Quantum data compression, quantum information generation, and the density-matrix renormalization-group method

Ö. Legeza and J. Sólyom

Research Institute for Solid State Physics, P. O. Box 49, H-1525 Budapest, Hungary

We have studied quantum data compression for finite quantum systems where the site density matrices are not independent, i.e., the density matrix cannot be given as direct product of site density matrices and the von Neumann entropy is not equal to the sum of site entropies. Using the density-matrix renormalization-group (DMRG) method for the one-dimensional Hubbard model, we have shown that a simple relationship exists between the entropy of the left or right block and dimension of the Hilbert space of that block as well as of the superblock for any fixed accuracy. The information loss during the RG procedure has been investigated and a more rigorous control of the relative error has been proposed based on Kholevo's theory. Our results are also supported by the quantum chemistry version of DMRG applied to various molecules with system lengths up to 60 lattice sites. A sum rule that relates site entropies and the total information generated by the renormalization procedure has also been given, which serves as an alternative test of convergence of the DMRG method.

I. INTRODUCTION

In the past few years nonlocal generalizations[1,2] of the density-matrix renormalization-group (DMRG) method[3,4] have gained much attention since the method allows us to study interacting spin and electron systems in momentum space[1,5,6] or molecules within the context of quantum chemistry.[2,6–11]

Recently the concepts of information theory have reappeared in the study of solid state physics and statistical physics problems.[6,12–15] On the one hand, it has been pointed out[12,13] that the von Neumann entropy of a finite system of length N,

$$S(\rho^{(N)}) = - \mathrm{Tr}(\rho^{(N)} \ln \rho^{(N)}), \qquad (1)$$

where $\rho^{(N)}$ is the density matrix, diverges logarithmically with the system size if the system is critical and the spectrum is gapless, while it saturates as N increases for noncritical, gapped models. On the other hand, in an independent work,[6] the von Neumann entropy has been used to improve the performance of the momentum space and quantum chemistry versions of the DMRG method. It has been shown that an optimal ordering of states can be obtained from the distribution of site von Neumann entropies of "lattice sites." Quite recently, the study of entanglement has also led to an alternative procedure using the periodic boundary condition[16] that increases dramatically the performance of the real space DMRG (RS-DMRG) method and to the development of the time-dependent DMRG algorithm.[17,18]

An important problem in quantum information theory is quantum data, compression,[19–22] i.e., how the dimension of the typical subspace should be chosen to achieve a prescribed fidelity. The DMRG method is based on a similar reduction of the Hilbert space. In the standard DMRG procedure, the dimension, the number of states to be kept, is fixed before the calculation. In the dynamical block-state selection (DBSS) approach[10] the truncation error is kept fixed and the dimension is allowed to vary. It has been shown[6] that

the prescribed accuracy can be achieved with the least number of block states, if the entropy-optimized ordering is used. In this paper, we study this problem from the point of view of quantum data compression and show that a simple relationship exists between the von Neumann entropy of subsystem blocks and the size of Hilbert space of the blocks or alternatively the size of Hilbert space that one has to diagonalize in DMRG. We argue that the accessible information[23,24] of mixed-state ensembles can be interpreted in the context of DMRG as the information loss due to the truncation procedure. Relying on this we propose an alternative method to improve the convergence of DMRG.

The main purpose of this paper is to study the properties of DMRG from the point of view of quantum information theory and to use the obtained results to optimize the method. The setup of the paper is as follows. In Sec. II we describe the theoretical background of data compression, accessible information for mixed-state ensembles, and their relation to the DMRG method. In Sec. III these concepts and quantities are studied by applying the RS-DMRG to the half-filled one-dimensional (1D) Hubbard chain using open and periodic boundary conditions and the nonlocal generalization of DMRG to various molecules. In Sec. IV the obtained results are interpreted as quantum information generation in the DMRG method. Our conclusions are presented in Sec. V.

II. QUANTUM DATA COMPRESSION

In his landmark work,[25] Shannon has shown how much a message constructed from N independent letters (x_α), where each letter occurs with *a priori* probability p_α, can be compressed. In classical information theory messages are classified into so-called typical and atypical sequences. If for any $\epsilon > 0$ the sum of the probabilities of all typical sequences lies between $1 - \epsilon$ and 1, then for any $\delta > 0$ and for large enough values of N the number of typical sequences $n(\epsilon, \delta)$ lies between the bounds $2^{N(S+\delta)} \geq n(\epsilon, \delta) \geq (1-\epsilon)2^{N(S-\delta)}$, where S is

the Shannon entropy, $S=-\Sigma_\alpha p_\alpha \log_2 p_\alpha$. Therefore, a block code of length NS bits encodes all typical sequences irrespective of how the atypical sequences are encoded and the probability of error will still be less than ϵ. In quantum information theory the letters are density matrices and one has to distinguish two cases, namely when the density matrices correspond to ensembles of q pure states, $|\phi_\alpha\rangle$, or when they are formed from density matrices ρ_α, with probability p_α.

