OPEN ACCESS

Matrix product operator representations

New Journal of Physics
The open-access journal for physics

Matrix product operator representations

B Pirvu[1], V Murg[1,2], J I Cirac[2] and F Verstraete[1]

[1] Fakultät für Physik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria

[2] Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany

E-mail: frank.verstraete@univie.ac.at

New Journal of Physics **12** (2010) 025012 (13pp)

Received 15 August 2009

Published 26 February 2010

Online at http://www.njp.org/

doi:10.1088/1367-2630/12/2/025012

Abstract. We show how to construct relevant families of matrix product operators (MPOs) in one and higher dimensions. These form the building blocks for the numerical simulation methods based on matrix product states and projected entangled pair states. In particular, we construct translationally invariant MPOs suitable for time evolution, and show how such descriptions are possible for Hamiltonians with long-range interactions. We show how these tools can be exploited for constructing new algorithms for simulating quantum spin systems.

Contents

The study of strongly correlated quantum systems is currently receiving a lot of attention. To a large extent, this is due to the formidable progress that has been made in creating such systems under controlled laboratory conditions, such as in optical lattices and ion traps. From the theoretical point of view, major new insights have been obtained into characterizing the nature of the wavefunctions associated with those strongly correlated systems. The concept of matrix product states and their generalizations plays a central role in those new insights, as it provides a sound foundation and justification for the success of numerical renormalization group methods and especially of the density matrix renormalization group (DMRG) [1, 2]. Those insights have led to the development of new algorithms for simulating quantum spin systems; most notable are the algorithms for simulating time evolution [3], [5]–[7] and the ones generalizing DMRGs to higher dimensions [6].

In this work, we are concerned with the efficient construction of the so-called matrix product operators (MPOs), the basic building blocks for those novel algorithms. MPOs were introduced in the paper [7, 8] and form the operator analogue of matrix product states. We will show how to construct translationally invariant MPOs in 1 and 2 dimensions (1D and 2D) that approximate real or imaginary time evolution; in contrast to the time evolved block decimation (TEBD)/DMRG algorithms [4, 5], the translational symmetry is not broken in the Trotter step. This generalizes the constructions reported in [16]. Second, we construct MPO descriptions for general Hamiltonians with decaying long-range interactions. This is very interesting in the light of simulating quantum spin systems with long-range interactions.

Similar work for constructing MPO representations of Hamiltonians has independently been reported in [9, 11]. Reference [11] gives a good presentation of MPOs from the point of view of the DMRG, and also contains results on how to write spin chain Hamiltonians using MPOs. Crosswhite and Bacon [9] explore the connection between MPOs and Markov processes in depth, and also obtain some results on generalizations to higher dimensions. In [10], an algorithm is devised to simulate quantum spin chains with long-range interactions in the thermodynamic limit; it also contains similar results as reported here on the approximation of power law decaying interactions by sums of exponentials.

1. MPO descriptions of exponentials

1.1. Construction

Let us start with a simple example: suppose we want to simulate real/imaginary time evolution under the Ising Hamiltonian in a transverse field

$$\mathcal{H}_{\mathrm{Is}} = -\sum_{\langle ij\rangle} \sigma_i^z \otimes \sigma_j^z - B \sum_i \sigma_i^x,$$

where only nearest neighbour interactions are considered; both the 1D and 2D case will be considered. As usual, this evolution will be approximated using a Trotter expansion, but we want to do this in such a way that the translational invariance is not broken. Therefore, we split the Hamiltonian in two parts $\mathcal{H} = H_z + H_x$, where H_z contains all terms with σ^z operators and H_x the ones with σ^x. Obviously, all terms in H_z commute, and therefore $O_z = \exp(\epsilon H_z)$ can be calculated exactly. As we will show, O_z has a very simple and elegant MPO description, and of course O_x has a trivial MPO description as it is a product of strictly local operators. Time evolution can now be described within the formalism of matrix product states by evolving the matrix product states (MPS) under the action of the MPO $O_x O_z$.

Let us next show how the MPO of O_z can be constructed. First, observe that

$$\begin{aligned}\exp(\epsilon Z\otimes Z) &= \cosh(\epsilon) I\otimes I + \sinh(\epsilon) Z\otimes Z\\ &= \underbrace{\begin{pmatrix}\sqrt{\cosh\epsilon} & 0\end{pmatrix}}_{B_0^{\mathrm{T}}}\begin{pmatrix}\sqrt{\cosh\epsilon}\\ 0\end{pmatrix} I\otimes I + \underbrace{\begin{pmatrix}0 & \sqrt{\sinh\epsilon}\end{pmatrix}}_{B_1^{\mathrm{T}}}\begin{pmatrix}0\\ \sqrt{\sinh\epsilon}\end{pmatrix} Z\otimes Z\\ &= \sum_{ij}\left(B_i^{\mathrm{T}} B_j\right) Z^i\otimes Z^j.\end{aligned}$$

