On the choice of the density matrix in the stochastic TMRG

T Enss and U Schollwöck

Sektion Physik, Ludwig-Maximilians-Universität München, Theresienstr. 37/III, D-80333 München, Germany

E-mail: Tilman.Enss@physik.uni–muenchen.de, scholl@theorie.physik.uni–muenchen.de

Abstract
In applications of the density matrix renormalization group to nonhermitean problems, the choice of the density matrix is not uniquely prescribed by the algorithm. We demonstrate that for the recently introduced stochastic transfer matrix DMRG (stochastic TMRG) the necessity to use open boundary conditions makes asymmetrical reduced density matrices, as used for renormalization in quantum TMRG, an inappropriate choice. An explicit construction of the largest left and right eigenvectors of the full transfer matrix allows us to show why symmetrical density matrices are the correct physical choice.

PACS numbers: 02.50.Ey, 64.60.Ht, 02.70.−c, 05.10.Cc

1. Introduction

Since its inception in 1992 by White [1], the density matrix renormalization group (DMRG) has emerged as one of the most powerful numerical methods in the study of low-dimensional strongly correlated fermionic, bosonic or quantum magnetic systems [2]. The universality of its core idea of deducing a decimation prescription for state spaces by considering the renormalization flow of suitably reduced density matrices has led many researchers to extend the method to other fields of research in systems with many correlated degrees of freedom.

Originally, the method was applied to the renormalization of effective low-energy Hamiltonians to study static and dynamic $T = 0$ properties. Major progress occurred with Nishino's realization [3] that the DMRG can be used to renormalize the transfer matrix of semi-infinite two-dimensional strips, a method which we will refer to as TMRG (transfer matrix renormalization group) in the following. The use of a transfer matrix implies that the system is truly infinite in one spatial direction, whereas the power of the DMRG to treat large one-dimensional systems is reflected in the fact that the strip width can be chosen so large as to allow reliable finite-size extrapolations. In an analogy to the approach used in Quantum Monte Carlo, one-dimensional quantum systems were then mapped by a Trotter–Suzuki checkerboard

decomposition to a two-dimensional classical system [4]. The renormalization of the resulting quantum transfer matrix allows the study of the thermodynamics of quantum chains for infinite system sizes and very fine Trotter–Suzuki decompositions, i.e. down to very low temperatures.

As the Trotter–Suzuki decomposition generates a transfer matrix which is not invariant under a spatial reflection of the system, the quantum transfer matrix is asymmetrical and there are left and right eigenstates $\langle\psi_L|$ and $|\psi_R\rangle$ associated with the thermodynamically relevant largest eigenvalue of the transfer matrix. This raises the question which is the correct way to construct the optimal density matrix for the DMRG. In the usual hermitean DMRG, the density matrix is typically obtained by tracing out a part of the universe ('environment') in the ground state:

$$\rho_{sys} = \text{Tr}_{env}\, |\psi\rangle\langle\psi|. \tag{1}$$

In the non-hermitean case, Bursill *et al* [4] made the *symmetrical* choice

$$\rho_{sys} = \text{Tr}_{env}\, |\psi_R\rangle\langle\psi_R| \tag{2}$$

which can be easily extended to

$$\rho_{sys}^{symm} = \text{Tr}_{env}(|\psi_R\rangle\langle\psi_R| + |\psi_L\rangle\langle\psi_L|) \tag{3}$$

which treats the left and right eigenvectors on an equal footing (all composite density matrices have to be appropriately normalized, depending on the normalization chosen for the eigenvectors; we do not show these factors explicitly in the paper). This approach also has the property of minimizing the sum of the two squared distances between the original eigenstates and their projections onto the reduced state spaces after the renormalization step, which is the prescription from which the DMRG procedure for symmetrical matrices can be derived.

However, it was pointed out by Wang and Xiang [5] that the choice of an *asymmetrical* density matrix,

$$\rho_{sys}^{symm} = \text{Tr}_{env}\, |\psi_R\rangle\langle\psi_L| \tag{4}$$

is more physically adequate and yields numerically more satisfying results. This is now the accepted choice for the application of the TMRG to quantum problems [6–8].

A further field of applications was opened up by the application of the DMRG idea to the renormalization of genuinely asymmetrical problems [9–11] as they occur in non-equilibrium statistical mechanics. Carlon *et al* [11] exploited the fact that the time evolution of reaction–diffusion systems can be mapped to a Schrödinger-like equation. In this case, one can study the long-time behaviour of finite size systems quite precisely [11–14].

Here, the question of the correct choice of the density matrix arises as the transition matrix is genuinely asymmetrical itself, since there is no detailed balance. Carlon *et al* found the choice of the symmetrical density matrix to be most suitable, but essentially due to reasons of numerical stability: in this class of problems, very high numerical precision for the chosen basis states in the reduced basis was found to be important, which is less easily obtained for asymmetrical matrices. An attempt to include projectors like $|\psi_R\rangle\langle\psi_L|$ while building a symmetrical density matrix by studying

$$\rho_{sys}^{mix} = \text{Tr}_{env}(|\psi_L\rangle + |\psi_R\rangle)(\langle\psi_L| + \langle\psi_R|) \tag{5}$$

yielded inferior results.

Recently, the TMRG has been used by Kemper *et al* [15] to successfully renormalize stochastic transfer matrices, an approach which henceforth we will refer to as 'stochastic TMRG'. In their study of the Domany–Kinzel cellular automaton, they found that the use of a symmetrical density matrix yielded satisfactory results, while the asymmetrical choice led

Figure 1. Trotter–Suzuki decomposition of the full time evolution into local interactions τ ($2M = 6$ time steps shown). The lattice is spatially infinitely extended, while in time direction the system is finite. Two adjacent columns (boxed) constitute the basic building block: the transfer matrix $(T)^{l_1...l_{2M-1}}_{r_1...r_{2M-1}}$.

to incorrect results, which is at variance with the conventional (or 'quantum') TMRG. These findings were obtained on an entirely empirical basis.

In this paper, we want to investigate in depth the question of the correct choice of the density matrix, which is at the core of all asymmetrical DMRG applications for the new stochastic TMRG. We will give explicit constructions for the eigenvalues and selected eigenvectors of the stochastic transfer matrices, highlighting the central role of the boundary conditions in time direction. This will allow us to demonstrate that the content of the asymmetrical density matrix is essentially trivial, while the symmetrical density matrix is the physically adequate choice.