A. Pure-state ensemble

Considering the first case, that is, a pure-state ensemble, the density matrix of a message consisting of N letters is $\rho^{(N)}=\rho\otimes\rho\otimes\cdots\otimes\rho$, where $\rho=\Sigma_\alpha p_\alpha|\phi_\alpha\rangle\langle\phi_\alpha|$, and the von Neumann entropy of the message is simply related to the entropy of the ensemble, $S(\rho^{(N)})=NS(\rho)$. The optimal code that compresses the Hilbert space of the entire message, $\Lambda^{(N)}=\Lambda\otimes\Lambda\otimes\cdots\otimes\Lambda$ to a smaller Hilbert space without compromising the fidelity of the message for $N\rightarrow\infty$ has been obtained by b. Schumacher.[19] Analogously to classical information theory $\Lambda^{(N)}$ is divided into so-called typical ($\Lambda_{\mathrm{typ}}^{(N)}$) and atypical ($\Lambda_{\mathrm{atyp}}^{(N)}$) subspaces by applying a projection Π_{typ} and Π_{atyp}. If for any $\epsilon>0$ almost all weight of the ensemble lies within $\Lambda_{\mathrm{typ}}^{(N)}$ and $\mathrm{Tr}\,\Pi_{\mathrm{typ}}\rho^{(N)}\Pi_{\mathrm{typ}}>1-\epsilon$, while for the atypical subspace $\mathrm{Tr}\,\Pi_{\mathrm{atyp}}\rho^{(N)}\Pi_{\mathrm{atyp}}<\epsilon$, then for any $\delta>0$ and sufficiently large enough N the eigenvalues ω_{typ} of $\rho^{(N)}$ belonging to typical eigenstates fall within a narrow range:

$$e^{-N[S(\rho)+\delta]}<\omega_n<e^{-N[S(\rho)-\delta]}. \quad (2)$$

Therefore, the number of dimensions of the typical subspace lies between the bounds

$$(1-\epsilon)e^{N[S(\rho)-\delta]}\leqslant \dim\Lambda_{\mathrm{typ}}^{(N)}\leqslant e^{N[S(\rho)+\delta]}. \quad (3)$$

This means that the von Neumann entropy is the number of qubits of quantum information carried per letter of the message and unless $\rho=(1/q)\mathbb{1}$, a compression is always possible. We have to mention that Eqs. (2) and (3) have been reformulated in order to be consistent with Eq. (1).

B. Mixed-state ensemble

The result is less simple if the letters are chosen from density matrices of mixed states. In this case only an upper[23] or lower bound[26] can be derived for the accessible information. Therefore, if the source of information is constructed from messages represented by ρ_a states and *a priori* probability p_a, then the mutual information between the sender's and receiver's measurement is bounded by

$$I\leqslant S(\rho)-\sum_a p_a S(\rho_a), \quad (4)$$

where $\rho=\Sigma_a p_a\rho_a$ and S is the von Neumann entropy given by Eq. (1). If all signal states ρ_a are pure states, the upper bound on the accessible information reduces to $I\leqslant S(\rho)$.

The lower bound on accessible information also depends not only on the entire density matrix but also on the particular way ρ is realized as an ensemble of mixed states. As it has been given by Jozsa *et al.*,[26]
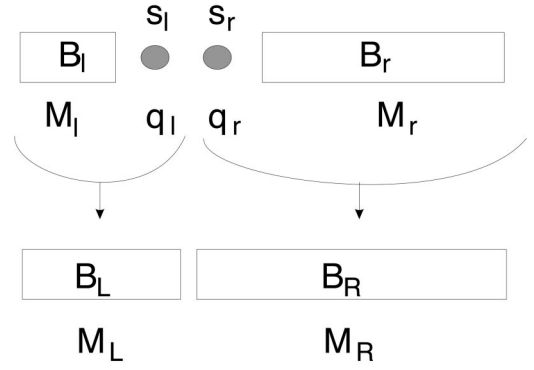

FIG. 1. Schematic plot of the system and environment block of DMRG. B_l and B_r denote the left and right block of length l and r, and of dimension M_l and M_r, respectively, • stands for the intermediate sites (s_l and s_r) with q_l and q_r degrees of freedom. The blocks $B_L=B_l\bullet, B_R=\bullet B_r$ have dimension M_L and M_R, respectively.

$$I\geqslant Q(\rho)-\sum_\alpha p_\alpha Q(\rho_\alpha), \quad (5)$$

where the subentropy Q is defined as

$$Q=-\sum_{\alpha=1}^{M}\left(\prod_{\alpha\neq\beta}\frac{\omega_\alpha}{\omega_\alpha-\omega_\beta}\right)\omega_\alpha\ln\omega_\alpha. \quad (6)$$

If two or more of the eigenvalues are equal, Q remains finite if one takes the limit $\omega_\alpha\rightarrow\omega_\beta$.