Here we used the notation $Z^0 = I$, $Z^1 = \sigma_z = Z$ and defined the vectors B_i. Let us now consider the translationally invariant 1D case of N spins with periodic boundary conditions

$$\begin{aligned}\exp\left(\epsilon\sum_i Z_i Z_{i+1}\right) &= \prod_i \exp(\epsilon Z_i Z_{i+1})\\ &= \sum_{i_1 j_1 i_2 j_2 \ldots j_N j_1}\left((B_{i_1}^{\mathrm{T}} B_{i_2})(B_{j_2}^{\mathrm{T}} B_{j_3})\ldots(B_{j_N}^{\mathrm{T}} B_{j_1})\right) Z_1^{i_1} Z_1^{j_1}\otimes Z_2^{i_2} Z_2^{j_2}\otimes\ldots\\ &= \sum_{i_1 j_1 i_2 j_2 \ldots}\mathrm{Tr}\left(B_{j_1} B_{i_1}^{\mathrm{T}} B_{i_2} B_{j_2}^{\mathrm{T}} B_{j_3}\ldots B_{i_N} B_{j_N}^{\mathrm{T}}\right) Z_1^{i_1+j_1}\otimes Z_2^{i_2+j_2}\otimes\ldots\\ &= \sum_{k_1 k_2\ldots}\mathrm{Tr}\left(\underbrace{\left(\sum_{i_1} B_{i_1\oplus k_1} B_{i_1}^{\mathrm{T}}\right)}_{C^{k_1}}\left(\sum_{i_2} B_{i_2\oplus k_2} B_{i_1}^{\mathrm{T}}\right)\ldots\right) Z_1^{k_1}\otimes Z_2^{k_2}\ldots\\ &= \sum_{k_1 k_2\ldots}\mathrm{Tr}\left(C^{k_1} C^{k_2}\ldots C^{k_N}\right) Z_1^{k_1}\otimes Z_2^{k_2}\ldots\end{aligned}$$

In the third step, we made use of the cyclic property of the trace, and in the fourth step, we made a change of variables $k_1 = i_1 \oplus j_1$ where binary arithmetic is assumed. We have therefore proven that $\exp\left(\epsilon\sum_i Z_i Z_{i+1}\right)$ has a very efficient matrix product description with the matrices C^k given by

$$\begin{aligned}C^0 &= \sum_i B_i B_i^{\mathrm{T}} = \begin{pmatrix}\cosh(\epsilon) & 0\\ 0 & \sinh(\epsilon)\end{pmatrix},\\ C^1 &= \sum_i B_{i\oplus 1} B_i^{\mathrm{T}} = \begin{pmatrix}0 & \sqrt{\sinh(\epsilon)\cosh(\epsilon)}\\ \sqrt{\sinh(\epsilon)\cosh(\epsilon)} & 0\end{pmatrix}.\end{aligned}$$

A big advantage of this precise MPO formulation is that it is symmetric; the spectral properties of the associated transfer operator are hence well behaved, which is important if used in algorithms with periodic boundary conditions [17].

In 2D, we can repeat exactly the same argument and obtain the projected entangled pair state (PEPS) description of the operator

$$\exp\left(\epsilon \sum_{\langle ij\rangle} Z_i Z_j\right) = \sum_{x_1 x_x \ldots} F(C^{x_1}, C^{x_2}, \ldots) Z_1^{x_1} \otimes Z_2^{x_2} \otimes \ldots$$

with tensors

$$C^x_{\alpha\beta\gamma\delta} = \sum_{i+j+k+l=x} B_i(\alpha) B_j(\beta) B_k(\gamma) B_l(\delta).$$

Here, $B_i(\alpha)$ means the α component of the vector B_i, $x \in \{0, 1\}$ and the sum is taken over $i, j, k, l = 0 : 1$ with the condition that $i + j + k + l = x$ in binary arithmetic. This proves that the PEPS description of the operator $\exp(\epsilon \sum_{\langle ij\rangle} Z_i Z_j)$ has bond dimension 2. Note that no approximations were made and as such this statement is valid for any value of ϵ. In particular, this gives the MPO description for the classical Ising partition function; its free energy can therefore be calculated by contracting the tensor network consisting of tensors C^0.

The previous analysis can trivially be generalized to the case of any Hamiltonian that is a sum of commuting terms: for this class of Hamiltonians, $\exp(\epsilon H)$ has a very simple MPO description. As this holds for any ϵ, it also holds for all thermal states, and by taking $\epsilon \to -\infty$ it is proven that all ground states of such Hamiltonians have exact MPO descriptions that can easily be constructed. Notable examples of this are the toric code state of Kitaev and the family of string net states [14, 15].

From numerical considerations, it is useful if the matrices/tensors occurring in the MPO description are real and symmetric. There are some tricks for how to achieve this. Consider for example the Heisenberg antiferromagnetic Hamiltonian

$$\mathcal{H}_{\text{Heis}} = \sum_{\langle ij\rangle} \left(X_i X_j + Y_i Y_j + Z_i Z_j\right).$$

