $\psi$, after solving for the ground state, $H\psi = E\psi$, and nothing is changed by adding $H\psi$ to the basis. To go beyond the basis, we need to construct the parts of $H\psi$ as the basis is built up. The crucial terms of $H\psi$ come from the terms of $H$ which connect the system and environment blocks.

For the current superblock configuration, write the Hamiltonian in the form

$$H = \sum_\alpha t_\alpha \hat{A}^\alpha \hat{B}^\alpha.$$  

(1)

Here the $\hat{A}^\alpha$ act only on the system block (including the site to be added to it), and the $\hat{B}^\alpha$ act only on the environment block (plus its site). All the terms which do not connect the blocks are contained in two terms of the sum which have either $A$ or $B$ equal to the identity operator, so that this form is completely general. (The other term in each case is the block Hamiltonian.) In order to put $H\psi$ into the basis, we need to target, in addition to $\psi$, the terms $\hat{A}^\alpha \psi$ for all $\alpha$. Let the states of the system have indices $s$, $p_i$, and $q$, and the states of the environment $e$. The state $\hat{A}^\alpha \psi$ can be written as

$$\sum_{se} \sum_p A^\alpha_{sp} \psi_{pe}|s\rangle\langle e|.$$  

(2)

Targetting this wavefunction means adding into the density matrix a term

$$\Delta\rho^\alpha_{se} = a_\alpha \sum_{epq} A^\alpha_{sp} \psi_{pe} \psi^*_{qe} A^\alpha_{qe} \psi^*_{pe}$$  

(3)

where $a_\alpha$ is an arbitrary constant determining how much weight to put into this additional state. The total contribution of all the terms is

$$\Delta\rho = \sum_\alpha a_\alpha \hat{A}^\alpha \rho \hat{A}^{\alpha\dagger}$$  

(4)

where $\rho$ is the density matrix determined in the usual way, only from $\psi$. This is the form of the correction that we use, with $a_\alpha = a \sim 10^{-3} - 10^{-4}$.

As a second derivation, we utilize perturbation theory. First, imagine that the environment block, but not the system block, is complete. We obtain the ground state exactly for this superblock, and then transform to the basis of density matrix eigenstates for the system block, and then also do the same for the environment block. Then the wavefunction can be written in the form

$$|$\psi'\rangle = \sum_s \psi_s|L_s\rangle|R_s\rangle.$$  

(5)

The reduced density matrix is

$$\rho = \sum_s \langle R_s|\psi'\rangle\langle\psi|R_s\rangle = \sum_s |\psi_s|^2|L_s\rangle\langle L_s|$$  

(6)

Now consider the realistic case where the environment block is not complete. Assume the incompleteness takes the simple form that some of the $|R_s\rangle$ are missing, labeled $\bar{s}$, whereas $s$ are present. Let $P$ be a projection operator for the environment block $P = \sum_s |s\rangle\langle s|$, and take $\bar{P} = 1 - P$. Let the unperturbed ground state, with energy $E_0$ and density matrix $\rho_0$, be obtained using the incomplete environment basis. We take as a perturbation the terms in the Hamiltonian which couple to the states $\bar{s}$, namely

$$H' = \sum_\alpha t_\alpha \hat{A}^\alpha (\bar{P} \bar{B}^\alpha \bar{P} + \bar{P} \bar{B}^{\alpha\dagger} \bar{P}).$$  

(7)

The first order perturbative correction to the wavefunction due to $H'$ is

$$|\psi'\rangle = \sum_\alpha t_\alpha (E_0 - H_0)^{-1} \hat{A}^\alpha \bar{P} \bar{B}^\alpha |\psi\rangle$$  

(8)

where $H_0 = H - H'$.