C. Relationship to DMRG

For lattice models studied in solid state physics problems, on the other hand, the message coded into the wave function of the system has different features. The site density matrices are in general not independent, thus $\rho^{(N)}\neq\rho\otimes\rho\otimes\cdots\otimes\rho$ and $S(\rho^{(N)})\neq NS(\rho)$. In this paper, this situation is studied using DMRG.

From the point of view of information theory DMRG is a numerical tool to select the typical subspace on which the target state (Ψ_{TG}) of a Hamiltonian of a finite system of length N can be represented. The Hilbert space of the system (called superblock Hilbert space) is defined on a bipartite system, $\Lambda_{\mathrm{typ}}=\Lambda_{\mathrm{typ}}^{(L)}\otimes\Lambda_{\mathrm{typ}}^{(R)}$, where $\Lambda_{\mathrm{typ}}^{(L)}$ and $\Lambda_{\mathrm{typ}}^{(R)}$ are the typical Hilbert spaces of the left and right blocks, B_L and B_R, which themselves are built from subblocks, B_l and B_r, with one extra site. The indices l and r denote at the same time the number of sites in the subblocks. The schematic plot of the DMRG configuration is shown in Fig. 1.

Although the target state is usually a pure state, the left and right blocks are in a mixed state described by the density matrices $\rho_L=\mathrm{Tr}_R\rho$, and $\rho_R=\mathrm{Tr}_L\rho$, respectively. Analogously one can define the density matrices of the B_l and B_r subblocks, $\rho_l=\mathrm{Tr}_{s_l}\mathrm{Tr}_R\rho$ and $\rho_r=\mathrm{Tr}_L\mathrm{Tr}_{s_r}\rho$, as well as that of the intermediate sites, given as $\rho_{s_l}=\mathrm{Tr}_l\mathrm{Tr}_R\rho$, $\rho_{s_r}=\mathrm{Tr}_L\mathrm{Tr}_r\rho$. It follows from the singular value decomposition theorem that for a pure target state and any choice of the length of the left subblock, l, the entropy of the left block of length $l+1$, $S_L(l+1)$ is identical to that of the right block, $S_R(r+1)$, where

$l+r+2=N$. $S \equiv S_L = S_R$ is also related to the mutual information of the blocks, $I = -2S$ since $S_L(l+1) + S_R(r+1) + I = S(N)$, where $S(N)$ is the entropy of the full system, which vanishes if the system is in a pure state. I quantifies the correlation between the two subsystems and vanishes if and only if the two subsystems are completely uncorrelated, i.e., when $\rho = \rho_L \otimes \rho_R$. S also measures the entanglement[27] and the amount of quantum information exchange between the blocks.[28]

The wave function of the finite lattice of length N is built up using the infinite lattice method followed by a systematic application of the sweeping procedure of the finite lattice method. The typical subspaces $\Lambda_{\mathrm{typ}}^{(L)}$ and $\Lambda_{\mathrm{typ}}^{(R)}$ are obtained by selecting that M renormalized states out of the $M \times q$ block states $(|\phi_\alpha^{(L)}\rangle, |\phi_\alpha^{(R)}\rangle)$, which have the largest eigenvalues, $\omega_\alpha^{(L)}$ and $\omega_\alpha^{(R)}$, of the density matrices ρ_L and ρ_R, respectively. The Π_{typ}^L and Π_{typ}^R projection operators are formed from the M chosen eigenstates of ρ_L and ρ_R. The accuracy of the truncation procedure is given by the so-called truncation error, $\mathrm{TRE} = 1 - \sum_{\alpha=1}^M \omega_\alpha$, where $\omega_\alpha \equiv \omega_\alpha^{(L)}$ or $\omega_\alpha^{(R)}$, depending on the direction of the sweep. Therefore, due to the truncation of basis states ϵ is always greater than zero.

Another main difference between conventional information theory and DMRG is that while in the former case the letter states have an *a priori* assigned probability density,[29,30] in the latter case the probability of a configuration depends on the target state. Therefore, the subset of states that define the code words is determined by the physical properties of the target state. When DMRG is applied to find the typical subspace of a target state, i.e., the typical code words, this subspace is reduced further. In addition, during the RG procedure the original basis states are also transformed into new renormalized states, thus code words change at each iteration step. In this respect, in DMRG the information of one block is coded into the basis function of the renormalized blocks rather than into a quantum channel of qubits.

In the standard DMRG procedure the number of states of the left and right blocks, M, is fixed in advance of the calculation. The superblock Hilbert space, Λ_{SB}, on which the target state is determined is not simply $\Lambda_{\mathrm{typ}}^{(L)} \otimes \Lambda_{\mathrm{typ}}^{(R)}$, since only states with prescribed quantum numbers have to be considered. Usually M is increased systematically during the finite lattice method; thus the accuracy depends on M in an uncontrolled way. In contrast to this in the DBSS approach[10] the truncation error is kept fixed and the number of retained states is varied dynamically, which allows us to control the accuracy more precisely. A threshold value for the minimum number of block states M_{min} has also been introduced in order to avoid higher-lying local minima.

