Large Eigenvalue Problems: Computation of Ground States

Konrad Waldherr

June 2010

Joint work with Thomas Huckle
Overview

1. Problem Setting: Computation of Ground States
2. Efficient Representations of Vectors
   - Alternating Least Squares
   - Matrix Product States
   - Generalizations of Classical Decomposition Schemes
3. Numerical Results
4. Conclusions / Outlook
Problem setting: computation of ground states

Given:

- Physical system with $N$ particles (e.g. 1D spin chain)

\[1 \quad \ldots \quad 2 \quad \ldots \quad 3 \quad \ldots \quad 4 \quad \ldots \quad 5\]
Problem setting: computation of ground states

Given:

- Physical system with $N$ particles (e.g. 1D spin chain)
  ![Diagram of 1D spin chain with $N$ particles]

- Interaction within the system (e.g. nearest-neighbour interaction)
  ![Diagram of nearest-neighbour interaction]

Goal:
- Find ground state, i.e. the state related to the smallest energy of the system.
Problem setting: computation of ground states

Given:

- Physical system with \( N \) particles (e.g. 1D spin chain)

- Interaction within the system (e.g. nearest-neighbour interaction)

- External interaction (e.g. exterior magnetic field)
Problem setting: computation of ground states

Given:
- Physical system with $N$ particles (e.g. 1D spin chain)
- Interaction within the system (e.g. nearest-neighbour interaction)
- External interaction (e.g. exterior magnetic field)

Goal:
- Find ground state, i.e. the state related to the smallest energy of the system.
Mathematical Model

- Any state of the system is represented by a vector $x \in \mathbb{C}^{2^N}$.
- The physical system can be described by the Hamiltonian $H \in \mathbb{C}^{2^N \times 2^N}$.
Mathematical Model

- Any state of the system is represented by a vector \( x \in \mathbb{C}^{2^N} \).
- The physical system can be described by the Hamiltonian \( H \in \mathbb{C}^{2^N \times 2^N} \).
- The Hamiltonian may be formulated as a weighted sum of Kronecker products of Pauli and identity matrices.
- Pauli matrices:
  \[
  \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
  \]
- Kronecker product:
  \[
  A \otimes B := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}
  \]
Mathematical Model

- Any state of the system is represented by a vector \( x \in \mathbb{C}^{2^N} \).
- The physical system can be described by the Hamiltonian \( H \in \mathbb{C}^{2^N \times 2^N} \).
- The Hamiltonian may be formulated as a weighted sum of Kronecker products of Pauli and identity matrices.
- Pauli matrices:
  \[
  \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
  \]
- Kronecker product:
  \[
  A \otimes B := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}
  \]
- Then, the ground state corresponds to the eigenvector related to the smallest eigenvalue.
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I \]

\[ + I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I + I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I + I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I \]

\[ + I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I \]

\[ + I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I \]

\[ + I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z \]

\[ + I \otimes I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I + I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I + I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I + I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z + \sigma_x \otimes I \otimes I \otimes I \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I + I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I + I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I + I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z + \sigma_x \otimes I \otimes I \otimes I \otimes I + I \otimes \sigma_x \otimes I \otimes I \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I + \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I + \sigma_z \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I + \sigma_z \otimes \sigma_z \otimes \sigma_z \otimes \sigma_z \otimes I + \sigma_x \otimes I \otimes I \otimes I \otimes I + \sigma_x \otimes I \otimes \sigma_x \otimes I \otimes I + \sigma_x \otimes \sigma_x \otimes I \otimes I \otimes I + \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes I \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I \]

\[ + \; I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I \]

\[ + \; I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I \]

\[ + \; I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z \]

\[ + \; \sigma_x \otimes I \otimes I \otimes I \otimes I \]

\[ + \; I \otimes \sigma_x \otimes I \otimes I \otimes I \]

\[ + \; I \otimes I \otimes \sigma_x \otimes I \otimes I \]

\[ + \; I \otimes I \otimes I \otimes \sigma_x \otimes I \]

\[ + \; I \otimes I \otimes I \otimes I \otimes \sigma_x \otimes I \]
Example: Ising-type Hamiltonian

\[ H = \sigma_z \otimes \sigma_z \otimes I \otimes I \otimes I 
+ I \otimes \sigma_z \otimes \sigma_z \otimes I \otimes I 
+ I \otimes I \otimes \sigma_z \otimes \sigma_z \otimes I 
+ I \otimes I \otimes I \otimes \sigma_z \otimes \sigma_z 
+ \sigma_x \otimes I \otimes I \otimes I \otimes I 
+ I \otimes \sigma_x \otimes I \otimes I \otimes I 
+ I \otimes I \otimes \sigma_x \otimes I \otimes I 
+ I \otimes I \otimes I \otimes \sigma_x \otimes I 
+ I \otimes I \otimes I \otimes I \otimes \sigma_x \]
Sparsity Pattern of the Ising-type Hamiltonian
Problem Setting