In order to make progress we assume that each perturbation term $A^\alpha \bar{P} \bar{B}^\alpha$ acting on the ground state creates a
set of nearly degenerate excited states, with average energy $E_\alpha$. This assumption is equivalent to saying that the spectral function associated with each term is dominated by a narrow peak at $E_\alpha$. This significant approximation is reasonable because the correction to the density matrix is only used to enlarge the basis, to improve DMRG convergence. Correspondingly, we approximate $(E_0 - H_0)^{-1}$ as $(E_0 - E_\alpha)^{-1} \approx 1/\varepsilon_\alpha$. This gives

$$|\psi'\rangle \approx \sum_s \psi_s \sum_{\alpha} \frac{t_\alpha}{\varepsilon_\alpha} \hat{A}^\alpha \hat{P} \hat{B}^\alpha |L_s\rangle |R_s\rangle. \quad (9)$$

There are no first order corrections to the density matrix from $|\psi'\rangle$, since $\hat{P}|\psi\rangle = 0$. The lowest order correction to $\rho$ can be written as

$$\Delta \rho = \sum_{ss'} \psi_s \psi^*_s \sum_{\alpha\alpha'} \frac{t_{\alpha}}{\varepsilon_\alpha} \frac{t_{\alpha'}}{\varepsilon_{\alpha'}} \hat{A}^\alpha \hat{P} \hat{B}^\alpha |L_s\rangle \langle L_{s'}| \hat{A}^{\alpha'} \hat{M}_{s's'a'a'} \quad (10)$$

where

$$M_{s's'a'a'} = \langle R_{s'}| \hat{B}^{\alpha'} \hat{P} \hat{B}^\alpha |R_s\rangle \quad (11)$$

Here if $A$ is the unit operator, the term adds nothing to the basis. If $B$ is the unit operator, $M_{s's'a'a'}$ vanishes. For the nontrivial pairs of operators $A$ and $B$, this matrix element somewhat resembles a correlation function and it is natural to assume that the diagonal terms are dominant, where $\alpha = \alpha'$ and $s = s'$. We expect the off diagonal terms $\alpha \neq \alpha'$ to describe coherence between different perturbation terms which would tend to reduce the number of basis functions needed to describe the system block; therefore, ignoring the offdiagonal terms is a conservative assumption. Accordingly, we take

$$M_{s's'a'a'} \approx \delta_{ss'} \delta_{\alpha\alpha'} \delta_{\alpha a} \quad (12)$$

This gives Eq. (4) with $a_\alpha = b_\alpha |t_\alpha|^2 / \varepsilon^2_\alpha$, and where we omit block-Hamiltonian terms.

In practice, we take $a_\alpha$ to be a small constant independent of $\alpha$. Construction of the correction to $\rho$ take a calculation time for a single step proportional to $m^3$ times the number of connecting terms, which is typically significantly smaller than the other parts of the DMRG calculation, although the scaling is the same. Larger values of $a$ introduce more “noise” into the basis, speeding convergence, but also limiting the final accuracy. Note that it is just as easy to apply the correction within the two-site method as the single-site method, which may be useful in some very difficult cases. We do not present results for this combination here.

As a first test calculation, we consider the $S = 1$ Heisenberg model

$$H = \sum_j \tilde{S}_j \tilde{S}_{j+1}, \quad (13)$$

where we have set the exchange coupling $J$ to unity. The correction consist of the following: for each boundary site $i$ of a block, i.e. a site directly connected to the other block, we add into the density matrix

$$\Delta \rho = a(S^+_i \rho S^-_i + S^-_i \rho S^+_i + S^z_\alpha \rho S^z_\alpha). \quad (14)$$