III. NUMERICAL RESULTS

The most general fermionic Hamiltonian can be given as

$$\mathcal{H} = \sum_{ij\sigma} T_{ij} c_{i\sigma}^\dagger c_{j\sigma} + \sum_{ijkl\sigma\sigma'} V_{ijkl} c_{i\sigma}^\dagger c_{j\sigma'}^\dagger c_{k\sigma'} c_{l\sigma}, \qquad (7)$$

where T_{ij} denotes the matrix elements of the one-particle Hamiltonian and V_{ijkl} stands for the matrix elements of the
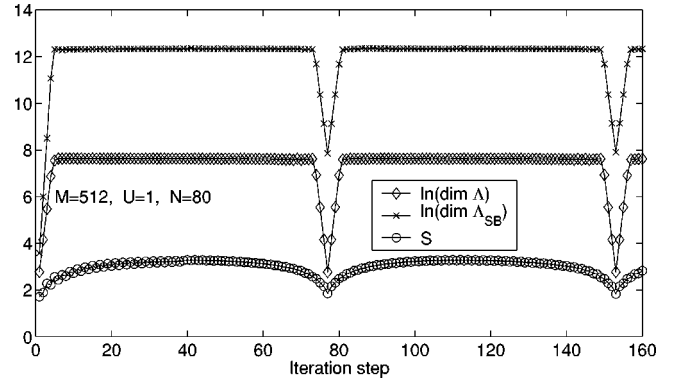

FIG. 2. Block entropy, S, and $\ln(\dim\Lambda)$, $\ln(\dim \Lambda_{SB})$ as a function of iteration steps for the half-filled 1D Hubbard model with $N=80$ for $U=1$ with fixed number of block states ($M=512$) using PBC. For an iteration step less than 39 the infinite lattice method has been used.

electron interaction operator. Depending on the structure of T_{ij} and V_{ijkl} this Hamiltonian can describe fermionic models in real space with open or periodic boundary conditions, a molecule, or a usual fermionic model in momentum space.

A special case of Eq. (7) is the 1D Hubbard model,

$$\mathcal{H} = \sum_{i\sigma} t(c_{i\sigma}^\dagger c_{i+1\sigma} + c_{i+1\sigma}^\dagger c_{i\sigma}) + U \sum_i c_{i\uparrow}^\dagger c_{i\uparrow} c_{i\downarrow}^\dagger c_{i\downarrow}. \qquad (8)$$

In this case the $|\phi_\alpha\rangle$ states are $|0\rangle, |\downarrow\rangle, |\uparrow\rangle, |\downarrow\uparrow\rangle$. For homogeneous lattice models with periodic boundary conditions each lattice site carries the same information (S_i). For $U=0$, $S_i = 4 \ln 4 = 1.38629$, while for $U \to \infty$ the model is equivalent to the spin-$1/2$ Heisenberg model and only the $|\downarrow\rangle$ and $|\uparrow\rangle$ states have finite weight, thus $S_i = 2 \ln 2 = 0.69314$. In the numerical work we have used the real space version of DMRG with open (OBC) and periodic (PBC) boundary conditions.

A. The standard DMRG procedure

First we have investigated the block entropy in the standard RS-DMRG method when calculating the ground-state energy of the half-filled 1D Hubbard model with $t=1$ for a chain with $N=80$ sites, for $U=1, 10, 100$ using a fixed number of block states, namely, $M=256, 512, 1024$. In Fig. 2 we have plotted the block entropy and the logarithm of the dimension of Λ and the Hilbert space of the superblock, $\ln(\dim \Lambda)$ and $\ln(\dim \Lambda_{SB})$, respectively, as a function of iteration steps for $U=1$ and $M=512$. It is clear from the figure that for the *infinite lattice* method (for iteration steps less than 39), the block entropy, S, increases with the block length. According to Vidal and co-workers[12,13] a logarithmic divergence is expected if no truncation had been applied during the RG procedure. Deviations occur for two reasons. First the truncation procedure has been applied from the fifth iteration step. More importantly, in DMRG the chains are built up systematically, and for chains of length N the finite lattice method has to be used after $N/2-1$ initialization steps. At this point the entropy of the blocks starts to decrease. In the successive sweeps the entropy reaches its maximum when the sizes of the two blocks become equal. It
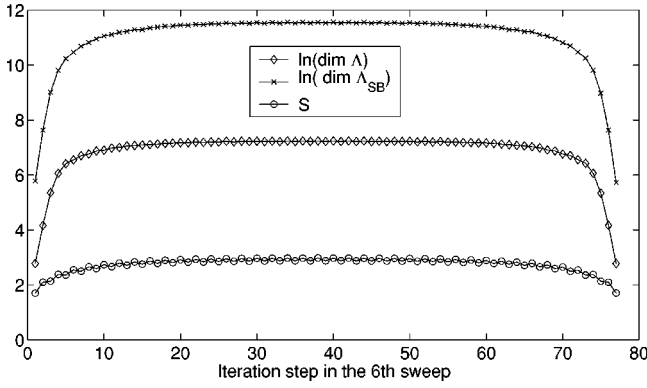
FIG. 3. Same as Fig. 2 but using DBSS approach for $\text{TRE}_{max} = 10^{-4}$ and $M_{min} = 16$.