2. Explicit construction of the unrenormalized TMRG transfer matrix

2.1. The TMRG transfer matrix

In the stochastic TMRG (for a description of the very similar TMRG applied to quantum systems, see [2]), one considers a stochastically evolving system extended infinitely in one spatial dimension with (for simplicity) local update rules involving neighbouring sites that only allow a finite number of states (n). The local interaction between two (or similarly, a few) lattice sites is given by the local transfer matrix

$$(\tau)^{l_2 r_2}_{l_1 r_1} = [\exp(\Delta t \cdot H_{\text{local}})]^{l_2 r_2}_{l_1 r_1} = \quad \tag{6}$$

where the stochastic 'Hamiltonian' H_{local} gives the transition rates between states and the arrow indicates the direction of the time step Δt. We assume spatial parity invariance

$$(\tau)^{l_2 r_2}_{l_1 r_1} = (\tau)^{r_2 l_2}_{r_1 l_1}. \tag{7}$$

In complete analogy to Quantum Monte Carlo and the conventional TMRG, the real time evolution operator for the full lattice is mapped approximately by a Trotter–Suzuki decomposition to a checkerboard structure of local interactions (figure 1).

In the time direction the evolution of the system is simulated during a finite interval. As illustrated in figure 1, the full evolution operator is then written as $\exp(2M\Delta t \cdot H) = \lim_{N\to\infty} T^N$, where M is the Trotter number, N the (diverging) system size and $H = \lim_{N\to\infty} \sum_{i=-N}^{N} H_{\text{local}}(i, i+1)$. T is called the transfer matrix which consists of two adjacent columns of τ's, where we take as upper (lower) index the 'history' (time evolution) of the lattice site at the left (right) side of the zig-zagged column,

$$(T)^{l_1...l_{2M-1}}_{r_1...r_{2M-1}} = \underbrace{\sum_{m_0 r_0} w(m_0)\, w(r_0)}_{\text{open initial b.c.}} \sum_{m_1...m_{2M-1}} \underbrace{\sum_{l_{2M} m_{2M}}}_{\text{open final b.c.}} \prod_{k=0}^{M-1} (\tau)^{m_{2k+1} r_{2k+1}}_{m_{2k} r_{2k}} (\tau)^{l_{2k+2} m_{2k+2}}_{l_{2k+1} m_{2k+1}} \tag{8}$$

where l, m, and r denote the left, middle and right lattice sites of T, respectively. One uses stochastic initial conditions: the weight of each possible initial configuration is given by the product of local weights $w(s_i)$ of the states s_i at each lattice site i,

$$w(\ldots, s_1, s_2, s_3, \ldots) = \prod_i w(s_i) \qquad \sum_{s_i=1}^{n} w(s_i) = 1 \quad \text{(norm.)} \tag{9}$$

such that if each state is weighted equally, there is no bias in the initial configuration. At the end, one uses open final boundary conditions, i.e. all final states are allowed without bias. In the diagrams, we thus implicitly trace over all internal indices, sum over all final indices and perform a weighted sum over all initial indices.

2.2. Spectrum and eigenvectors of the transfer matrix

As the local stochastic transfer matrix τ must obey probability conservation, the probability of ending up in *any* state is unity:

$$\sum_{l_2 r_2} (\tau)^{l_2 r_2}_{l_1 r_1} = 1 \quad \forall l_1, r_1 \qquad \qquad \qquad \qquad \text{(trivial).} \tag{10}$$

This implies

$$\text{Tr}\, T = 1 \tag{11}$$

where tracing contracts the upper index of T with the lower one; thus the probability that the lattice sites at the left and right boundary of T have the same time evolution is unity. In pictorial language, the trace rolls up the transfer matrix into a cylinder, with its left and right boundary identified and the middle site at the opposite side of the cylinder (cf. figure 2).

But this is equivalent to having only the local interaction between the left/right lattice site on one side and the middle one on the other, for $2M$ time steps Δt, which is known exactly (e.g. $2M = 6$):

$$(T)^{l_1...l_5}_{r_1...r_5} = \sum_{m_0 r_0} w(m_0)\, w(r_0) \sum_{l_6 m_6} \sum_{m_1...m_5} (\tau)^{l_6 m_6}_{l_5 m_5} (\tau)^{m_5 r_5}_{m_4 r_4} \cdots (\tau)^{l_2 m_2}_{l_1 m_1} (\tau)^{m_1 r_1}_{m_0 r_0} \tag{12}$$

$$\begin{aligned}
\text{Tr}\, T &= \sum_{l_1...l_5} (T)^{l_1...l_5}_{l_1...l_5} \\
&= \sum_{m_0 r_0} w(m_0)\, w(r_0) \sum_{l_6 m_6} \sum_{l_1...l_5} \sum_{m_1...m_5} (\tau)^{l_6 m_6}_{l_5 m_5} (\tau)^{m_5 l_5}_{m_4 l_4} \cdots (\tau)^{m_1 l_1}_{m_0 l_0} \\
&= \sum_{m_0 r_0} w(m_0)\, w(r_0) \sum_{l_6 m_6} [\tau^{2M}(\Delta t)]^{l_6 m_6}_{r_0 m_0} \qquad \text{(by parity invar. (7))}
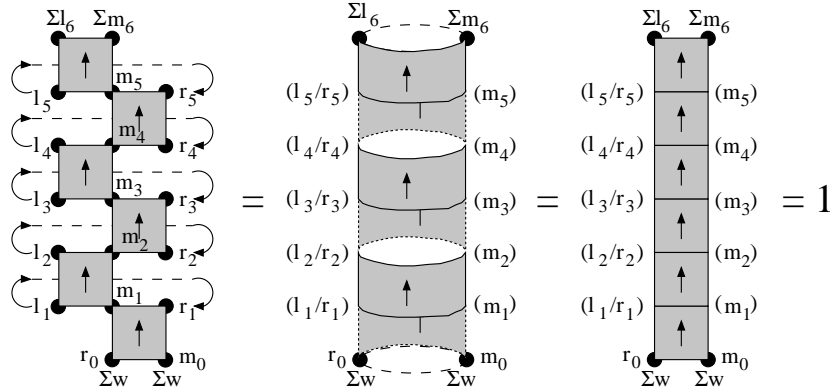\end{aligned}$$

Figure 2. $\text{Tr}(T)^{l_1 \ldots l_{2M-1}}_{r_1 \ldots r_{2M-1}}$ is equivalent to T being rolled up on a cylinder with a solely local interaction.
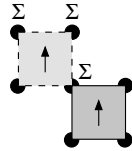

