Matrix product operators

Boulder Summer School 2010

Frank Verstraete
University of Vienna

- References:

  - F. Verstraete, J.I. Cirac, V. Murg, Matrix Product States, Projected Entangled Pair States, and variational renormalization group methods for quantum spin systems, Adv. Phys. 57,143 (2008)
  - Ian P. McCulloch, J. Stat. Mech. (2007) P10014
  - V. Murg, J.I. Cirac, B. Pirvu, F. Verstraete, Matrix product operator representations, New J. Phys. 12 025012 (2010)
  - J. I. Cirac, F. Verstraete, Renormalization and tensor product states in spin chains and lattices, J. Phys. A: Math. Theor. 42, 504004 (2009)

Content

- Transfer matrices and matrix product operators MPO in statistical physics
  - Transfer matrices for classical partition functions
  - Quantum Hamiltonians as Matrix product operrators
  - Exponentials of local Hamiltonians as MPO
  - MPO and the Algebraic Bethe Ansatz

- Matrix product states as variational ansatz
  - Why does the ansatz makes sense?
  - DMRG as an alternating least squares optimization problem
  - Time-evolution as an alternating least squares optimization problem
  - Higher dimensional generalizations

Matrix Product States as variational states for simulating strongly correlated quantum systems

- Why?
  - History of Quantum Mechanics is to a large extent one in which we try to find approximate solutions to Schrödinger equation
  - Most relevant breakthroughs in context of many-body physics: guess the right wavefunction (BCS, Laughlin, …)
  - Is there a way to come up with a systematic way of parameterizing the wavefunctions arising in relevant Hamiltonians?
    - In case of 1-D quantum spin chains: NRG / DMRG : MPS
    - In case of 2-D quantum spin systems: PEPS / MERA / ….
  - Central concept: matrix product operators

Transfer matrices in classical spin systems

- Consider partition function for ferromagnetic Ising model

$$Z = \sum_{\{s_i\}} \exp\left(\beta \sum_{\langle i,j \rangle} s_i s_j \right)$$

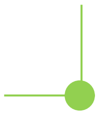
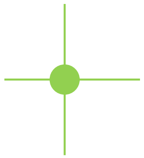
- By introducing dual variables, this can be turned into a product of tensors:

$$\begin{bmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{bmatrix} = \begin{bmatrix} \sqrt{\cosh\beta} \\ \sqrt{\cosh\beta} \end{bmatrix} \begin{bmatrix} \sqrt{\cosh\beta} & \sqrt{\cosh\beta} \end{bmatrix} + \begin{bmatrix} \sqrt{\sinh\beta} \\ -\sqrt{\sinh\beta} \end{bmatrix} \begin{bmatrix} \sqrt{\sinh\beta} & -\sqrt{\sinh\beta} \end{bmatrix}$$

$$x_{s,\mu} = \begin{bmatrix} \sqrt{\cosh\beta} & \sqrt{\sinh\beta} \\ \sqrt{\cosh\beta} & -\sqrt{\sinh\beta} \end{bmatrix}$$

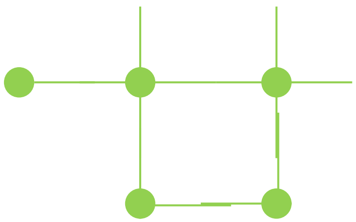
$$Z = \sum_{\{\mu_\alpha\}} \sum_{\{s_i\}} x_{s_i \mu_\alpha}$$

Much easier to work with diagrams:

- Vector: x_μ 
- Matrix: $x_{\mu\nu}$  
- Tensor with 4 legs: $x_{\alpha\beta\gamma\delta}$ 
- …

- Multiplying matrices: $y_{\mu\pi} = \sum_i x_{\mu\nu} x_{\nu\pi}$ 

- Contracting tensors:



1-dimensional Ising model

$$Z = \sum_{\{\mu_\alpha\}} \sum_{\{s_i\}} x_{s_i \mu_\alpha}$$

$x_{s\mu} =$ 

Z= 

This is a product of matrices: transfer matrix T= 

$$\mathrm{T}_{\mu_1\mu_2} = \sum_s x_{s\mu_1} x_{s\mu_2} = \begin{bmatrix} 2\cosh\beta & 0 \\ 0 & 2\sinh\beta \end{bmatrix}$$

$$Z = \mathrm{Tr}\left(T^N\right)$$

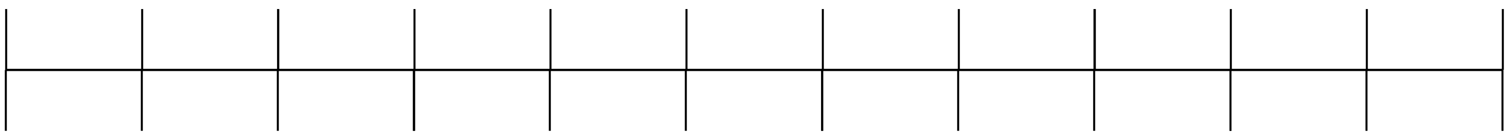
Therefore the partition function can efficiently be calculated