The operator $\exp(-\beta \mathcal{H}_{\text{Heis}})$ can be decomposed in Trotter steps consisting of H_x, H_y, H_z, and every Trotter term involves operators of the form $\exp(-\epsilon H_x)$. As we saw in the previous section, the associated matrices involve terms like $\sqrt{\sinh(\epsilon)}$, which becomes complex when $\epsilon > 0$. What we can do however is a change of basis on every second site (this obviously only works for bipartite lattices), where we rotate the spins with the unitary operator $Y = \sigma_y$; this maps $X_{2n} \to -X_{2n}$, $Y_{2n} \to Y_{2n}$, $Z_{2n} \to -Z_{2n}$. On the level of the Hamiltonian, this flips the sign of the H_x and H_z interactions, for which the associated operators $\exp(+\epsilon H_x)$ have indeed real and symmetric MPO descriptions. The problem seems to remain however with the operator $\exp(-\epsilon H_y)$. This can however be easily cured by defining the real antisymmetric matrix $\tilde{Y} = \mathrm{i}Y$ for which $H_{\tilde{y}} = -H_y$ when we replace all operators Y by $\tilde{Y}$. Next, $\exp(+\epsilon H_{\tilde{y}})$ can again be expressed as a MPO; however, we have to be careful as $\tilde{Y} \cdot \tilde{Y} = -I$ as opposed to $+I$. Looking back at the derivation of the MPO for the Ising case, we can easily see that this sign can be absorbed into C, and we can express

$$\exp\left(-\epsilon \sum_i Y_i Y_{i+1}\right) = \sum_{k_1 k_2 \ldots} \mathrm{Tr}(\tilde{C}^{k_1} \tilde{C}^{k_2} \ldots) \tilde{Y}_1^{k_1} \otimes \tilde{Y}_2^{k_2} \ldots$$

as an MPO with matrices

$$\tilde{C}^0 = \sum_i B_i B_i^{\mathrm{T}} \cdot (-1)^i = \begin{pmatrix} \cosh(\epsilon) & 0 \\ 0 & -\sinh(\epsilon) \end{pmatrix},$$

$$\tilde{C}^1 = C^1 = \sum_i B_{i+1} B_i^{\mathrm{T}} = \begin{pmatrix} 0 & \sqrt{\sinh(\epsilon)\cosh(\epsilon)} \\ \sqrt{\sinh(\epsilon)\cosh(\epsilon)} & 0 \end{pmatrix}.$$

Of course the same can be done in 2D. Here we obtain

$$\tilde{C}^x_{\alpha\beta\gamma\delta} = \sum_{i+j+k+l=x} B_i(\alpha) B_j(\beta) B_k(\gamma) B_l(\delta) \sqrt{-1}^{(x+i+j+k+l)},$$

where the sum in the power $\sqrt{-1}^{(x+i+j+k+l)}$ is not in binary arithmetic. This clearly leads to a real translationally invariant MPO parameterization.

1.2. Algorithms

It is now obvious how to turn those MPO-descriptions to our advantage for constructing new algorithms for the simulation of quantum spin chains.

Let us first consider the case of imaginary time evolution, where the goal is to evolve a state in imaginary time such as to simulate a thermal (finite β) or ground state ($\beta \to \infty$). Obviously, we will use the Trotterization described in the previous section. The big advantage there is that the translational invariance is never broken, and furthermore that the matrices involved in the MPS description of the MPO are real and symmetric. In particular, this means that, if we start with a translationally invariant MPS with real symmetric MPS description, then it will stay like that during the whole course of the evolution. This has a dramatic effect on the numerical conditioning and stability of the algorithm.

The algorithm for time evolution is now as follows: given a translationally invariant MPS with matrices $\{A^i\}$ with bond dimension D and MPO with matrices $\{B^i\}$, $\{X^i\}$ of dimension D', we want to find a way of representing cutting the bond dimension of the MPS $\{C^i\}$ given by

$$C^i = \sum_{jk} A^j \otimes B^k \langle i|X^k|j\rangle$$

in an optimal way. This can easily be done as follows: calculate the leading eigenvector x of the transfer operator $E = \sum_i C^i \otimes C^i$ (note that E is symmetric and as such this is a very well conditioned problem). Rewriting x as a $DD' \times DD'$ positive semidefinite matrix, we can easily calculate its singular value decomposition $x = U\Sigma U^\dagger$. We now define the projector/isometry P as the rectangular matrix consisting of the first D rows of U, and act with this P on the matrices C^i. The updates to matrices A^i are therefore obtained by $A^i \equiv P^\dagger C^i P$ which is obviously still symmetric and real. Clearly, all those steps have to be done in such a way as to exploit the sparse nature of the problem, such as done in DMRG, which leads to a complexity that scales like D^3. Also, if the eigenvalues that are thrown away are not small enough, we can always increase the bond dimension.

The big advantage of this procedure is that it is extremely well conditioned and very efficient to implement. This allows it to go to very large bond dimensions. Notably, as compared