Figure 3. Open b.c. lead to trivial τ's that cause also the final indices of the previous τ's to be summed over, thus 'propagating' the open b.c. backwards in time.

$$
\begin{aligned}
&= \sum_{m_0 r_0} w(m_0)\, w(r_0) \sum_{l_6 m_6} [\exp(2M\Delta t \cdot H_{\text{local}})]^{l_6 m_6}_{r_0 m_0} \\
&= \sum_{m_0 r_0} w(m_0)\, w(r_0) \sum_{l_6 m_6} [\tau(2M\Delta t)]^{l_6 m_6}_{r_0 m_0} \\
&= \sum_{m_0 r_0} w(m_0)\, w(r_0) \cdot 1 = 1 \qquad\qquad \text{(by (10) and (9))}.
\end{aligned}
$$

Next we consider the spectrum and eigenvector decomposition of T. As all entries of T are non-negative, the Frobenius theorem implies that the largest eigenvalue λ_0 is non-degenerate and real. The associated left and right eigenvectors $\langle\psi_{\text{L}}|$ and $|\psi_{\text{R}}\rangle$ (where $|\psi_{\text{R}}\rangle^\dagger \neq \langle\psi_{\text{L}}|$ in general, since T is not symmetrical) have nonnegative entries. Now the decisive difference between conventional and stochastic TMRG is that in the conventional TMRG one has periodic boundary conditions in (imaginary) time, but open final boundary conditions in the stochastic TMRG. This latter property implies

$$
T^{2M-1} = |\psi_{\text{R}}\rangle\langle\psi_{\text{L}}| = T^k \qquad\qquad \forall k \geqslant 2M-1 \tag{13}
$$

i.e. sufficiently large powers of T are identical and decompose into an outer product of the two eigenvectors of $\lambda_0 = 1$, normalized as $\langle\psi_{\text{L}}|\psi_{\text{R}}\rangle = 1$, while all other eigenvalues of T vanish.

To see this, consider figure 3. Because τ is stochastic, it becomes trivial by summing over its final indices (cf. (10); its value does not depend on its initial indices, such that the contraction with the final index of a *previous* τ yields an unweighted sum over the previous τ's final index. Therefore in figure 4, the trivial τ's 'propagate' the open final b.c. backwards in time, resulting in a cone of trivial τ's that has no influence on the value of the transfer matrix T^k.

The Trotter–Suzuki decomposition introduces a finite speed of action ('speed of light') into the model: during each time step, only nearest neighbours that are connected by an edge
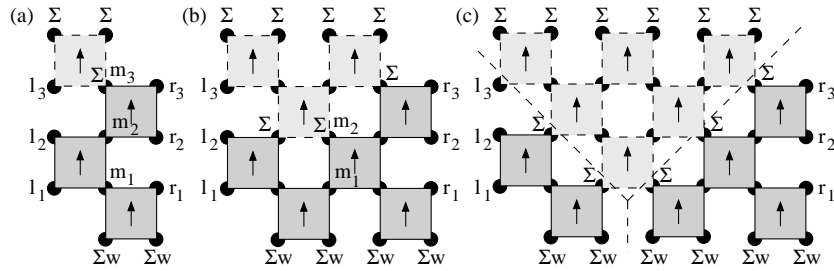
Figure 4. T^k decompose into a left and right part of ever smaller rank until T^{2M-1} decomposes into two lightcones which cannot interact during $2M$ time steps (here $2M = 4$): (a) $T = (\psi_R)_{m_1 m_2}^{l_1 l_2 l_3} \cdot (\psi_L)_{r_1 r_2 r_3}^{m_1 m_2}$ with two connecting indices (m_1, m_2), (b) $T^2 = (\psi_R)_{m_1}^{l_1 l_2 l_3} \cdot (\psi_L)_{r_1 r_2 r_3}^{m_1}$ with one connecting index (m_1), (c) $T^{2M-1} = (\psi_R)^{l_1 l_2 l_3} \cdot (\psi_L)_{r_1 r_2 r_3} = |\psi_R\rangle\langle\psi_L|$ with no connecting index.

of a (*non-trivial*) τ can interact, so interactions can only propagate at a speed of one lattice site per time step Δt. Hence we speak of the 'light cone' of a certain site as those sites in time and space that influence the probabilities of different occupations for this site.

In T (cf. figure 4(a)), the last τ is trivial, hence it has the value 1 indepedent of l_3 and m_3. Thus we effectively sum over the internal index m_3, and l_3 is a 'dummy index': the value of $(T)_{r_1 r_2 r_3}^{l_1 l_2 l_3}$ does not depend on it. This leaves only two indices (m_1, m_2) connecting the left and right part of T, so the transfer matrix can be decomposed into a product of two matrices ψ_R and ψ_L of rank n^2 (n is the number of possible states per site).

When we multiply this with another T (cf. figure 4(b)), we observe that a third τ becomes trivial as the final τ's propagate the final open b.c. backwards. This leads to a sum also over m_2, leaving m_1 as the only connecting index, so T^2 decomposes into a product of two matrices of rank n.

Finally in T^{2M-1} (cf. figure 4(c)), the trivial τ's propagate all the way to the initial τ's, effectively separating the left and right boundary: now there are no connecting indices left between ψ_R and ψ_L, turning them into column and row vectors $|\psi_R\rangle$ and $\langle\psi_L|$. From the viewpoint of causality, the left and right boundary is separated by $4M - 2$ lattice sites, so no initial lattice site can influence *both* the left and the right boundary, i.e. the light cone for the lattice sites on the left boundary $(l_1, \ldots, l_{2M-1})$ does not intersect the light cone for the lattice sites on the right boundary $(r_1, \ldots, r_{2M-1})$. This means that the probability of a particular history on the left boundary is completely independent of the probability of a particular history on the right.

We can now explicitly compute these column and row vectors $|\psi_R\rangle^{l_1 \ldots l_{2M-1}}$ and $\langle\psi_L|_{r_1 \ldots r_{2M-1}}$: we take the *left* light cone (of triangular shape, delimited by the dashed line) and contract the included non-trivial local transfer matrices τ in all internal, initial, and final indices, keeping the *free* indices $l_1 \ldots l_{2M-1}$, and denote this object as (*right* eigenvector, see below) $|\psi_R\rangle^{l_1 \ldots l_{2M-1}}$. Analogously, we denote the same construction on the right with free indices $r_1 \ldots r_{2M-1}$ as $\langle\psi_L|_{r_1 \ldots r_{2M-1}}$. Note that the last index on the left, l_{2M-1}, is by construction a dummy index which does not change the value of $|\psi_R\rangle^{l_1 \ldots l_{2M-1}}$.