FIG. 4. Parameter β as a function of rescaled iteration step using the DBSS approach for $\text{TRE}_{max} = 10^{-4}$, $M_{min} = 16$, $U = 1$, $N = 40, 60, 80$.

is worth mentioning that S became almost totally symmetric after the second half sweep already. Since S is related to the mutual information of the blocks, this result is in agreement with the long known fact that the best accuracy of RS-DMRG is always obtained for the symmetric block configurations. For larger U values the entanglement between the blocks is reduced as indicated by the decreasing value of S and S_i, while $\ln(\dim \Lambda)$ remained almost the same, and the relative error of the calculations improved significantly. This is the first sign that the entropy, the dimension of the Hilbert space, and the accuracy are related to each other in DMRG.

B. DBSS approach

This standard procedure, however, does not allow an *a priori* control of the accuracy, while this control is possible in the DBSS approach.[10] Moreover, the latter approach allows us to study the relationship between the von Neumann entropy of the block and the dimension of the Hilbert space of the superblock qualitatively. Therefore, we have repeated the DMRG calculations on the same model using the DBSS approach with $M_{min} = 16$ for various values of the truncation error, namely for $\text{TRE}_{max} = 10^{-2}, 10^{-3}, \dots, 10^{-8}$. In Fig. 3 we have plotted $\ln(\dim \Lambda)$, $\ln(\dim \Lambda_{SB})$, and the block entropy obtained after the sixth sweep for $U = 1$, $N = 80$ as a function of iteration step shifted modulo 80. It is clearly seen in the figure that due to the truncation of basis states, the logarithmic corrections are cut, and S and $\ln(\dim \Lambda)$ are almost constant for a wide range of block lengths. A simple relationship seems to exist between the two quantities, which is analogous to Eq. (3). We define the shift between $\ln(\dim \Lambda)$ and S as

$$\beta \equiv \ln(\dim \Lambda) - S. \tag{9}$$

Figure 4 shows the value of β for $U = 1$, $\text{TRE}_{max} = 10^{-4}$ for three different chain lengths, $N = 40, 60, 80$. β is plotted as a function of the shifted iteration step rescaled by the length of the chain. We have found that after the sixth sweep except for very short block lengths β is practically independent of N for a given accuracy. A similar result is expected for noncritical systems since according to Vidal *et al.*[12] in this case the block entropy saturates with increasing block length. It is
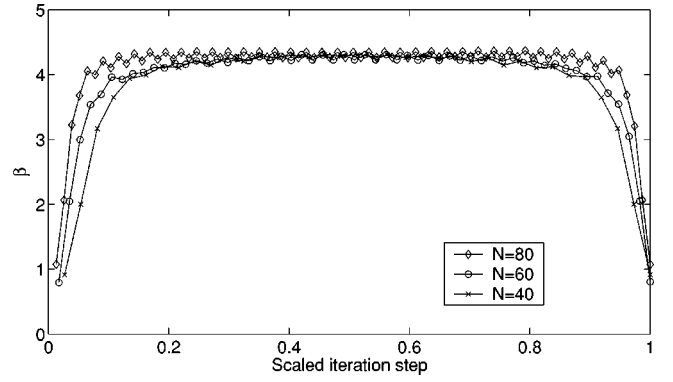
worth mentioning that β' defined analogously to Eq. (9) but with $\ln(\dim \Lambda_{SB})$,

$$\beta' \equiv \ln(\dim \Lambda_{SB}) - S, \tag{10}$$

is also independent of N for large enough N.

Both β and β' are, however, functions of TRE_{max}. This dependence has been tested for $U = 1$ for $\text{TRE}_{max} = 10^{-3} - 10^{-6}$, when $\max[\ln(\dim \Lambda_{SB})]$ changed from 6×10^3 to 2.1×10^6. In this range of the truncation error β increases proportionally to $-\ln \text{TRE}_{max}$. Similar results have been obtained with open boundary conditions but with significantly smaller block entropies and much less block states. For larger U values and for small values of TRE_{max} the target state has often been lost and the convergence depended very much on the minimum number of block states M_{min}.

C. An alternative truncation procedure

The simple relationships given by Eqs. (9) and (10) seem to indicate that a better DMRG procedure that avoids the above-mentioned problems should rely on the control of the von Neumann entropy of the blocks. In order to control the weight of retained information during the RG procedure in a more rigorous way, the reduced density matrix of the left or right subsystem is written as the sum of the density matrices of mutually orthogonal mixed states belonging to the typical and atypical subspaces,

$$\rho^{(L)} = p_{\text{typ}}\rho_{\text{typ}}^{(L)} + (1 - p_{\text{typ}})\rho_{\text{atyp}}^{(L)}, \tag{11}$$