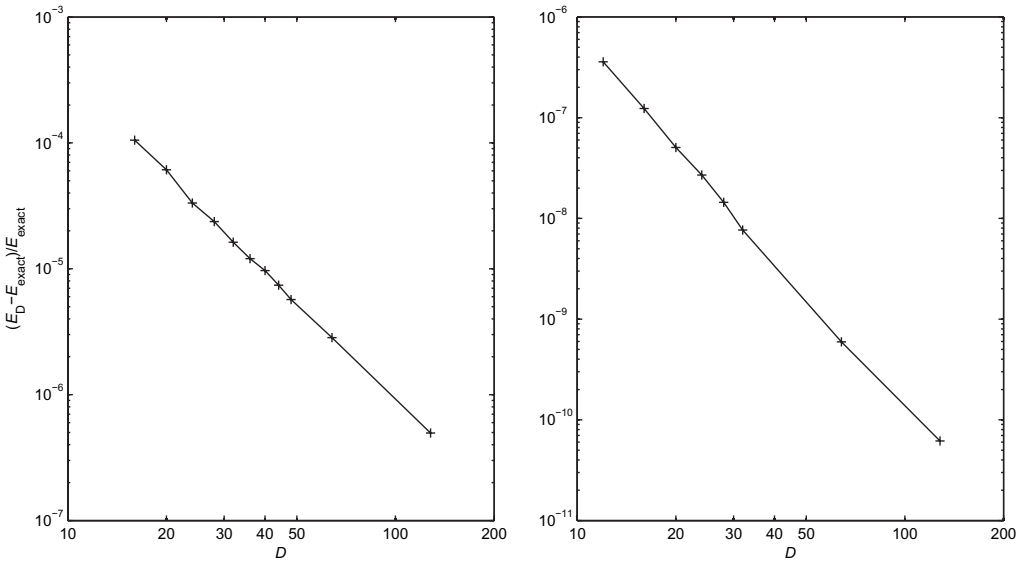

Figure 1. Dependence of the precision $(E_{\mathrm{D}} - E_{\mathrm{exact}})/E_{\mathrm{exact}}$ on the bond dimension D for the Heisenberg antiferromagnetic spin chain (left) and the critical Ising chain in a transverse field (right).

to the original formulation of the TEBD algorithm, we do not have to take inverses at any time (because the gauge degrees of freedom are trivial as they consist of unitary matrices), and furthermore it works equally well if the MPO is very far from the identity operator (this is important in the context of PEPS algorithms).

The same ideas can of course be used in the case of real time evolution. In that case, the matrices involved become complex symmetric, and it might be beneficial to apply some gauge conditions to optimize the stability. This can be done as follows: given x, we want to find the complex (symplectic) matrix Q, $Q \cdot Q^{\mathrm{T}} = \mathbb{1}$ such that the condition number (i.e. smallest divided by largest singular value) of the matrix $QxQ^{\dagger}$ is as large as possible. This optimization problem can be solved recursively as follows: calculate the singular value decomposition of $x = vsv^{\dagger}$, choose the generator $G_k = -G_k^{\mathrm{T}}$ as $G_k = \mathrm{Im}(v_D v_D^{\dagger} - v_1 v_1^{\dagger})$, and make the substitution $x \to \exp(-\mathrm{i}\epsilon G)x\exp(\mathrm{i}\epsilon G)$ for small enough ϵ, and repeat this until convergence. Convergence is equivalent to the derivative of the condition number being zero. The final gauge transform to be implemented is the product of all infinitesimal transformations $Q = \prod \exp(\mathrm{i}\epsilon G_k)$. Note again that all of this becomes trivial in the case of real symmetric matrices (such as occurring in imaginary time evolution): in that case the Q cannot change the condition number as they are unitary.

We have tested these new algorithms on the critical Ising and Heisenberg spin chain models, and obtained results that are consistent with what we expected. In particular, for the Heisenberg antiferromagnetic spin chain, we obtain a precision of $(E_{D=64} - E_{\mathrm{exact}})/E_{\mathrm{exact}} = 2.83 \times 10^{-6}$ with very modest calculations. In the case of the critical Ising chain in a transverse field, we obtain $(E_{D=64} - E_{\mathrm{exact}})/E_{\mathrm{exact}} = 5.95 \times 10^{-10}$. The dependence of the precision on the bond dimension D can be gathered from figure 1.

The algorithms for the 2D analogue will be discussed elsewhere [17].

2. MPO descriptions of Hamiltonians with long-range interactions

2.1. Construction

Let us next investigate how to represent Hamiltonians with long-range interactions of the form

$$\mathcal{H} = \sum_{ij} f(i-j) Z_i Z_j$$

with $f(i-j)$ some decaying function. The first question to ask is whether it is still possible to find an exact MPO description of $\hat{O} = \exp(\epsilon \mathcal{H})$. It can easily be seen that this is not possible if the function $f(x)$ does not vanish at some finite distance: otherwise, the action of $\hat{O}$ on a MPS could increase the Schmidt number over any cut with an arbitrary large amount, and hence no finite MPO description is possible. This is the reason why the transfer matrix approach in classical 1D spin systems breaks down for such long-range interactions.

So let us be less ambitious and try to find an MPO description of the Hamiltonian itself. This is interesting for several reasons: first, it is useful in constructing algorithms for time evolution using iterative methods like Lanczos, and second, it allows us to calculate quantities like $\langle\psi|H^2|\psi\rangle$ efficiently.