If we consider even higher powers of T than $2M-1$, we only add more trivial τ's in the middle that lie outside either light cone and therefore do not change the value of the transfer matrix. Thus we have proved $T^{2M-1} = T^k \ \forall k \geqslant 2M - 1$, which implies for the eigenvalues λ_i of T that $\lambda_i^{2M-1} = \lambda_i^{2M} \ \forall i$, such that λ_i can only assume the values 0 or 1. From ([11]) it follows that $\lambda_0 = 1$, $\lambda_i = 0 \ \forall i > 0$.

Figure 5. $|\psi_R\rangle$ is indeed the right eigenvector of T corresponding to the eigenvalue $\lambda_0 = 1$: for $2M = 4$, $T^{l_1 l_2 l_3}_{m_1 m_2 m_3} |\psi_R\rangle^{m_1 m_2 m_3} = |\psi_R\rangle^{l_1 l_2 l_3}$.

The above decomposition of T^{2M-1} into an outer product of two vectors $|\psi_R\rangle\langle\psi_L|$ is in fact the explicit eigenvector decomposition corresponding to the eigenvalue $\lambda_0 = 1$: multiplying the right (left) eigenvector with T^{2M-1} from the right (left) is, in pictorial language, attaching the respective light cone to the transfer matrix with a trace over all overlapping indices. But now the final open b.c. can propagate through all τ's on that side, making them trivial (which automatically gives the correct normalization $\langle\psi_L|\psi_R\rangle = 1$) and leaving just the light cone on the other side, which reproduces exactly the previously attached eigenvector. The eigenvalue 1 is determined by this normalization condition:

$$T^{2M-1}|\psi_R\rangle = |\psi_R\rangle\langle\psi_L|\psi_R\rangle = |\psi_R\rangle \cdot 1. \tag{14}$$

This concludes the proof of (13). Similarly, we check that $|\psi_R\rangle$ and $\langle\psi_L|$ are also the eigenvectors of T corresponding to the same eigenvalue (cf. figure 5).

Let us finally emphasize that for small powers of T, T^k with $k < 2M - 1$, this simple decomposition does not hold. But this is not relevant, as we take k to infinity in the TMRG.

To arrive at a prescription for the stochastic TMRG density matrix, we must now consider how actual physical information is extracted from the stochastic transfer matrix. This is done by computing expectation values of local operators, averaged over the whole lattice at final time $2M \cdot \Delta t$ by multiplying τ in the last time step with the local operator, e.g. for the local particle number operator $n(i)$ one has

$$(\tau_n)^{l_2 r_2}_{l_1 r_1} := (\tau)^{l_2 r_2}_{l_1 r_1} \cdot n(l_2) \tag{15}$$

where we denote the transfer matrix with its last τ replaced by τ_n as T_n (coloured black in the diagram). Then the expectation value of the particle number operator can be calculated as

$$\langle n_{t=2M \cdot \Delta t}\rangle = \lim_{N\to\infty} \frac{\text{Tr}(T^N T_n)}{\text{Tr}(T^{N+1})} = \lim_{N\to\infty} \frac{\text{Tr}\left(\text{[diagram]} \right)}{\text{Tr}\left(\text{[diagram]} \right)} \tag{16}$$

$$= \frac{\text{Tr}(|\psi_R\rangle\langle\psi_L|T_n)}{\text{Tr}(|\psi_R\rangle\langle\psi_L|)} = \frac{\langle\psi_L|T_n|\psi_R\rangle}{\underbrace{\langle\psi_L|\psi_R\rangle}_{1}} = \left(\text{[diagram]} \right) \tag{17}$$
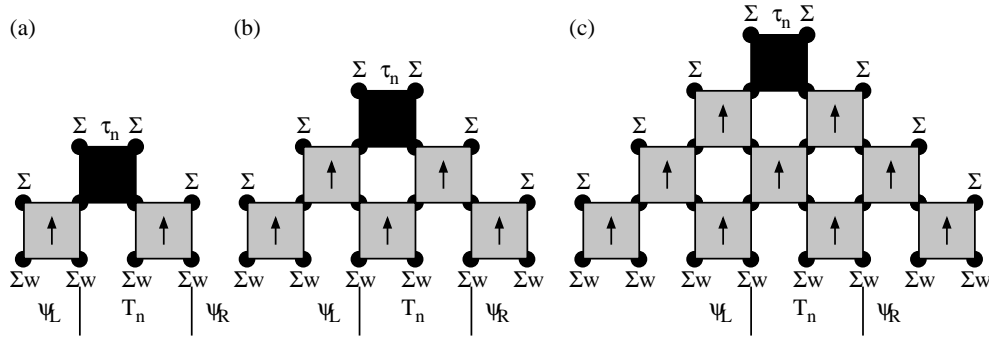
Figure 6. The light cones that are needed to compute the expectation value of the particle number operator for the first few time steps: (a) $2M = 2$, (b) $2M = 3$, (c) $2M = 4$.

where we have omitted local transfer matrices that become trivial by summing and used the pictorial construction of the eigenvectors. The computation of expectation values is thus reduced to the computation of the light cone of lattice sites that can possibly influence the particular site at which the operator is measured.

Figure 6 shows the light cones for the first few time steps. In this manner one can compute the exact evolution of $\langle n(t) \rangle$, the only inaccuracy stemming from the Trotter–Suzuki decomposition. It is interesting to observe that for determining both eigenvectors and expectation values, a matrix diagonalization can be completely avoided.

For the choice of the density matrix the important observation in figure 6 now is that by parity invariance (7) and adding one dummy index (i_{2M-1}),

$$|\psi_R\rangle_{2M}^{i_1 \ldots i_{2M-1}} = (\langle \psi_L |_{2M-1, \, i_1 \ldots i_{(2M-1)-1}})^T =: |\psi_L\rangle_{2M-1}^{i_1 \ldots i_{2M-2}} \tag{18}$$

i.e. *the left eigenvector for a given time step is identical to the right eigenvector in the next (half) time step.*

3. Density matrix renormalization of the stochastic transfer matrix

We have seen how, in principle, physical information can be extracted from the stochastic transfer matrix, but both memory usage and computation time grow exponentially with the number of time steps. The transfer matrix operates on the space of possible time evolutions of particular sites, which has to be decimated to remain numerically manageable. This is the idea of the TMRG. As the probability of a partial time evolution is determined not only by its past, but also by its future continuation, and the infinite future is not yet known, one contends oneself with a 'lookahead' of several time steps which is called the 'environment', while the past time evolution—that is being projected onto a smaller Hilbert space—is called the 'system'.