2-dimensional Ising model

- Equivalent construction:

$$Z = \mathrm{Tr}\left(T^{N}\right) \qquad \mathrm{A}_{\alpha\beta\gamma\delta} = \sum_{s} x_{s\alpha} x_{s\beta} x_{s\gamma} x_{s\delta}$$

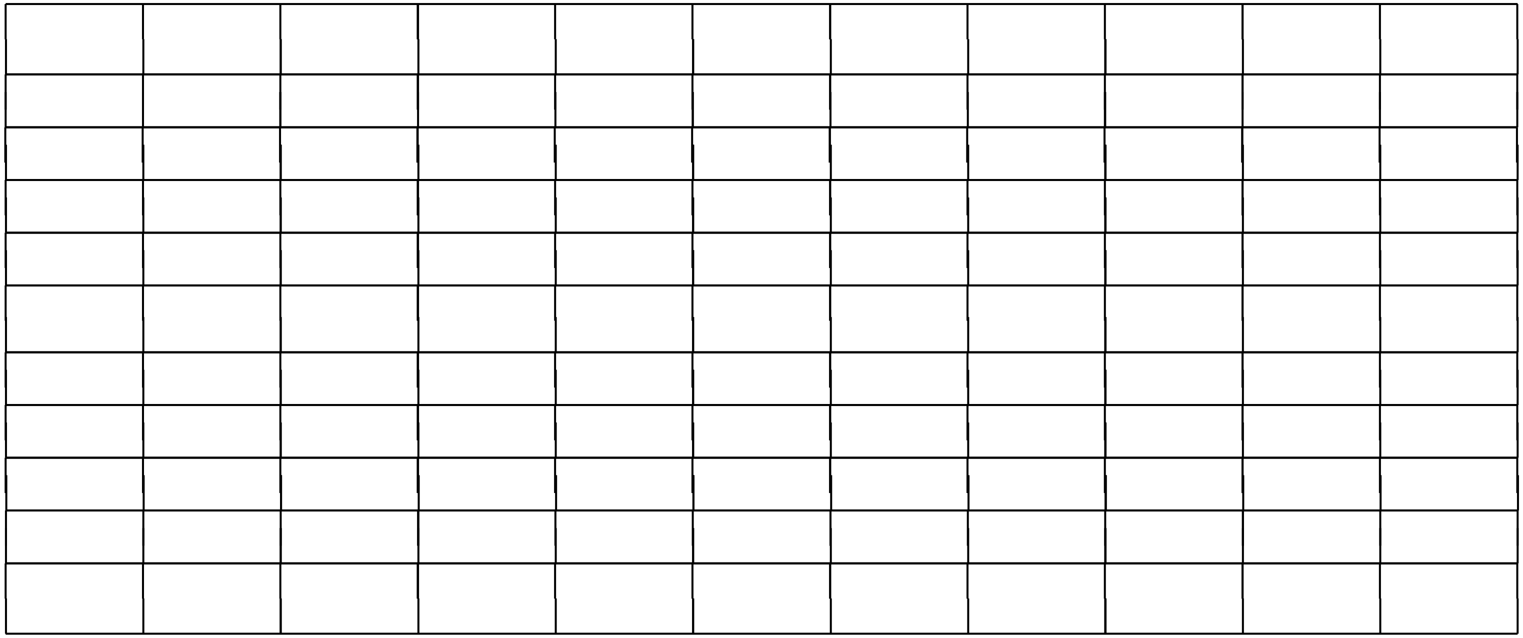
T=



- T is the transfer matrix, and can be written in the form of a matrix product operator: $T = \sum_{i_1 i_2 i_3 \ldots} Tr\left\{A^{i_1} A^{i_2} A^{i_3} \ldots\right\} O_{i_1} \otimes O_{i_2} \otimes O_{i_3} \otimes \ldots$

$$\left| x^{\pm} \right\rangle = \begin{bmatrix} \sqrt{\cosh(\beta)} \\ \pm\sqrt{\sinh(\beta)} \end{bmatrix} \qquad \mathrm{A}^{\pm} = \left| x^{\pm} \right\rangle \left\langle x^{\pm} \right| \qquad \mathbf{O_i} = A^{i}$$