As a start, let us consider a general 1D spin 1/2 Hamiltonian with nearest-neighbour interactions. If the Hamiltonian is translationally and reflection invariant, then there always exists a basis such that the Hamiltonian can be written as

$$\mathcal{H} = \sum_{\alpha,i} \mu_\alpha \sigma_\alpha^i \otimes \sigma_\alpha^{i+1} + \sum_j \hat{O}^j,$$

where $\hat{O}$ can be any one-qubit operator. Similarly to the construction of MPS descriptions of the W-state [12], a MPO can be constructed to represent this $\mathcal{H}$ by making use of nilpotent matrices:

$$\begin{aligned}
\mathcal{H} &= \sum_{i_1 i_2 \ldots} \left(v_{\mathrm{l}}^{\mathrm{T}} B_{i_1} B_{i_2} \ldots B_{i_N} v_{\mathrm{r}}\right) X_{i_1} \otimes X_{i_2} \otimes \ldots X_{i_N}, \\
X_0 &= I, \quad X_1 = \sigma_x, \quad X_2 = \sigma_y, \quad X_3 = \sigma_z, \quad X_4 = \hat{O}, \\
v_{\mathrm{l}} &= |0\rangle, \quad v_{\mathrm{r}} = |4\rangle, \\
B_0 &= |0\rangle\langle 0| + |4\rangle\langle 4|, \\
B_1 &= |0\rangle\langle 1| + \mu_1 |1\rangle\langle 4|, \quad B_2 = |0\rangle\langle 2| + \mu_2 |2\rangle\langle 4|, \quad B_3 = |0\rangle\langle 3| + \mu_3 |3\rangle\langle 4|, \\
B_4 &= |0\rangle\langle 4|.
\end{aligned}$$

The simplest way of deriving this is to think about the Hamiltonian as a Markov process with five possible symbols (remember that MPS can be constructed using Markov processes), such that a symbol X_1, X_2, X_3 is always followed by itself and then all zeros X_0 and X_4 by all zeros. As such, one can easily prove that $D = 5$ is optimal in this case because this is the operator Schmidt number of the Hamiltonian when splitting it into two pieces. Note that if only Ising interactions would have been considered, then $D = 3$ would have been sufficient and we could have chosen

$$B_0 = |0\rangle\langle 0| + |2\rangle\langle 2|, \quad B_1 = |0\rangle\langle 1| + \mu_1 |1\rangle\langle 2|.$$

Note that there is no need for B_2, B_3 and B_4 in that case.

It is obvious how to generalize this description to the case of higher dimensional systems and to the case of exponentially decaying interactions. Let us first look at the case of exponentially decaying interactions. By adding diagonal terms to B_0

$$B_0 = |0\rangle\langle 0| + \lambda_x |1\rangle\langle 1| + \lambda_y |2\rangle\langle 2| + \lambda_z |3\rangle\langle 3| + |4\rangle\langle 4|,$$

we can immediately check that the corresponding Hamiltonian/MPO is given by

$$\mathcal{H} = \sum_{\alpha, i<j} \mu_\alpha \lambda_\alpha^{i-j} \sigma_\alpha^i \otimes \sigma_\alpha^j + \sum_j \hat{O}^j$$

which is a spin chain with exponentially decaying interactions.

Unfortunately, it is impossible to obtain exact MPO descriptions when the interactions are decaying following a power law. However, inverse polynomials can pretty well be approximated by sums of exponentials (this is the reason why the DMRG is able to reproduce the correlations in critical models pretty well). Hamiltonians with power law decay of correlations can therefore be well approximated by sums of MPOs, which is itself an MPO. Actually, very few exponentials are needed to obtain a good approximation, even at large distances. The problem of finding the optimal weights and exponents for such an approximation problem for a general function $f(k)$, i.e.

$$\min_{x_i, \lambda_i} \sum_{k=1}^{N} |f(k) - \sum_{i=1}^{n} x_i \lambda_i^k|,$$

is not completely trivial. In the appendix, we present a simple method that solves this optimization problem for general $f(k)$ and a given number of exponentials n and a number of sites $N > n$ (the method works for any functions, and returns complex exponents in the case of oscillating functions as should be). If we choose a power law decay with cube power 3, $N = 1000$ and $n = 10$ then the above cost function is 10^{-5} (the maximal difference between the function and the approximated one is 5×10^{-8}). This maximal difference falls to 3×10^{-6} for power 2 and 3×10^{-4} for power 1.

In conclusion, we found the exact MPO description for Hamiltonians with exponentially decaying interactions. Hamiltonians with power law decay can be approximated very well using sums of this MPO. The MPOs obtained for the description of Hamiltonians are of a very different form than the ones obtained by taking the exponential. The main difference is that the corresponding transfer matrices will always contain a Jordan block structure, and one has to be careful in dealing with such situations when considering the thermodynamic limit.

Let us now turn to the 2D case. We again first consider the square lattice with only nearest neighbour interactions. There is a very simple way of writing down a PEPS description that achieves the task: first, consider the MPS

$$|W\rangle = \sum_k |0\rangle_1 |0\rangle .. |0\rangle |1\rangle_k |0\rangle \ldots |0\rangle_{N^2}$$

which is the equal superposition of having one spin up and all other ones down over all sites. Note that this MPS has bond dimension 2, and can therefore trivially be represented as a PEPS with bond dimension 2. The idea is that this excitation specifies where to put an interaction. Let

us next consider the tensors

$$B^0_{i;\alpha,\beta,\gamma,\delta} = |0\rangle\langle 0000|,$$

$$B^1_{i;\alpha,\beta,\gamma,\delta} = |1\rangle\langle 00|\,(\langle 01| + \langle 10|) + |0\rangle\,(\langle 01| + \langle 10|)\,\langle 00|,$$

where we assume that the indices α, β are the left respectively top indices, and the associated operators

$$X^0 = I,$$

$$X^1 = Z.$$

It can readily be seen that we obtain the Ising Hamiltonian if we act with the $|W\rangle$ state on the fifth index of the tensors: the $|W\rangle$ state puts one index i equal to one, and the other terms are such that an interaction to the right and below it will be created. The total bond dimension of the corresponding PEPS (including the $|W\rangle$) is therefore 4.