Traditionally, one obtains this basis by computing the reduced density matrix for the system, ρ_{sys}, by taking a partial trace over the environment of the outer product of right and left T eigenvectors; the new Hilbert space basis consists of ρ_{sys}'s eigenvectors associated with the largest eigenvalues (highest probabilities). Inspired by traditional TMRG (with periodic b.c. in time direction), one could try $\rho_{\text{sys}}^{\text{asym}} = \text{Tr}_{\text{env}} |\psi_R\rangle\langle\psi_L|$. But in the case of open b.c., this is just a partial trace of the transfer matrix,

$$\rho_{\text{sys}}^{\text{asym}} = \text{Tr}_{\text{env}} |\psi_R\rangle\langle\psi_L| = \text{Tr}_{\text{env}} T^{2M-1} = |\psi_R\rangle_{\text{sys}}\langle\psi_L|_{\text{sys}} \tag{19}$$

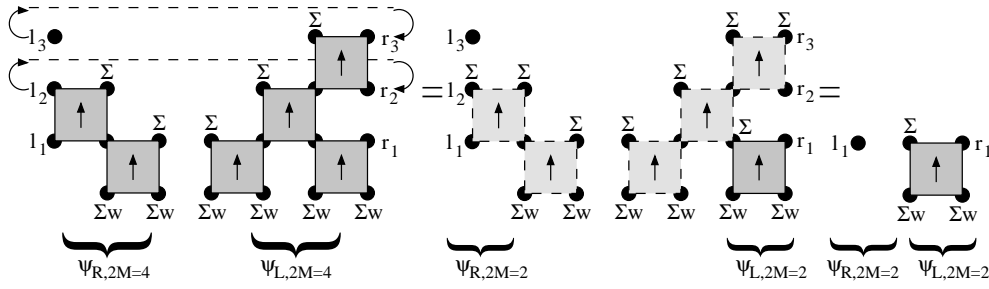**Figure 7.** Computing the asymmetrical density matrix $\rho_{\mathrm{sys}}^{\mathrm{asym}}$: the trace makes the environment trivial, $\mathrm{Tr}_{\mathrm{env}}\,|\psi_{\mathrm{R}}\rangle\langle\psi_{\mathrm{L}}| = |\psi_{\mathrm{R}}\rangle_{\mathrm{sys}}\langle\psi_{\mathrm{L}}|_{\mathrm{sys}}$.

having a spectrum of

$$\mathrm{Spec}\left(\rho_{\mathrm{sys}}^{\mathrm{asym}}\right) = \{1, 0, \ldots, 0\}. \tag{20}$$

We demonstrate this for $2M = 4$, taking e.g. the first index as the system index and the following two indices as environment indices. Then the trace over the environment $\mathrm{Tr}_{\mathrm{env}} = \sum_{l_3 r_3 l_2 r_2} \delta_{l_3}^{r_3} \delta_{l_2}^{r_2}$, as illustrated in figure 7, makes all τ's in the environment trivial: when we contract the dummy index l_3 with r_3, the final τ of $\langle\psi_{\mathrm{L}}|$ becomes trivial, so r_2 now becomes a dummy index. This in turn gives a sum over l_2, also making the final τ of $|\psi_{\mathrm{R}}\rangle$ trivial, etc. In the end we are left with $|\psi_{\mathrm{R}}\rangle$ and $\langle\psi_{\mathrm{L}}|$ of an earlier time step, namely the one with only the system indices l_1, r_1.

Thus this asymmetrical density matrix provides only a single basis vector for the Hilbert space projection, which is obviously not enough. The reason is that the open b.c. 'average away' the physical interaction in the environment by making irrelevant how the system will react in the future. In these boundary conditions resides the essential difference between quantum and stochastic TMRG.

Each renormalization step using $\rho_{\mathrm{sys}}^{\mathrm{asym}}$ therefore projects onto a one-dimensional Hilbert space. No longer-ranged correlation implicit in differently weighted basis vectors is being built up. We therefore expect to observe only the local evolution of the model, which is exactly what we get from numerical simulation using the asymmetrical density matrix.

As an example, we use an exactly solvable reaction–diffusion model: $AA \to 00$ with rate 2α (reaction) and $A0 \leftrightarrow 0A$ with rate D (diffusion). For $2\alpha = D$, the exact solution is given by Spouge [16]: for initial pratical number $n(t = 0) = \frac{1}{2}$ (unbiased $w(0) = w(A) = \frac{1}{2}$, i.e. all lattice sites have an independent probability $\frac{1}{2}$ of occupancy),

$$\langle n(t)\rangle = \tfrac{1}{2}\exp(-2Dt)\left[I_0(2Dt) + I_1(2Dt)\right] \sim \frac{1}{\sqrt{t}} \quad \text{(asympt.).} \tag{21}$$

This agrees well only with the simulation using the symmetrized density matrix $\rho_{\mathrm{sys}}^{\mathrm{symm}}$ (cf. figure 8). But instead, when we use the asymmetrical density matrix $\rho_{\mathrm{sys}}^{\mathrm{asym}}$, we observe a $1/t$ decay. This resembles very closely what we expect from a local interaction: the exact 4-site interaction (i.e. renormalizing as soon as the system has n states) is

$$\langle n(2M + 1)\rangle = \langle n(2M)\rangle - [1 - \exp(-D\Delta t)] \cdot \langle n(2M)\rangle^2 \tag{22}$$

($n(0) = \frac{1}{2}$) which goes asymptotically as $1/t$.