where $\rho_{\text{typ}}^{(L)}$ is formed from the M largest eigenvalues of $\rho^{(L)}$ and $\rho_{\text{atyp}}^{(L)}$ from the remaining eigenvalues with $\text{Tr }\rho_{\text{typ}}^{(L)} = \text{Tr }\rho_{\text{atyp}}^{(L)} = 1$. A similar decomposition holds for the right block, and therefore in what follows the superscript L and R are dropped. In usual information theory, if the message contained ρ_{typ} or ρ_{atyp} with the appropriate probabilities, the accessible information for such a binary channel would be less than the Kholevo bound[23]

$$I \leq S(\rho) - p_{\text{typ}}S(\rho_{\text{typ}}) - (1 - p_{\text{typ}})S(\rho_{\text{atyp}}), \tag{12}$$
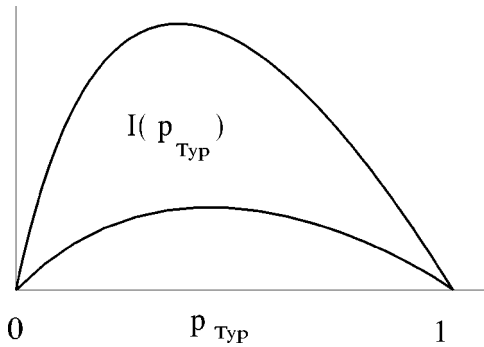
and larger than the Jozsa-Robb-Wootters lower bound[26]

FIG. 5. Schematic plot of the upper and lower bonds on accessible information for a binary channel.

$$I \geq Q(\rho) - p_{\text{typ}}Q(\rho_{\text{typ}}) - (1 - p_{\text{typ}})Q(\rho_{\text{atyp}}). \quad (13)$$

The schematic plot of the two bounds as a function of p_{typ} is shown in Fig. 5. It is worth mentioning that other ensemble-dependent bounds have been derived by Fuchs and Caves.[24]

In DMRG, however, the atypical subspace is neglected, which leads to a loss of information. For p_{typ} close to unity this loss is argued to be equal to the accessible information, which is negligible if the probability of the atypical subspace is negligible. This can be used to optimize the DMRG procedure. A modified DBSS procedure is proposed, where instead of fixing the truncation error, the Kholevo bound is required to be less than an ϵ fixed in advance, and the number of states M is chosen accordingly in every step.

We have run independent calculations for the 1D Hubbard model for different threshold values on the upper bound of accessible information, denoted by χ, ranging from 10^{-2} to 10^{-7} using three sweeps. In Fig. 6 we have plotted the relative error as a function of TRE and χ on a logarithmic scale for $U=1, 10, 100$ with $M_{min}=4$, where the relative error is given as $(E_{\text{DMRG}} - E_{\text{exact}})/E_{\text{exact}}$. We have found that this selection rule is very stable even for very small values of the retained eigenvalues of ρ_{typ}. Assuming that the linear relationship apparent on the right panel holds for other models as well, an alternative extrapolation procedure can be obtained. Instead of using the truncation error as proposed in Refs. 3
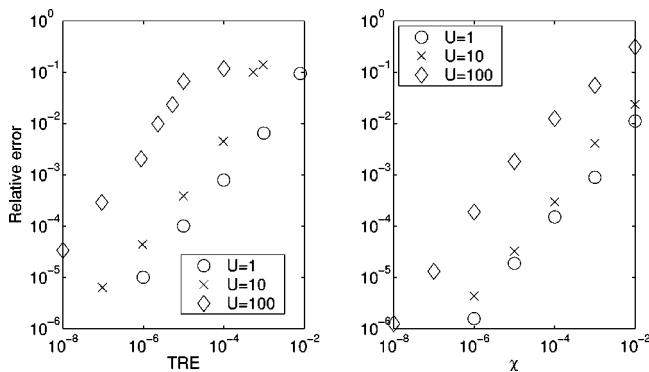

FIG. 6. The relative error as a function of the truncation error and the threshold value of Kholevo's bound on accessible information obtained with the DBSS method and by Eq. (12), respectively.
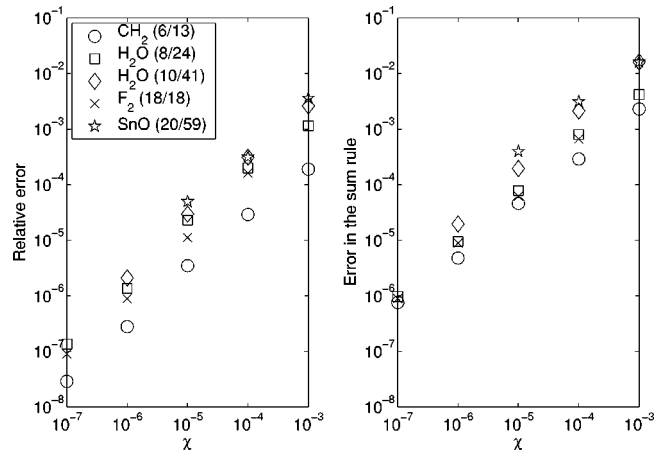

FIG. 7. The first panel shows the relative error as a function of the threshold value of Kholevo's bound on accessible information obtained by Eq. (12). The second panel shows the error in the sum rule given by Eq. (19).

and 31–34, the energy of the target state is extrapolated as a function of χ.