Z=

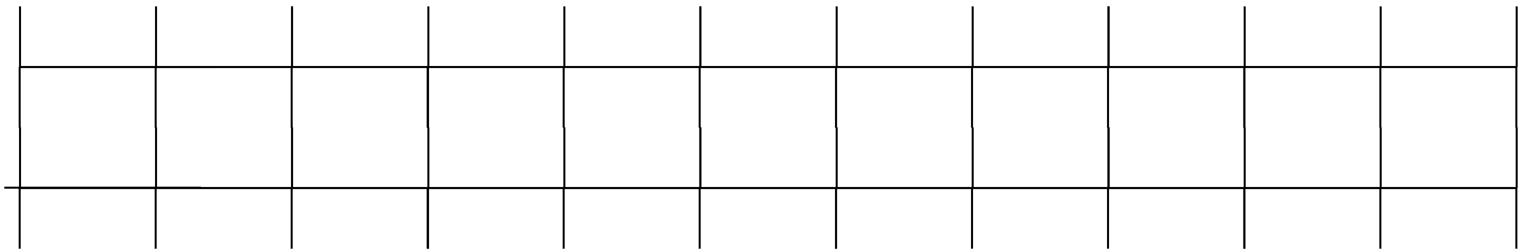


Calculating the partition function for the infinite system is equivalent to finding leading eigenvalue of the transfer matrix T

$$\lambda_0 = \max_{|\Psi\rangle} \frac{\langle \Psi | T | \Psi \rangle}{\langle \Psi | \Psi \rangle} \;\Rightarrow\; \mathrm{f} = -\beta \log \lambda_0$$

Properties of MPO

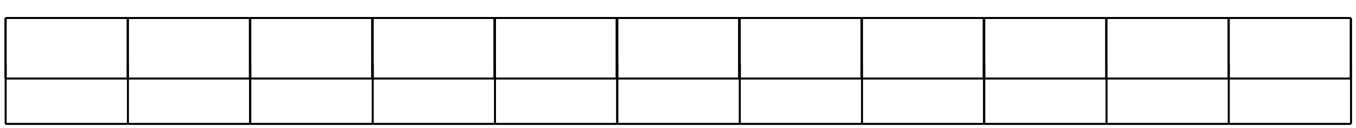
- Algebra of MPO: product of 2 MPO yields another MPO with larger bond dimension



- Gauge transformations leave MPO invariant
- Matrix product states: subclass of MPO in the sense that operators Oi are vectors:

  – Multiplication of MPS with MPO yields MPS with larger bond dimension

- Central property: expectation values of MPO with respect to MPS can be calculated efficiently (simple matrix multiplication):



Related problem: find ground states of 1-D quantum Hamiltonian

- Local quantum Hamiltonians of the form

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

can be written in MPO form:

$$\mathcal{H} = \sum_{i_1 i_2 \ldots} \left(v_l^T B_{i_1} B_{i_2} \ldots B_{i_N} v_r\right) X_{i_1} \otimes X_{i_2} \otimes \ldots X_{i_N}$$

$$\begin{aligned}
X_0 &= I \qquad X_1 = \sigma_x \qquad X_2 = \sigma_y \qquad X_3 = \sigma_z \qquad X_4 = \ddot{O} \\
v_l &= |0\rangle \qquad v_r = |4\rangle \\
B_0 &= |0\rangle\langle 0| + |4\rangle\langle 4| \\
B_1 &= |0\rangle\langle 1| + \mu_1 |1\rangle\langle 4| \qquad B_2 = |0\rangle\langle 2| + \mu_2 |2\rangle\langle 4| \qquad B_3 = |0\rangle\langle 3| + \mu_3 |3\rangle\langle 4| \\
B_4 &= |0\rangle\langle 4|
\end{aligned}$$

- Finding ground states of 1-D quantum Hamiltonians is therefore equivalent to the variational problem

$$\min_{|\Psi\rangle} \frac{\langle \Psi | T | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

- Note: Hamiltonians with exponentially decaying interactions of the form

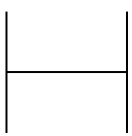
$$\mathcal{H} = \sum_{\alpha,i<j} \mu_\alpha \lambda_\alpha^{i-j} \sigma_\alpha^i \otimes \sigma_\alpha^j + \sum_j \hat{O}^j$$

still have exact simpe MPO description: just replace

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

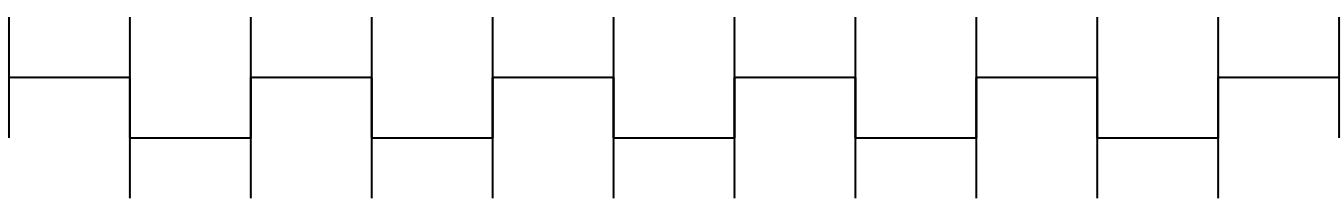
Even more MPO's: exponentials of local quantum Hamiltonians