For a chain with open boundaries, there is one site $i$; for periodic boundaries, there are two. One could argue that this expression should be adjusted with factors of 2 between the $z$ term and the other two terms, but this is not likely to make a significant difference. Note that the $S^+$, $S^-$ terms automatically increase the range of quantum numbers (i.e. total $S^z$) with nonzero density matrix eigenvalues. Figure 2 shows the convergence of the energy for a 100 site chain with open boundaries as a function of the sweep, keeping $m = 50$ states, relative to the numerically exact result obtained with $m = 200$ and 10 sweeps. One can see the excellent convergence of the standard approach. The single-site method without corrections does not do too badly in this case, but still gets stuck significantly above the two-site energy. Adding the corrections, in this case with $a = 10^{-4}$, dramatically improves the convergence, making the single site method converge nearly as fast as the two site method. The two site method is roughly a factor of three slower than the single site method. Thus, even in this simple 1D case where the standard approach works extremely well, there are advantages to using the corrected single site method.

The results change significantly if we consider periodic boundary conditions. Here we consider the same superblock configuration as with open boundary conditions, but simply add in the connection to the Hamiltonian between the first and last sites. There are better configurations for periodic boundaries, such as considering it to be a ladder with the interchain couplings turned off except at the ends. These other configurations are superior only in the sense of improved convergence with
FIG. 3: Error in the total energy for a 100 site Heisenberg spin-one chain, with periodic boundaries. The number of states kept per block is indicated, and is the same for all three methods; four sweeps were made for each m. The correction parameter a was taken to be $10^{-4}$ for sweeps 1 - 8, and $10^{-6}$ for later sweeps. A somewhat slower convergence is visible for $a = 10^{-6}$. The reference energy used was 100 times the infinite energy per site, $-1.401484038971(4)$\[4\]. The corrected single site method using $m = 4000$ states gives a slightly lower total energy, due to exponentially small finite size effects, of $-140.14840390392$.

the number of sweeps, not improved with respect to the number of states for a large number of sweeps. This naive configuration thus provides a difficult test for the single site method with corrections. In Fig. 3, we show the results for the same three cases as in Fig. 2. In this case, in the early sweeps, both uncorrected methods are stuck, ignoring the extra link between the first and last sites. The extra link eventually appears in the basis, but there is still sticking two or three times in higher energy states. In contrast, the corrected single site method never gets stuck and shows excellent convergence.

A very useful DMRG technique is the extrapolation of the energy with the truncation error, i.e. the weight in the states which are thrown out. If the truncation error were measured exactly, with a complete basis for the environment, then the energy error would be proportional to the truncation error, allowing a linear extrapolation to zero truncation error. In practice, the apparent truncation error from the two site method may often be an underestimate, but one often finds that it is very consistent and still allows excellent extrapolation, even on fairly wide ladders. The truncation error within the corrected single-site method depends on a: as $a \to 0$, the apparent truncation error goes to zero and is unrelated to the exact truncation error. However, if a is not too small, linearity and excellent extrapolation are possible.

Figure 4 shows results for the 100 site periodic system with a larger value of a, $10^{-2}$, suitable for extrapolation. The results show excellent linearity. The extrapolation gives -140.148416, off by $1.2 \times 10^{-5}$, whereas the sweep with $m = 340$ gave -140.148279, off by $1.2 \times 10^{-4}$. We have found that typically an order of magnitude improvement in the estimate for the energy is obtained by extrapolation in good cases; here we see similar improvement.

In performing these extrapolations one always need to check the linearity for the system being studied.

In summary, we have demonstrated a correction to the density matrix which allows the single-site DMRG method to converge well, and which improves the convergence dramatically for hard-to-converge systems.

We thank J. Rissler and F. Verstraete for helpful conversations. We acknowledge the support of the NSF under grant DMR03-11843.

[1] S.R. White, Phys. Rev. Lett. 69, 2863 (1992); Phys. Rev. B 48, 10345 (1993).
[2] U. Schollwöck, Rev. Mod. Phys. 77, 259 (2005).
[3] For example, see S.R. White and D.J. Scalapino, Phys. Rev. Lett. 80, 1272 (1998).
[4] S.R. White and D.A. Huse, Phys. Rev. B 48, 3844 (1993).