The origin of this shortcoming becomes clear in the following consideration: figure 9 shows the left eigenvectors during consecutive time steps. Suppose we want to find a good basis for renormalizing $\langle\psi_{\mathrm{L}}|_{2M=5}$, taking its first two indices (l_1, l_2) as system and the latter
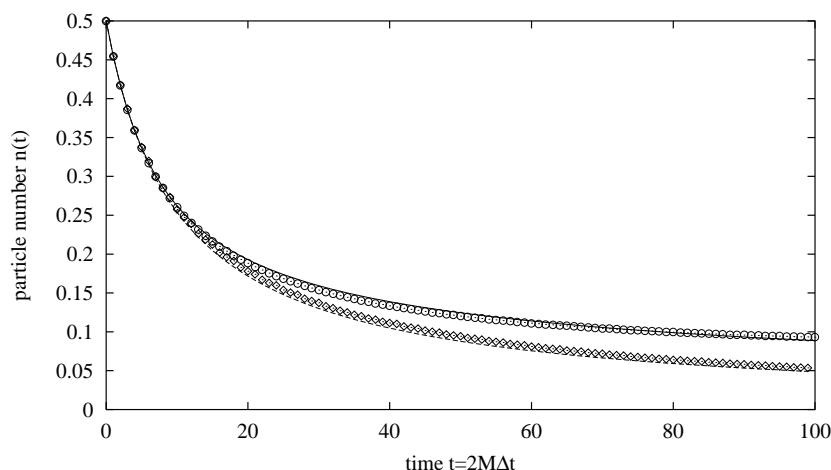
Figure 8. Comparing the simulation of a reaction–diffusion model for symmetrized vs. asymmetrical density matrices. (——) exact solution, (- - - -) exact local (4-site) interaction, (◯) simulation, symmetrized density matrix, (◇) simulation, asymmetrical density matrix ($D = 0.1$ in inverse time units, and $\Delta t = 1$).
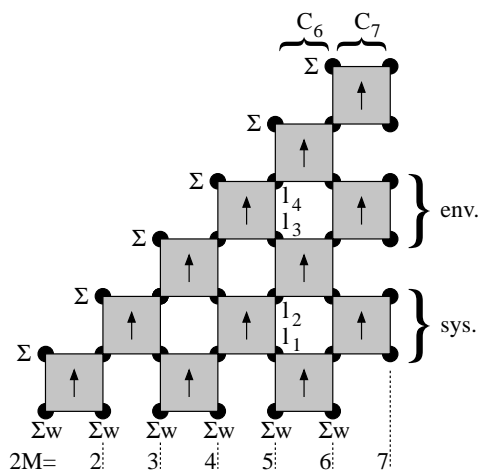

Figure 9. The evolution of the largest left T eigenvector $\langle\psi_L|_{2M}$ for the first few Trotter steps.

two (l_3, l_4) as environment. After somehow tracing over the environment, the new basis vectors will have indices l_1, l_2, and will roughly resemble $\langle\psi_L|_3$, having a strong correlation between l_1 and l_2 because they are connected by an edge of τ.

Now in order to advance one time step from $\langle\psi_L|_5$ to $\langle\psi_L|_6$, we see from figure 9 that we need to multiply $\langle\psi_L|_5$ by the matrix represented by the column C_6. This reveals the problem: the first two indices of C_6 are not connected by an edge of τ and are therefore completely independent. Thus the renormalization basis obtained from $\langle\psi_L|_5$ with a strong correlation between the first pair of indices fits very poorly for renormalizing C_6. The states in the basis are essentially orthogonal to those that would describe a renormalized C_6 well. Looking ahead, C_7 which would bring us to $\langle\psi_L|_7$ has again a structure similar to $\langle\psi_L|_5$ and could be renormalized well.

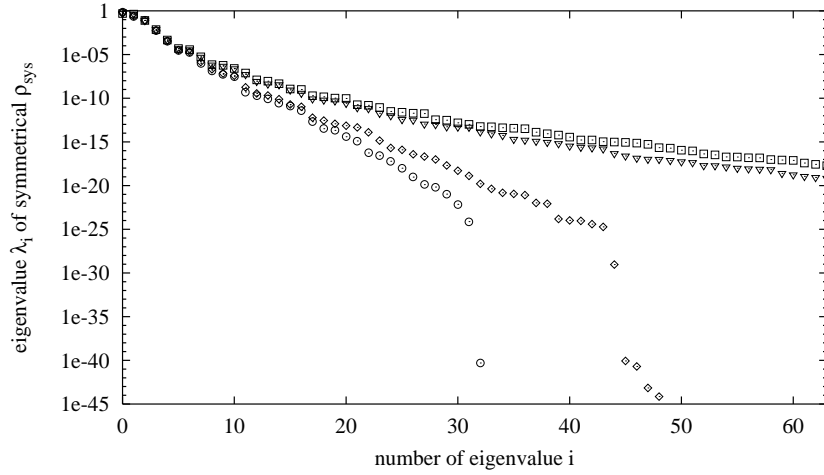**Figure 10.** Spectrum of the reduced system density matrix $\rho_{\mathrm{sys}}^{\mathrm{symm}}$ at specific steps $2M$ ($m = 64$). ($\bigcirc$) $2M = 18$ (trunc. 10^{-40}), ($\Diamond$) $2M = 20$ (trunc. 10^{-29}), (∇) $2M = 60$ (trunc. 10^{-21}), ($\Box$) $2M = 280$ (trunc. 10^{-21}).

This problem could be remedied by adding to the renormalization basis also basis vectors of a form similar to $\langle\psi_{\mathrm{L}}|_4$ (tracing over the last index as environment) which shows little correlation between the first two indices. Noting that $\langle\psi_{\mathrm{L}}|_4 = (|\psi_{\mathrm{R}}\rangle_5)^T$ (18) we have to mix left and right eigenvectors, $\langle\psi_{\mathrm{L}}|_5$ and $\langle\psi_{\mathrm{R}}|_5$, in order to obtain a good basis. Based on the original argument by White [1] about the optimal density matrix in DMRG, Carlon *et al* [11] have shown that in order to express *two* different vectors $|\psi_{\mathrm{L}}\rangle$, $|\psi_{\mathrm{R}}\rangle$ most accurately (in the *same* basis), one has to use an eigenbasis of $\rho_{\mathrm{sys}} = \mathrm{Tr}_{\mathrm{env}}(|\psi_{\mathrm{R}}\rangle\langle\psi_{\mathrm{R}}| + |\psi_{\mathrm{L}}\rangle\langle\psi_{\mathrm{L}}|)$. The symmetrical form of this density matrix arises because they used as a measure of distance between exact and approximated vectors the symmetrical norm $\||\psi_{\mathrm{R}}\rangle - |\tilde{\psi}_{\mathrm{R}}\rangle\|^2 + \||\psi_{\mathrm{L}}\rangle - |\tilde{\psi}_{\mathrm{L}}\rangle\|^2$. With this same argument, we now conclude that *this symmetrized density matrix is also optimal in the case of explicit eigenvectors*.