Decaying interactions between one spin and all other ones can however be obtained in a much more elegant way; as we will show, it is even possible to model power law decay of interactions exactly. The idea is as follows: the critical classical Ising model in 2D has power law decay of correlations. Consider the quantum state

$$|\psi_\beta\rangle = \exp\left(-\beta \sum_{\langle ij\rangle} Z_i Z_j\right) (|+\rangle)^{\otimes N},$$

where $|+\rangle$ stands for the superposition $|0\rangle + |1\rangle$. This is obviously a PEPS, as it is obtained by acting with an MPO (see earlier) on a product state. The partition function of the Ising model at temperature β is obtained by calculating the overlap

$$(\langle +|)^{\otimes N} |\psi(\beta)\rangle,$$

and correlation functions between two spins are obtained by replacing the corresponding $\langle +|$ at the left side of this expression by $\langle -| = \langle 0| - \langle 1|$. Instead of the $|W\rangle$ state in the previous example, we will use a state $|W''\rangle$ that is the equal superposition of two excited spins as opposed to one. The MPS description of $|W''\rangle$ has bond dimension 3 and is similar to the ones derived for the 1D Hamiltonians with exponential decay where we put the parameter $\lambda = 1$ (i.e. $B_0 = I$). This gives us all the necessary ingredients to construct the MPO description of the Hamiltonian:

$$\mathcal{H} = \sum_{i_1 i_2 \ldots i_N} \left(\langle x^{i_1}| \langle x^{i_2}| \ldots \langle x^{i_N}|\right) \left(|\psi_\beta\rangle |W''\rangle\right) X^{i_1} \otimes X^{i_2} \ldots \otimes X^{i_N},$$

$$|x^0\rangle = |0\rangle|+\rangle, \quad |x^1\rangle = |1\rangle|-\rangle, \quad X^0 = I, \quad X^1 = Z.$$

Here the vectors $|x^i\rangle$ act on two qubits, one on the corresponding qubit of $|\psi_\beta\rangle$ and the other on $|W''\rangle$. The $|W''\rangle$ state enforces that exactly two operators X^i will be nontrivial, and $|\psi_\beta\rangle$ gives the right weight to the associated interaction. As those are products of PEPS, the result is a PEPS with bond dimension $2 \times 3 = 6$. This is an amazing result: as opposed to the 1D case, there is an exact PEPS description for Hamiltonians with two-body interactions that decay as the $r^{-\nu}$ with $\nu = 1$ the critical exponent of the Ising model.

Obviously, this construction can be repeated for any classical spin model, and hence many different exponents can in principle be taken. It is as yet an open question how to engineer the PEPS so as to obtain a specific exponent, although very good approximations can again be obtained by making use of sums of exponentials.

2.2. Algorithms

It is obvious how to make use of all this in algorithms for simulating quantum spin chains. First of all, it is clear how to extend the variational MPS method described in [13] to the present case. For this, we have to consider finite systems with open boundary conditions and matrix product states that have site-dependent matrices in their MPS description. The optimization $\langle\psi|H|\psi\rangle$ can then be done using the alternating least squares method described in [13]. As expected, numerical tests showed very good convergence properties.

The problem of time evolution is a little more challenging, as we cannot use the Trotterization tricks. However, Krylov-based methods can be used (see [18] for a review), and are the method of choice here.

These finite-dimensional algorithms cannot readily be generalized to the infinite case however. A sensible way for determining ground state energies in that limit would be to first use the finite dimensional algorithm just described, and then use a brute-force gradient-based optimization method for optimizing the infinite case. For this we need to be able to calculate expectation values $\langle\psi|\hat{O}|\psi\rangle$ in the thermodynamic limit, i.e. when $|\psi\rangle$ is an infinite MPS with matrices $\{A^i\}$ and $\hat{O}$ the MPO description of a Hamiltonian with exponentially decaying interactions. The idea is to consider a family of MPOs O_N whose support is limited to N sites (i.e. the Hamiltonian only acts on N sites), calculate the energy with respect to the infinite MPS (this energy will scale linearly in N), and then take the thermodynamic limit to calculate the energy per site. It holds that

$$\langle\psi|\hat{O}_N|\psi\rangle = \langle L|E_{\mathrm{H}}^N|R\rangle,$$

where $|L\rangle = |x_{\mathrm{l}}\rangle|v_{\mathrm{l}}\rangle$, $|R\rangle = |x_{\mathrm{r}}\rangle|v_{\mathrm{r}}\rangle$ with $|x_{\mathrm{l}}\rangle, |x_{\mathrm{r}}\rangle$ the left/right eigenvectors of the transfer matrix $E_0 = \sum_i A_i \otimes \bar{A}_i$; the vectors $v_{\mathrm{l}}, v_{\mathrm{r}}$ are the ones used in the MPO description of the Hamiltonian, and

$$E_{\mathrm{H}} = \sum_{ijk} A_i \otimes B_j \otimes \bar{A}_k \langle k|X_j|i\rangle.$$