D. Application to quantum chemistry DMRG

We have tested the truncation procedure presented above by performing DMRG calculations on various molecules up to 59 lattice sites. We have used the dynamically extended active space procedure[6] (DEAS) and the DBBS approach by controlling the change of von Neumann entropy during the truncation procedure at each renormalization step. In all calculations $M_{min}=4$ was used and during the first half sweep the right block was represented by 256–512 states. The maximum number of block states that our program could treat is in the range of 2000–3500, strongly depending on the system size. Our results on various molecules are shown in Fig. 7. We have found again that on a logarithmic scale the relative error is a linear function of the accessible information even for large threshold values.

For very small system sizes, like CH_2 (6/13) the number of block states has been overestimated by using the upper bound on the accessible information. Therefore, we have also tried to control the lower bound on accessible information. For large threshold values, the lower bond on accessible information has also provided similar results but with less block states. For smaller threshold values, however, we have found it to be very unstable for increasing M with larger values of p_{typ}.

When molecules are studied in quantum chemistry applications, "lattice sites" carry different quantum information as described by a nonconstant value of the site entropy function. This means that the entropy profile of the bipartite partitioning of the finite system depends very much on the ordering of the lattice sites. Since the molecules are noncritical models, one can obtain the entropy profile of the exact solution of the finite system with an exponentially small error using a limited number of block states. Since the block entropy is related to the size of the superblock Hilbert space, in order to increase the efficiency of the DMRG method one

has to reduce the block entropy profile of the exact solution by optimizing ordering. For larger systems we have found that the optimization procedure outlined in Ref. 6, based on the site entropy profile alone, does not lead to optimal ordering and it often blocked the convergence of the DMRG method. This problem, related to the competition between entanglement localization and interaction localization, will be investigated in a subsequent paper.

IV. RELATIONSHIP TO QUANTUM INFORMATION GENERATION

In the DMRG procedure, during the renormalization step the block B_L is formed of the subblock B_l and the $(l+1)$th site. Denoting by $S_L(l)$ the entropy of the left subblock of length l and by S_{l+1} the entropy of the $(l+1)$th site, the change of the block entropy by forming a larger block, $B_L(l+1)$, is given as

$$S_L(l) + S_{l+1} + I_L(l) = S_L(l+1), \quad (14)$$

where the so-called mutual information $I_L(l)$ quantifies the correlation between the subsystem and the site. A similar relation holds for the right block, given as

$$S_R(r) + S_{l+2} + I_R(r) = S_R(r+1). \quad (15)$$

The unitary operation applied on the basis states of the $B_l \bullet$ composite system is a type of local quantum operations or classical communications (LOCC) operation, i.e., it cannot increase the entanglement between B_L and B_R blocks which has also been related to entanglement distillation protocols.[35] Therefore, the entropy of the $B_l \bullet$ composite system remains unchanged or decreased by forming the larger block B_L. The quantum information generated by the renormalization procedure of the forward sweep can be measured using Eq. (14) as

$$I_L(l) = S_L(l+1) - S_L(l) - S_{l+1}, \quad (16)$$

where l runs from 1 to $N-1$. Analogously the information generated by the backward sweep can be derived using Eq. (15).

If an effective system of length $N+2$ is formed by adding two noninteracting sites to the right end of the chain, all blocks containing 1 to N lattice sites of the original system can be formed by the forward sweep. The total information gain of a full half sweep can be calculated as $\Sigma_{l=1}^{N-1} I_L(l)$. The same holds for the backward sweep as well when the two noninteracting sites are attached to the left end of the chain. It is easy to show that if all q^l and q^r basis states of the blocks are kept at each iteration step, i.e., no truncation is applied, a sum rule holds, which relates the total information gain within a full half sweep and the sum of site entropies given as

$$\sum_{l=1}^{N-1} I_L(l) = -\sum_{l=1}^{N} S_l, \quad (17)$$

where we have used $S_L(1)=S_1$ and $S_L(N)=0$. This equality, however, does not hold in practical DMRG calculations.

First the blocks contain only a limited subset of the block states; thus for a given site S_l changes for each half sweep as the method converges to the attractor. In addition, during the renormalization process $S_L(l+1)$ is reduced to $S_L^{\text{Trunc}}(l+1)$ due to the truncation of basis states, thus $I_L(l)$ is also a function of subsequent sweeps. However, once the DMRG method has converged, i.e., subsequent DMRG sweeps do not change $S_L(l)$ and S_l, the following equality should hold to a good accuracy:

$$\sum_{l=1}^{N-1} I_L(l) \simeq -\sum_{l=1}^{N} S_l + \sum_{l=2}^{N} [S_L(l) - S_L^{\text{Trunc}}(l)]. \quad (18)$$

An analogous relationship holds for the backward sweep as well.