- Exponential $\exp\left(\epsilon \sum_i Z_i Z_{i+1}\right)$ can also be represented as a MPO
  - Lies at the heart of all time-evolution algorithms for DMRG/MPS
- First step: exponential of local interaction as an MPO:



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

- Similar for any Hamiltonian: follows from singular value decomposition
- For case of TEBD using Trotter expansion, the evolution operator is hence a MPO:

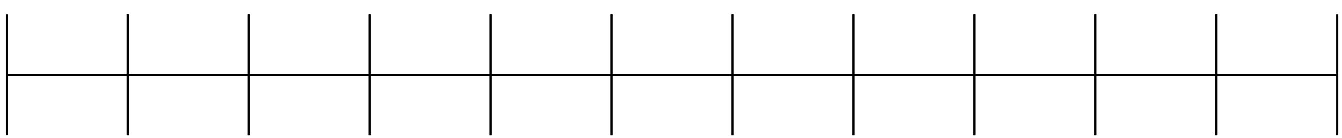


Exponentials of local commuting quantum Hamiltonians

- In case of Hamiltonian which is a sum of commuting terms: whole thing is one big MPO

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

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



Interludum: Matrix Product Operators and the Bethe ansatz:

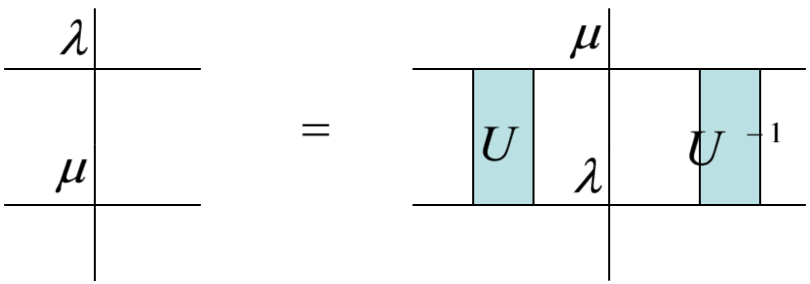
- Algebraic Bethe ansatz is all about MPO:

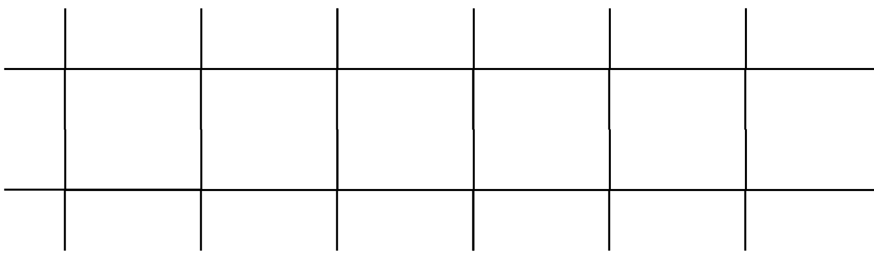
$$\tau(\lambda) = \sum_{i_1 i_2 i_3 \ldots} Tr\left\{A^{i_1} A^{i_2} A^{i_3} \ldots\right\} \; O_{i_1} \otimes O_{i_2} \otimes O_{i_3} \otimes \ldots$$

$$A^0 = \lambda . I \qquad A^1 = -\frac{i}{2} X \qquad \mathrm{A}^2 = -\frac{i}{2} Y \qquad \mathrm{A}^3 = -\frac{i}{2} Z$$

$$O^0 = I \qquad O^1 = X \qquad \mathrm{O}^2 = Y \qquad \mathrm{O}^3 = X$$

- Crucial Property of this family of MPO: they all commute (==Yang-Baxter equation):





$$U = (\lambda - \mu) I - ic\left(X \otimes X + Y \otimes Y + Z \otimes Z\right)$$

  – Gauge transformation of MPS/MPO leave it invariant!