The eigenstructure of E_{H} is nontrivial because it has Jordan blocks. In the present case, the only relevant blocks are of size 2 (larger blocks would lead to an energy that scales superlinearly with the size of the support of H, which cannot be). Using the previous notation, one sees that the left/right eigenvector corresponding to the largest eigenvalue in magnitude d_0 is given by $\langle q_{\mathrm{l}}| = \langle x_{\mathrm{l}}|\langle 0|/|q_{\mathrm{r}}\rangle = |x_{\mathrm{r}}\rangle|0\rangle$. The generalized eigenvectors can now be found by solving the equation $(E_{\mathrm{H}} - d_0 I)|\tilde{q}_{\mathrm{r}}\rangle = |q_{\mathrm{r}}\rangle$, $\langle\tilde{q}_{\mathrm{l}}|\,(E_{\mathrm{H}} - d_0 I) = \langle q_{\mathrm{l}}|$. We next define the matrices $Q_{\mathrm{l}} = (|q_{\mathrm{l}}\rangle, |\tilde{q}_{\mathrm{l}}\rangle)$, $Q_{\mathrm{r}} = (|q_{\mathrm{r}}\rangle, |\tilde{q}_{\mathrm{r}}\rangle)$ and the 2×2 matrix $Q = (Q_{\mathrm{l}}^{\mathrm{T}} Q_{\mathrm{r}})^{-1}$. In the limit of large N, it holds that

$$E_{\mathrm{H}}^N \simeq Q_{\mathrm{r}} \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} Q Q_l^{\mathrm{T}}.$$

As $Q_{\mathrm{l}}^{\mathrm{T}}|q_{\mathrm{r}}\rangle = \begin{pmatrix}1\\0\end{pmatrix} = Q_{\mathrm{r}}^{\mathrm{T}}|q_{\mathrm{l}}\rangle$, the expectation value $\langle\psi|\hat{O}_N|\psi\rangle$ is given by

$$\begin{pmatrix}1 & N\end{pmatrix} Q \begin{pmatrix}1\\0\end{pmatrix}$$

in the limit of large N. The energy per site is therefore given by

$$Q_{12} = -1/\langle\tilde{q}_{\mathrm{l}}|q_{\mathrm{r}}\rangle = -1/\langle q_{\mathrm{l}}|\tilde{q}_{\mathrm{r}}\rangle = -1/\langle q_{\mathrm{l}}|\,(E_{\mathrm{H}} - d_0 I)^{-^{\dagger}}\,|q_{\mathrm{r}}\rangle.$$

In a similar way, it is possible to calculate expectation values of the operator $\langle\psi|(H - \lambda)^2|\psi\rangle$. This is relevant because it gives an exact bound on how far a given MPS $|\psi\rangle$ is from an exact eigenvector of H. As we have a squared term, Jordan blocks of dimension three will be encountered. As before, we define

$$E_{\mathrm{H}^2} = \sum_{ijkl} A_i \otimes B_j \otimes \bar{B}_k \otimes \bar{A}_l \langle l|X_k^{\dagger} X_j|i\rangle.$$

The relevant right eigenvector is again of the form $|q_{\mathrm{r}}\rangle = |x_{\mathrm{r}}\rangle|0\rangle|0\rangle$ and we can find the eigenstructure of the associated Jordan block as follows: start by calculating $|\tilde{q}_{\mathrm{r}}\rangle$ as we did in the previous section using the operator E_{H} instead of E_{H^2}; next define $|\tilde{q}'\rangle = \mathrm{sym}(|\tilde{q}_{\mathrm{r}}\rangle|0\rangle)$ where the symbol 'sym' means symmetrization with respect to the part of the state acting on the Hamiltonian part of the MPO (the antisymmetrized wavefunction turns out to be an irrelevant eigenvector of E_{H^2} with eigenvalue d_0); finally solve the linear set of equations $(E_{\mathrm{H}^2} - d_0 I)\,|\tilde{q}''_{\mathrm{r}}\rangle = |\tilde{q}'_{\mathrm{r}}\rangle$. The relevant eigenstructure is now given by the matrix $Q_{\mathrm{r}} = \left(|q_{\mathrm{r}}\rangle, |\tilde{q}'_{\mathrm{r}}\rangle, |\tilde{q}''_{\mathrm{r}}\rangle\right)$. Similarly, one can calculate Q_{l} and $Q = (Q_{\mathrm{l}}^{\mathrm{T}} Q_{\mathrm{r}})^{-1}$. The final expectation value is then given by

$$(1, N, N(N-1)/2)\; Q \begin{pmatrix}1\\0\\0\end{pmatrix}.$$

The matrix Q therefore contains all the relevant information about the energies and their scaling when $N \to \infty$.

The energy can now easily be optimized with a brute-force gradient-based optimization routine.

Concerning the 2D MPO representing Hamiltonians, it turns out that they are very valuable for speeding up actual calculations done by the PEPS method: the calculation of the expectation value of the Hamiltonian with respect to a given PEPS can be calculated in one run using this idea, and we no longer have to calculate the expectation value for every term individually.