- General formulation of the Hamiltonian $H$:

$$H = \sum_{k=1}^{M} \alpha_k Q_1^{(k)} \otimes \cdots \otimes Q_N^{(k)}, \quad Q_i^{(k)} \in \{id, \sigma_x, \sigma_y, \sigma_z\} \subset \mathbb{C}^{2 \times 2}$$
Problem Setting

- General formulation of the Hamiltonian $H$:

$$H = \sum_{k=1}^{M} \alpha_k Q_1^{(k)} \otimes \ldots \otimes Q_N^{(k)}, \quad Q_i^{(k)} \in \{id, \sigma_x, \sigma_y, \sigma_z\} \subset \mathbb{C}^{2\times2}$$

- Variational ansatz: Minimization of the Rayleigh quotient:

$$\min_{x \neq 0} \frac{x^\dagger H x}{x^\dagger x}$$
**Problem Setting**

- General formulation of the Hamiltonian $H$:

\[
H = \sum_{k=1}^{M} \alpha_k Q^{(k)}_1 \otimes \cdots \otimes Q^{(k)}_N, \quad Q^{(k)}_i \in \{id, \sigma_x, \sigma_y, \sigma_z\} \subset \mathbb{C}^{2 \times 2}
\]

- Variational ansatz: Minimization of the Rayleigh quotient:

\[
\min_{x \neq 0} \frac{x^\dagger Hx}{x^\dagger x}
\]

- Problem: The vector space $\mathcal{V} = \mathbb{C}^{2^N}$ grows exponentially in $N$.

- Idea: Choose appropriate subset $\mathcal{U} \subset \mathcal{V}$ and consider

\[
\min_{x \in \mathcal{U}} \frac{x^\dagger Hx}{x^\dagger x}
\]
Tensor decompositions

- Any state $x$ may be considered as a $N$th order tensor:

$$x = (x_i)_{i=1,...,2^N} = (x_{i_1,i_2,...,i_N})_{i_j=0,1}.$$

- Goal: Find convenient tensor decompositions
- This decomposition should allow
  - polynomial representations,
  - efficient computations of $Hx$ and $y^\dagger x$,
  - good approximations of the eigenvector.

Decomposition schemes:
- Tucker decomposition,
- Canonical decomposition,
- Matrix Product States (MPS).
Tensor decompositions

- Any state $x$ may be considered as a $N$th order tensor:

$$x = (x_i)_{i=1,...,2^N} = (x_{i_1,i_2,...,i_N})_{i_j=0,1}.$$ 

- Goal: Find convenient tensor decompositions
- This decomposition should allow
  - polynomial representations,
  - efficient computations of $Hx$ and $y^\dagger x$,
  - good approximations of the eigenvector.

- Decomposition schemes:
  - Tucker decomposition,
  - Canonical decomposition,
  - Matrix Product States (MPS).
Alternating Least Squares

- Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).
- Minimize the Rayleigh quotient

\[
\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
\]
Alternating Least Squares

- Suppose $x = x_1 \otimes x_2 \otimes \cdots \otimes x_q$.
- Minimize the Rayleigh quotient

$$\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}$$

- Suppose $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_q$ to be fixed and find optimal $x_i$:

$$\min_{x_i} \frac{(x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots x_i \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)} = \min_{x_i} \frac{x_i^H H_i x_i}{x_i^H N_i x_i}$$
Alternating Least Squares

- Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).
- Minimize the Rayleigh quotient

\[
\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
\]

- Suppose \( x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_q \) to be fixed and find optimal \( x_i \):

\[
\min_{x_i} \frac{(x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_q)} = \min_{x_i} \frac{x_i^H H_i x_i}{x_i^H N_i x_i}
\]

- Leads to a generalized eigenvalue problem in the size of \( x_i \).
Alternating Least Squares

• Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).