This explains why a symmetrized density matrix, e.g. $\rho_{\mathrm{sys}}^{\mathrm{symm}} = \mathrm{Tr}_{\mathrm{env}}(|\psi_{\mathrm{R}}\rangle\langle\psi_{\mathrm{R}}| + |\psi_{\mathrm{L}}\rangle\langle\psi_{\mathrm{L}}|)$ as used by Kemper *et al* [15], works so well. Indeed, within the precision of the method, the exact result for the example considered is reproduced when we use this symmetrized reduced density matrix $\rho_{\mathrm{sys}}^{\mathrm{symm}}$.

If we consider for our example the eigenvalue spectrum of the symmetrized density matrix, initially it falls off approximately exponentially, $\lambda_i \sim \alpha^{-i}$ for some constant α (cf. figure 10). After many renormalization steps, the spectrum falls off rather according to a power law with large negative exponent, $\lambda_i \sim i^{-\beta}$. The symmetrized density matrix therefore indeed fulfills two basic requirements: It should have as many non-zero eigenvalues as possible in order to keep rich non-trivial information about the future evolution in the environment (contrary to (20)), but at the same time they should drop off as quickly as possible so that when we truncate the spectrum at the m states that are retained, we lose as little information on the system as possible.

Now, in practical applications of the quantum TMRG, system and environment are treated on an equal basis, renormalizing the system with a density matrix traced over the environment and vice versa. In the stochastic TMRG, the explicit construction of T and the eigenvectors opens the possibility of renormalizing only the system and keeping the environment as a 'lookahead' of constant finite length. This simplifies the algorithm, since the ψ's can be
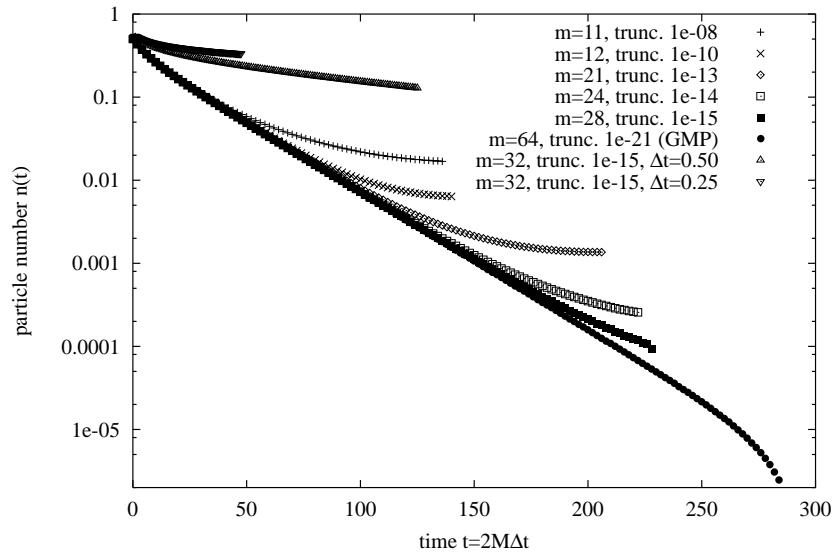
Figure 11. Branching–fusing model [11] at the critical point $p = 0.8403578$ for different numbers of states kept (m) and associated truncations in the spectrum of ρ_{sys}.

renormalized explicitly as shown above, and there is no need to ever solve the eigenvalue problem of the transfer matrix; on the other hand, due to the exponential growth of the matrices with a longer lookahead, this approach is in practice limited (see our numerical example below), and one reverts to the conventional approach of renormalizing both system and environment of the transfer matrix T and then finding its largest eigenvectors anew in each step by using a suitable algorithm for the diagonalization of large non-symmetrical matrices.

Our considerations above were all for the case without the renormalized environment. Renormalizing it, too, leads to an increasingly longer, but less precise lookahead. As our considerations hold formally for all lengths of lookaheads, the result on the choice of the density matrix should also hold in that case, which is indeed what we observe numerically. To check whether our special choice of a symmetrical density matrix is indeed optimal among symmetrical density matrices, we have also considered numerically several other symmetrical density matrices for comparison, in particular

$$\text{Tr}_{env}(|\psi_L\rangle\langle\psi_R| + |\psi_R\rangle\langle\psi_L|) \tag{23}$$

$$\text{Tr}_{env}(|\psi_L\rangle + |\psi_R\rangle)(\langle\psi_L| + \langle\psi_R|) \tag{24}$$

$$\text{Tr}_{env}(|\psi_L\rangle - |\psi_R\rangle)(\langle\psi_L| - \langle\psi_R|) \tag{25}$$

but the convergence turned out to be worse than for ρ_{sys}^{symm}. All results shown in the following are therefore for the 'canonical' choice of the symmetrical density matrix.

As an application, we tried the branching–fusing process explained in [11]: diffusion $A0 \leftrightarrow 0A$ (rate D), and several reactions $AA \rightarrow 00$ (rate 2α), $AA \rightarrow 0A, A0$ (rate γ), $0A, A0 \rightarrow 00$ (rate δ), and $0A, A0 \rightarrow AA$ (rate β). The relations between the rates are fixed as $D = 2\alpha = \gamma = \delta = 1 - p, \beta = p$. There is a critical point of the directed percolation universality class at $p_c \approx 0.8403578$. First we check the dependence of our TMRG simulation (using the conventional algorithm) for different values of m (number of states kept during projection into a smaller Hilbert space) (cf. figure 11) at the critical point, the most difficult point for the stochastic TMRG.

We observe that the particle number generally decays exponentially, despite criticality, because we operate at very short times due to the rather rough Trotter decomposition. On the lower branch, for a fixed Trotter time ($\Delta t = 1$), our results become more accurate as we increase m, effectively truncating the spectrum of ρ_{sys} at a smaller eigenvalue, until we reach the machine precision (53 mantissa bits). To compute more time steps, one option is to use more accurate floating point numbers (see below).