3. Conclusion

In conclusion, we constructed several examples of interesting families of MPOs in 1D and 2D. Those descriptions turn out to be very valuable for constructing stable and scalable algorithms for simulating quantum spin systems, in 1D and 2D.

Acknowledgments

This work was supported by the SFB (FoQuS) and CoQuS (FWF doctoral program Complex Quantum Systems W1210), by the DFG (Forschergruppe 635) and by the EU (STREP Quevadis and ERC grant QUERG).

Appendix

In this appendix, we show how to solve the problem of approximating any function $f(k)$ as a sum of exponentials for $k = 1 \dots N$:

$$\min_{x_i,\lambda_i} \sum_{k=1}^{N} |f(k) - \sum_{i=1}^{n} x_i \lambda_i^k|.$$

First, construct the rectangular $N-n+1 \times n$ matrix

$$F = \begin{pmatrix} f(1) & f(2) & f(3) & \dots & f(n) \\ f(2) & f(3) & \dots & & \\ f(3) & \dots & & & \\ \vdots & & & & f(N-1) \\ f(N-n+1) & \dots & & f(N-1) & f(N) \end{pmatrix}$$

$$\simeq \underbrace{\begin{pmatrix} \lambda_1^0 & \lambda_2^0 & \dots & \lambda_n^0 \\ \lambda_1^1 & \lambda_2^1 & & \\ \lambda_1^2 & & & \\ \vdots & & & \\ \lambda_1^{N-n} & \lambda_2^{N-n} & \dots & \lambda_n^{N-n} \end{pmatrix}}_{W} \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & & \\ \vdots & & & \\ 0 & & & x_n \end{pmatrix} \begin{pmatrix} \lambda_1^0 & \lambda_1^1 & \dots & \lambda_1^n \\ \lambda_1^0 & \lambda_2^1 & \dots & \lambda_2^n \\ \dots & & & \\ \lambda_n^0 & \dots & & \end{pmatrix}.$$

Note that W is a Vandermonde matrix. We observe that F and W span the same space (note that N is typically much larger than n), such that there exists a $n \times n$ matrix Q s.t. $FQ \simeq W$. Define F_1 as the rectangular matrix which consists of the first $N-n$ rows of F and F_2 as the one with the last $N-n$ rows. Due to the Vandermonde structure of W, it must be approximately true that $F_1 Q \Lambda \simeq F_2 Q$ with Λ the diagonal matrix containing the exponents. Therefore, $\Lambda \simeq Q^{-1} F_1^{-\dagger} F_2 Q$ ($F_1^{-\dagger}$ denotes the pseudoinverse of U_1): the exponents $\{\lambda_i\}$ hence correspond to the eigenvalues of the matrix $F_1^{-\dagger} F_2$ which can be calculated very easily.

This method can be made more robust by making use of the so-called QR-decomposition. This can be done by first calculating the (economical) QR decomposition of $F = UV$ and by defining U_1 as the rectangular matrix which consists of the first $N-n$ rows and n columns of U and U_2 as the one with the last $N-n$ rows: there must again exist a Q such that $UQ \simeq W$. The exponents Λ can therefore be easily calculated as the eigenvalues of the matrix $U_1^{-\dagger} U_2$. The advantage of using the QR-decomposition is that the pseudoinverse of U_1 is much better conditioned than that of F_1.

Once those exponents are found, a simple least-squares algorithm can be used to find the corresponding weights $\{x_i\}$. It happens that this method is very efficient and reliable, even when

oscillating functions are involved. A similar method is known in the field of signal processing under the name Hankel singular value decomposition.

References

[1] White S R 1992 *Phys. Rev. Lett.* **69** 2863–6
[2] Schollwoeck U 2005 *Rev. Mod. Phys.* **77** 259
[3] Vidal G 2003 *Phys. Rev. Lett.* **91** 147902
[4] Daley A J, Kollath C, Schollwoeck U and Vidal G 2004 *J. Stat. Mech.* (2004) P04005
[5] White S R and Feiguin A E 2004 *Phys. Rev. Lett.* **93** 076401
[6] Verstraete F and Cirac J I 2004 arXiv:cond-mat/0407066
[7] Verstraete F, Garcia-Ripoll J J and Cirac J I 2004 *Phys. Rev. Lett.* **93** 207204
[8] Zwolak M and Vidal G 2004 *Phys. Rev. Lett.* **93** 207205
[9] Crosswhite G M and Bacon D 2007 arXiv:0708.1221
[10] Crosswhite G M, Doherty A C and Vidal G 2008 arXiv:0804.2504
[11] McCulloch I P 2007 *J. Stat. Mech.* P10014
[12] Verstraete F, Cirac J I, Latorre J I, Rico E and Wolf M M 2005 *Phys. Rev. Lett.* **94** 140601
[13] Verstraete F, Porras D and Cirac J I 2004 *Phys. Rev. Lett.* **93** 227205
[14] Kitaev A Yu 2003 *Ann. Phys.* **303** 2
[15] Michael Levin A and Wen X-G 2005 *Phys. Rev.* B **71** 045110
[16] Verstraete F, Wolf M M, Perez-Garcia D and Cirac J I 2006 *Phys. Rev. Lett.* **96** 220601
[17] In preparation
[18] Garcia-Ripoll J J 2006 *New J. Phys.* **8** 305