As one sees in Fig. 5 the decrease of the loss of information goes together with the decrease of the contribution of the atypical states. Although the actual shape of the upper and lower bounds depends on U, we have found that when only the first few eigenvalues of ρ are retained, p_{typ} is already very close to unity. Therefore, the loss of information could be equally controlled by requiring that $\chi \equiv S(\rho) - S(\rho_{\text{typ}})$ should become less than ε in which case Eq. (18) leads to

$$\sum_{l=1}^{N-1} I_L(l) + \sum_{l=1}^{N} S_l < (N-1)\epsilon. \quad (19)$$

Equation (19) can be used as an alternative procedure to check the convergence of the DMRG method.

Numerically we have found a very similar result to the one shown in the preceding section, that the relative error and $\Sigma_l I(l) + \Sigma_l S_l$ are linear functions of χ on a logarithmic scale. The second panel of Fig. 7 shows the error in the sum rule given by Eq. (19) for the calculation presented in Sec. III D. This has been found to be one order of magnitude worse than the discarded weight of the von Neumann entropy, as expected.

V. SUMMARY

We have studied the density-matrix renormalization-group method from the point of view of quantum data compression. The method has been applied to the half-filled 1D Hubbard model in real space and to various molecules. Our findings are listed below.

(1) We have shown an explicit relationship between the dimension of the Hilbert space of DMRG blocks and block entropy for any fixed accuracy or alternatively between the superblock Hilbert space and block entropy.

(2) We have presented a more rigorous truncation procedure based on Kholevo's theory on accessible information. This method also provides an alternative extrapolation method. We have found numerically that on a logarithmic scale the relative error is a linear function of the threshold value on the accessible information.

(3) We have also studied models when lattice site entropies are not equivalent and demonstrated the efficiency of our new truncation procedure and criteria of convergence.

(4) We have presented a sum rule that relates the sum of site entropies to the total information generated within a half sweep of the DMRG method. It has been shown that this could be used as an alternative test for convergence.

The application of other concepts of quantum information theory[36] to DMRG might also prove to be useful when more complicated systems are studied. A more rigorous study of entanglement distillation and purification protocols for systems when site entropies are not equivalent could lead to further optimization of the DMRG method. Another natural extension of the present work would be to use nonorthogonal states; however, the DMRG implementation of such a problem is rather complicated.[37]

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

[2] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1998).

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

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

[5] S. Nishimoto, E. Jeckelmann, F. Gebhard, and R. M. Noack, Phys. Rev. B **65**, 165114 (2002).

[6] Ö. Legeza and J. Sólyom, Phys. Rev. B **68**, 195116 (2003).

[7] S. Daul, I. Ciofini, C. Daul, and S. R. White, Int. J. Quantum Chem. **79**, 331 (2000).

[8] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[9] A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, J. Chem. Phys. **115**, 6815 (2001).

[10] Ö. Legeza, J. Röder, and B. A. Hess, Phys. Rev. B **67**, 125114 (2003).

[11] Ö. Legeza, J. Röder, and B. A. Hess, Mol. Phys. **101**, 2019 (2003).

[12] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).

[13] J. I. Latorre, E. Rico, and G. Vidal, Quantum Inf. Comput. **4**, 48 (2004).

[14] V. E. Korepin, Phys. Rev. Lett. **92**, 096402 (2004).

[15] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[16] F. Verstraete, D. Porras, and J. I. Cirac, cond-mat/0404706 (unpublished).

[17] S. R. White and A. Feiguin, cond-mat/0403310 (unpublished).

[18] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech.: Theory Exp. **2004**, 04005 (2004).

[19] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[20] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[21] R. Jozsa and B. Schumacher, J. Mod. Opt. **41**, 2343 (1994).

[22] R. Jozsa, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **81**, 1714 (1998).

[23] A. S. Kholevo, Probl. Inf. Transm. **177**, 9 (1973).

[24] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73**, 3047 (1994).

[25] E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[26] R. Jozsa, D. Robb, and W. K. Wootters, Phys. Rev. A **49**, 668 (1994).

[27] C. H. Bennett, D. P. DiVicenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[28] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[29] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).

[30] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **64**, 042304 (2001).

[31] Ö. Legeza and G. Fáth, Phys. Rev. B **53**, 14349 (1996).

[32] E. Jeckelmann and S. R. White, Phys. Rev. B **57**, 6376 (1998).

[33] J. Bonca, J. E. Gubernatis, M. Guerrero, E. Jeckelmann, and S. R. White, Phys. Rev. B **61**, 3251 (2000).

[34] T. Xiang, J. Lou, and Z. Su, Phys. Rev. B **64**, 104414 (2001).

[35] M. Martin-Delgado and M. Navascues, Eur. Phys. J. D **27**, 169 (2003).

[36] A. Galindo and M. A. Martin-Delgado, Rev. Mod. Phys. **74**, 347 (2002).

[37] A. O. Mitrushenkov, G. Fano, R. Linguerri, and P. Palmieri, cond-mat/0306284 (unpublished).