- What has this to do with the Heisenberg model?

$$H_{heis} = 2i \frac{d}{d\lambda} \ln(\tau(\lambda))\Big|_{\lambda=-i/2}$$

  - This can easily be seen because $\tau(-i/2)$ is the shift operator (shifts qubits 1,2,3,…N to 2,3,4,…1); taking the derivative replaces one of those "swaps" with the idenity; logarithmic derivative undoes all the other swaps, leaving the Heisenberg Hamiltonian!
  - It follows that $[H_{heis}, \tau(\lambda)] = 0$ and hence they have the same eigenvectors

- Let's now define new operators similar to $\tau(\lambda)$ but with OBC:

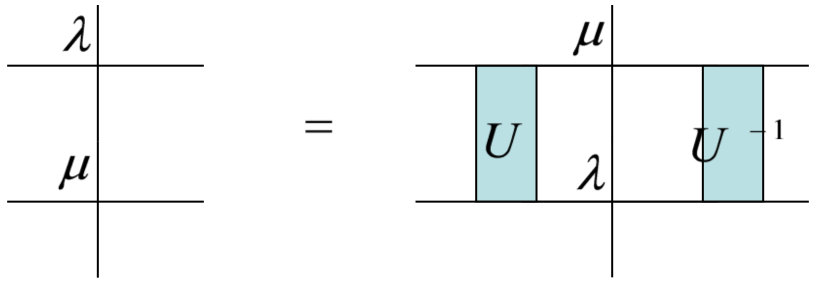
$$B(\lambda) = \sum_{i_1 i_2 i_3 \ldots} \langle 0 | A^{i_1} A^{i_2} A^{i_3} \ldots | 1 \rangle \quad O_{i_1} \otimes O_{i_2} \otimes O_{i_3} \otimes \ldots$$

  - These will play the role of creation operators and commute for all λ

- All eigenstates of the Heisenberg model are of the form

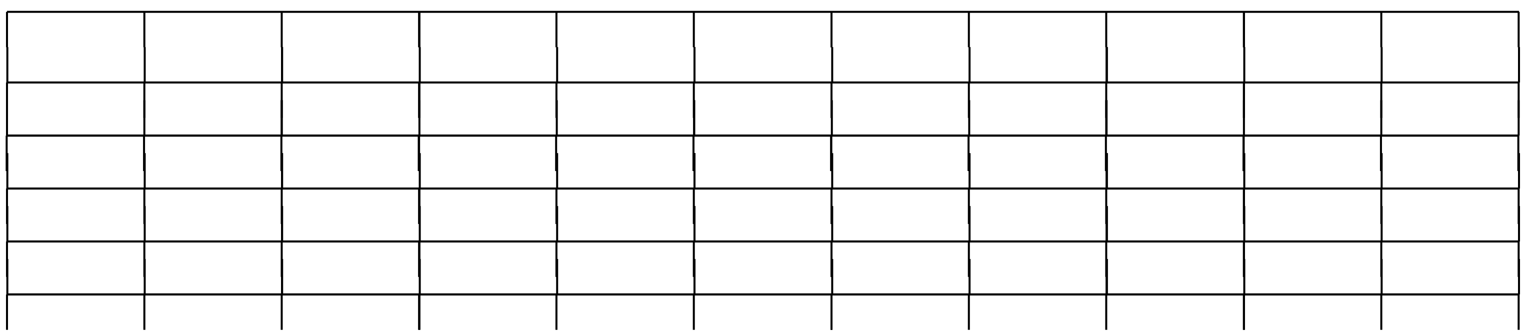
$$\left|\psi_{\lambda_1\lambda_2\lambda_3\ldots}\right\rangle = B(\lambda_1)B(\lambda_2)B(\lambda_3)\ldots\left|000\ldots\right\rangle \qquad \tau(\lambda)\left|000\ldots\right\rangle = f(\lambda)\left|000\ldots\right\rangle$$

  - The parameters $\{\lambda_i\}$ are found by imposing that these are eigenstates of $\tau(\lambda)$ = Bethe equations (follows simply from working out commutation relations; this leads to coupled equations between the $\{\lambda_i\}$)