For comparison, choosing a smaller time step Δt (i.e. a finer Trotter decomposition, last two curves in figure 11), while maintaining a large Hilbert space ($m = 32$), we get very different results: the particle number decreases much more slowly, which shows that our Trotter decomposition is still far from being sufficiently fine (of course, extrapolations $\Delta t \to 0$ are feasible). Unfortunately the number of time steps we can simulate also decreases, which makes such an extrapolation more difficult. This can be explained as follows: the shorter the time step Δt, the smaller the transition probabilities to different states, and the closer is τ to the identity matrix. But a transfer matrix composed of such τ's has eigenvectors with an exponentially growing Euclidean norm ($\langle \psi_R \psi_R \rangle \sim n^M$ and analogously for $|\psi_L\rangle$). This is not changed by renormalization because an orthogonal transformation conserves the norm and we only lose those Hilbert space directions that are very improbable. As renormalization keeps the number of components of $\langle \psi_L|$ and $|\psi_R\rangle$ bounded, the values of the components have to grow exponentially to obtain a growing norm. At the same time, $\langle \psi_L|\psi_R\rangle = 1$, so $\langle \psi_L|$ and $|\psi_R\rangle$ are almost orthogonal. Now the numerical problem is the following: while the components grow exponentially, they have to cancel each other very accurately in order to yield the correct normalization. This works only until the components are of the order of magnitude $\sqrt{1/\epsilon}$, where ϵ is machine accuracy. Choosing more accurate floating point numbers, by taking twice as many mantissa bits we can simulate very roughly twice as many time steps. This is illustrated in figure 11: if we use the GNU multiprecision library (GMP), truncating at 10^{-21} instead of 10^{-15}, we can simulate approximately 1.5 times as many steps.

Finally we want to compare the numerical stability of the conventional TMRG algorithm, renormalizing both system and environment, with the newly proposed method of explicitly constructing the eigenvectors (saving the iterations of the transfer matrix diagonalization which take up more than half of the time in the original algorithm). Here we renormalize only the system, keeping a lookahead of finite length as environment. As can be seen in figure 12, the result is much better than the conventional algorithm for an intermediate accuracy of $m = 11$ states kept, while it cannot compete with the most accurate (up to machine precision) values for $m = 28$. So, as already mentioned above, it is more proof of a concept and a tool for analytical derivations than for practical simulations. Neither more sites lookahead nor larger m improve the result significantly. The computation time for the displayed curves was in the range of a few seconds ($m = 11$) up to a few tens of seconds ($m = 32$) on a PIII (500 MHz), while memory usage was at most a few MBytes.

4. Conclusions and outlook

We have shown a new way of looking at stochastic TMRG with open b.c. by explicitly giving the transfer matrix and its left and right eigenvectors corresponding to the physically relevant eigenvalue 1. This allowed us to study the physical effect of various choices for the density matrix of the TMRG. The conclusion was that, unlike the TMRG for quantum transfer matrices, an unsymmetrical density matrix is not just less accurate, but fundamentally unsuitable for the stochastic TMRG. The core problem was that one has open boundary conditions in the future of the time evolution, averaging away the physical interaction, which is not the case in the periodically closed quantum transfer matrix DMRG. The symmetrical density matrix
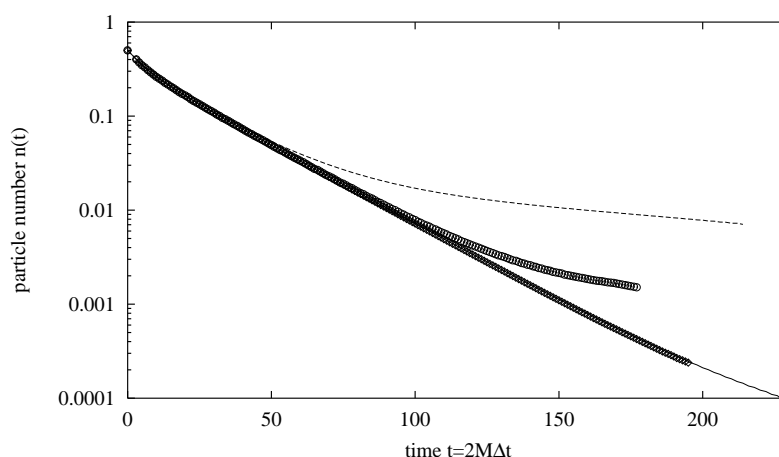
Figure 12. Branching–fusing model [11] at the critical point $p = 0.8403578$ for different TMRG methods. (——) conventional method, $m = 28$; (- - - -) conventional method, $m = 11$; (◯) explicit eigenvectors, $m = 8$ (four sites lookahead); (◇) explicit eigenvectors, $m = 11$ (six sites lookahead).

defined by demanding an optimal representation of both left and right eigenvectors in the truncated basis turned out to be the mandated choice. An important observation was that in certain problem classes where small transition probabilities lead to local transfer matrices close to unity, there is an inherent problem with the finite precision of computer arithmetic if one wants to have fine Trotter decompositions and to access comparatively large real times simultaneously.

Acknowledgments

References

[1] White S R 1992 *Phys. Rev. Lett.* **69** 2863
    White S R 1993 *Phys. Rev* B **48** 10345
[2] Peschel I, Wang X, Kaulke M and Hallberg K (eds.) 1999 *Density Matrix Renormalization* (Heidelberg: Springer)
[3] Nishino T 1995 *J. Phys. Soc. Japan* **64** 3598
[4] Bursill R J, Xiang T and Gehring G A 1996 *J. Phys.: Condens. Matter* **8** L583
[5] Wang X and Xiang T 1997 *Phys. Rev.* B **56** 5061
[6] Shibata N 1997 *J. Phys. Soc. Japan* **66** 2221
[7] Maisinger K and Schollwöck U 1998 *J. Phys. Rev. Lett.* **81** 445
[8] Nishino T and Shibata N 1999 *J. Phys. Soc. Japan* **68** 1537
[9] Hieida Y 1998 *J. Phys. Soc. Japan* **67** 369
[10] Kaulke M and Peschel I 1998 *Eur. J. Phys.* B **5** 727
[11] Carlon E, Henkel M and Schollwöck U 1999 *Eur. J. Phys.* B **12** 99
[12] Carlon E, Henkel M and Schollwöck U 2001 *Phys. Rev.* E **63** 036101
[13] Carlon E, Drzewinski A and van Leeuwen J M J 2000 *Preprint* cond-mat/0010177
[14] Henkel M and Schollwöck U 2001 *J. Phys. A: Math. Gen.* **34** 3333
[15] Kemper A, Schadschneider A and Zittartz J 2001 *J. Phys. A: Math. Gen.* **34** L279
[16] Spouge J L 1988 *Phys. Rev. Lett.* **60** 871