$$U = (\lambda-\mu)I - i(X\otimes X + Y\otimes Y + Z\otimes Z)$$

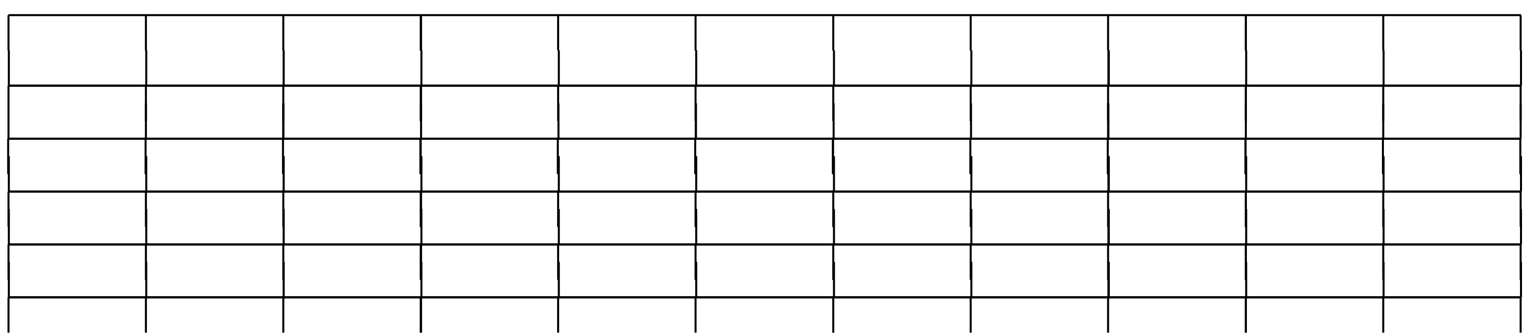
- In terms of MPS/MPO: all eigenstates can exactly be represented as



  - Note that the bond dimension increases exponentially with number of MPO's applied

Matrix product States

- Crucial insight: for gapped transfer matrices / Hamiltonians, a very good approximation to the extremal eigenvector will be obtained by subsequently applying the MPO to an arbitrary starting state (called power method in linear algebra)



  - Because of the algebra of MPO/MPS, this is just another MPS but with a larger bond dimension

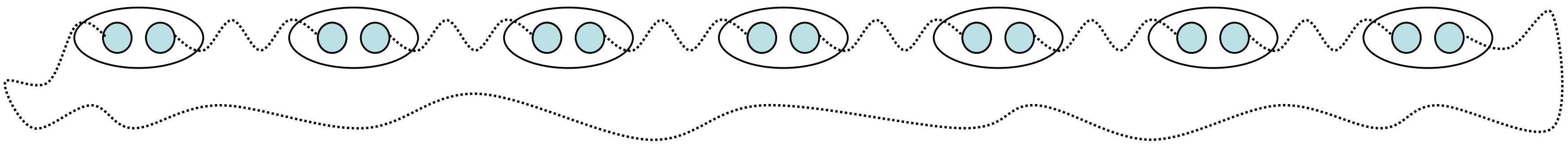
- Therefore MPS must capture all properties for representing the extremal eigenvectors!

- Therefore MPS seem to be natural candidates for variational wavefunctions of quantum Hamiltonians

  - Note: we can play the same game for systems with periodic boundary conditions, for systems in higher dimensions and for different systems involving e.g. fermions; this gives rise to DMRG with periodic boundary conditions, to 2-dimensional generalization of MPS, i.e. projected entangled pair states (PEPS), to the fermionic analogues of PEPS, …

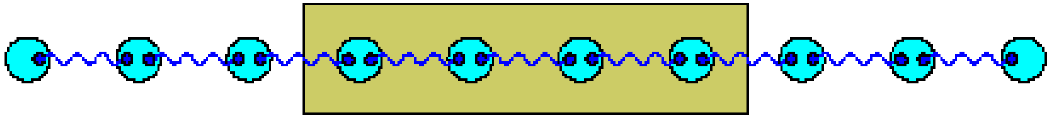
- Alternative justifications for the use of MPS:
  - Purifications of systems with finite correlation length
  - MPS represent optimal balance between strong local correlations and translational invariance
  - Area laws (even with logarithmic corrections) imply polynomial bond dimension for MPS (cfr Hastings)

Matrix Product States



- Valence bond picture: translational invariant by construction
- Has extremal local correlations
- Obeys area law by construction

- Theorem: if an area law is satisfied, then the state can be well approximated by a MPS:



$$S_\alpha\left(\rho_{1,2,\cdots,L}\right)\leq c\ln\left(L\right) \qquad \left\|\left|\psi_{ex}^{N}\right\rangle-\left|\psi_{D}^{N}\right\rangle\right\|\leq\varepsilon \qquad D_N\leq\frac{cst}{\varepsilon}N^{f(c)}$$

- In case of local gapped 1-D Hamiltonians: area law is guaranteed
- Conclusion: all states in finite 1-D chains can be represented by MPS: breakdown of exponential wall !

How to do the variation?

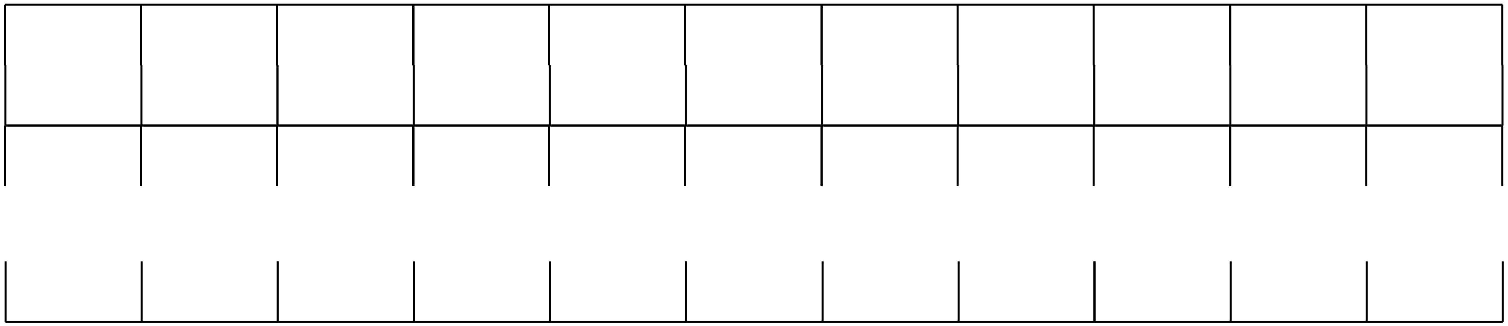
- Cost function for extremal eigenvalue is a multiquadratic problem on the variables of the MPS
  - Standard and pretty robust optimization method for solving such a problem is known as alternating least squares
  - Essentially equivalent to DMRG algorithm of White
  - Allows for simple generalization to e.g. PBC
- To make the algorithm better conditioned: exploit gauge degrees of freedom to orthonormalize vectors: denominator N becomes equal to the identity
  - Note: not possible to do this for PBC!

How to formulate time-evolution as a variational principle?

- Variational formulation of time evolution: variational dimensional reduction

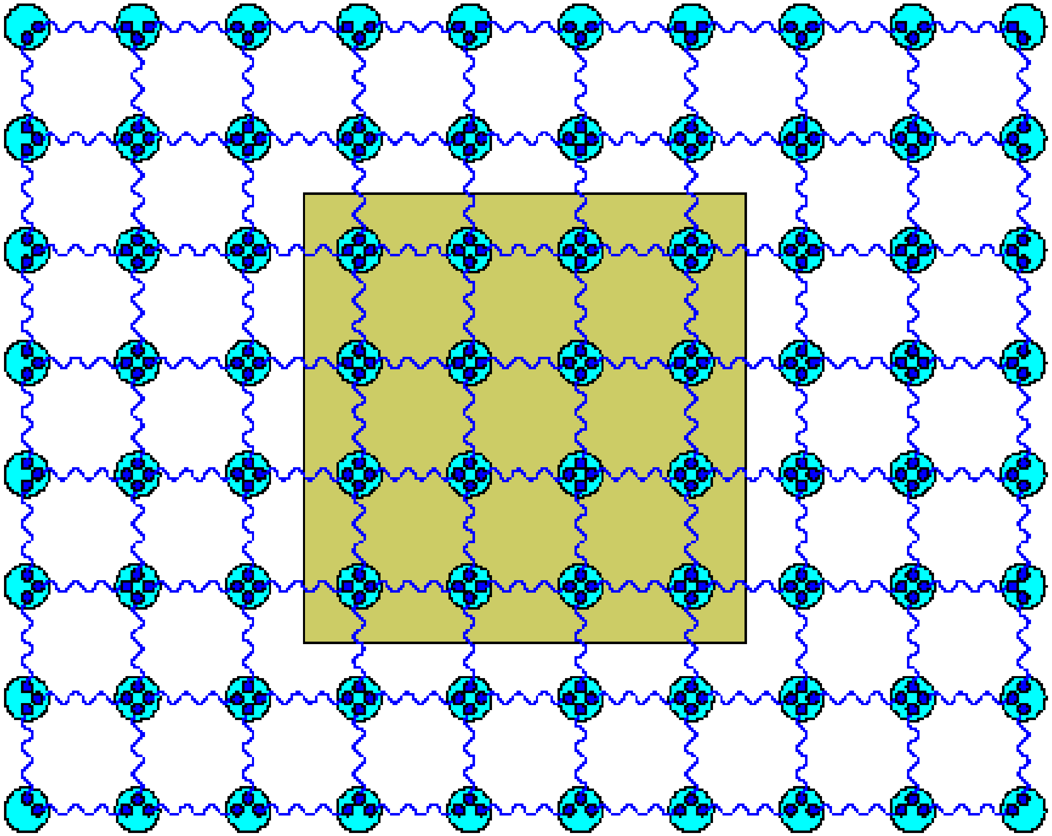
  given a MPS $|\chi\rangle$ and an MPO O, find the MPS $|\psi\rangle$ that minimizes

$$\min_{|\Psi\rangle} \left\| |\psi\rangle - O|\chi\rangle \right\|_2$$



  - It turns out that this is also a multiquadratic optimization problem that is very well conditioned and can be solved using DMRG-like sweeping!
  - Core method for simulating PEPS
  - The error in the approximation is known exactly!
  - Allows to do time evolution without breaking translational invariance

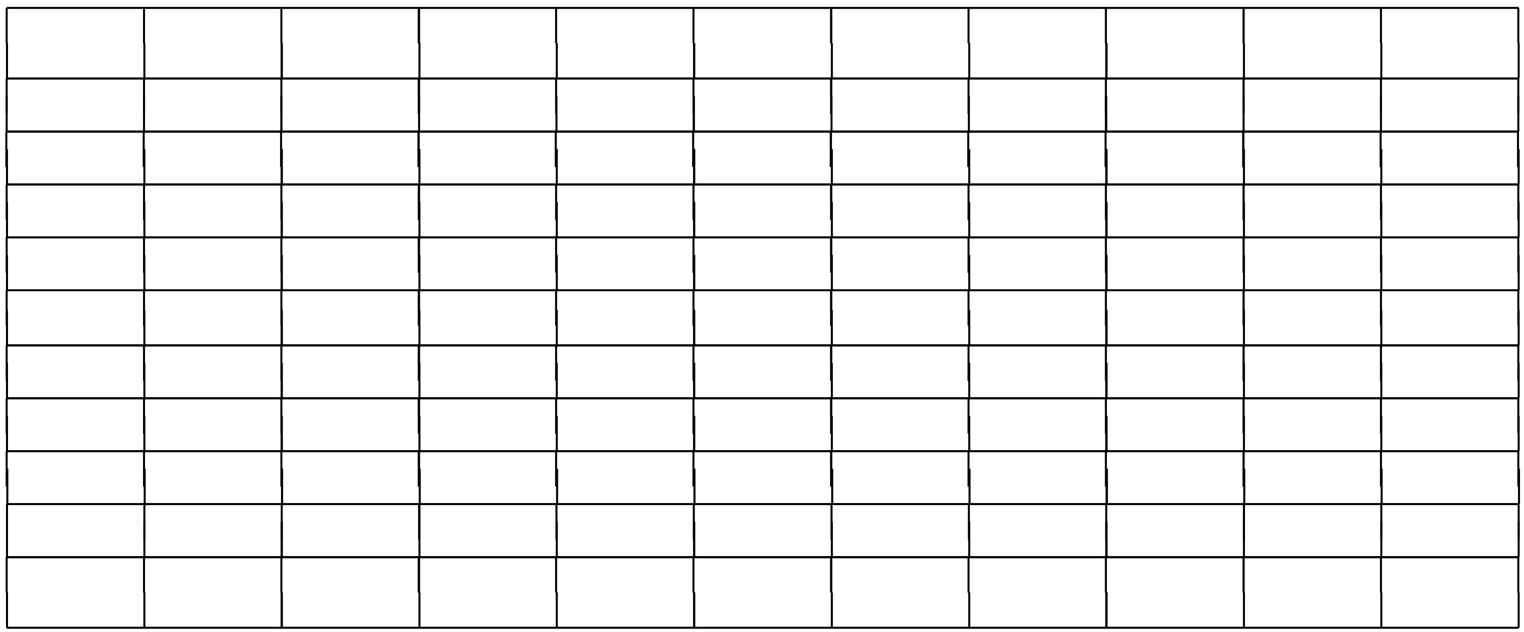
Generalizing MPS to higher dimensions: PEPS



- Area law is satisfied by construction : scalable!
- Precursors: AKLT, Nishino; PEPS introduced in context of measurement-based quantum computation

How to calculate expectation values?

- Equivalent to contracting tensor network consisting of MPS and MPO!



  - Obvious way of doing this: recursively use $\min_{|\Psi\rangle} \left\| |\psi\rangle - O|\chi\rangle \right\|_2$
  - Optimization: alternating least squares as in DMRG
    - Alternatively: imaginary time evolution ; infinite algorithm ; renormalization

Holographic principle: dimensional reduction

- Crucial property of MPS/PEPS: dimensional reduction
  - Start from quantum system in 2 dimensions (2+1)
  - The PEPS ansatz maps the quantum Hamiltonian to a state corresponding to a partition function in 2 dimensions (2+0)
  - The properties of such a state are described by a (1+1) dimensional theory (eigenvectors of transfer matrices)
  - Those eigenvectors are well described by MPS
  - Properties of MPS are trivial to calculate: reduction to a partition function of a 1-D system (1+0)