• Minimize the Rayleigh quotient

\[
\min_{x} \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
\]

• Suppose \( x_2, x_3, \ldots, x_q \) to be fixed and find optimal \( x_1 \):

\[
\min_{x_1} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_1} \frac{x_1^H H_1 x_1}{x_1^H N_1 x_1}
\]

• Leads to a generalized eigenvalue problem in the size of \( x_1 \).
Alternating Least Squares

- Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).
- Minimize the Rayleigh quotient
  \[
  \min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
  \]
- Suppose \( x_1, x_3, \ldots, x_q \) to be fixed and find optimal \( x_2 \):
  \[
  \min_{x_2} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_2} \frac{x_2^H H_2 x_2}{x_2^H N_2 x_2}
  \]
- Leads to a generalized eigenvalue problem in the size of \( x_2 \).
Alternating Least Squares

- Suppose $x = x_1 \otimes x_2 \otimes \cdots \otimes x_q$.
- Minimize the Rayleigh quotient

$$\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}$$

- Suppose $x_1, x_2, x_4, \ldots, x_q$ to be fixed and find optimal $x_3$:

$$\min_{x_3} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}$$

$$= \min_{x_3} \frac{x_3^H H_3 x_3}{x_3^H N_3 x_3}$$

- Leads to a generalized eigenvalue problem in the size of $x_3$. 
Alternating Least Squares

- Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).
- Minimize the Rayleigh quotient

\[
\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
\]

- Suppose \( x_1, x_2, \ldots, x_{q-1} \) to be fixed and find optimal \( x_q \):

\[
\min_{x_q} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_q} \frac{x_q^H H_q x_q}{x_q^H N_q x_q}
\]

- Leads to a generalized eigenvalue problem in the size of \( x_q \).
Alternating Least Squares

- Suppose \( x = x_1 \otimes x_2 \otimes \cdots \otimes x_q \).
- Minimize the Rayleigh quotient

\[
\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}
\]

- Suppose \( x_2, x_3, \ldots, x_q \) to be fixed and find optimal \( x_1 \):

\[
\min_{x_1} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_1} \frac{x_1^H H_1 x_1}{x_1^H N_1 x_1}
\]

- Leads to a generalized eigenvalue problem in the size of \( x_1 \).
Alternating Least Squares

- Suppose $x = x_1 \otimes x_2 \otimes \cdots \otimes x_q$.
- Minimize the Rayleigh quotient

$$\min_x \frac{x^\text{H} H x}{x^\text{H} x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^\text{H} H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^\text{H} (x_1 \otimes \cdots \otimes x_q)}$$

- Suppose $x_2, x_3, \ldots, x_q$ to be fixed and find optimal $x_1$:

$$\min_{x_1} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^\text{H} H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^\text{H} (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_1} \frac{x_1^\text{H} H_1 x_1}{x_1^\text{H} N_1 x_1}$$

- Leads to a generalized eigenvalue problem in the size of $x_1$. 
Alternating Least Squares

- Suppose $x = x_1 \otimes x_2 \otimes \cdots \otimes x_q$.
- Minimize the Rayleigh quotient

$$\min_x \frac{x^H H x}{x^H x} = \min_{x_1, \ldots, x_q} \frac{(x_1 \otimes \cdots \otimes x_q)^H H (x_1 \otimes \cdots \otimes x_q)}{(x_1 \otimes \cdots \otimes x_q)^H (x_1 \otimes \cdots \otimes x_q)}$$

- Suppose $x_2, x_3, \ldots, x_q$ to be fixed and find optimal $x_1$:

$$\min_{x_1} \frac{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)^H (x_1 \otimes x_2 \otimes x_3 \otimes \cdots \otimes x_q)} = \min_{x_1} \frac{x_1^H H_1 x_1}{x_1^H N_1 x_1}$$

- Leads to a generalized eigenvalue problem in the size of $x_1$.
- Iterate until convergence.
Matrix Product States (MPS)

- In MPS, the vector components are given by
  \[ x_i = x_{i_1}, \ldots, i_N = \text{trace} \left( A_{i_1}^{(i_1)} \cdot A_{i_2}^{(i_2)} \cdots A_{i_N}^{(i_N)} \right) \]
  
  with matrices \( A_i^{(0)}, A_i^{(1)} \in \mathbb{C}^{D_i \times D_{i+1}} \)

- Thus, the vector \( x \) may be written as
  \[ x = \sum_{i=1}^{2^N} x_i e_i = \sum_{i_1, \ldots, i_p} \text{trace} \left( A_1^{(i_1)} \cdot A_2^{(i_2)} \cdots A_p^{(i_N)} \right) e_{i_1, \ldots, i_N} \]
  
  \[ = x_i = x_{i_1}, \ldots, i_N \]
  
  \[ = e_i \]
Matrix Product States (MPS)

- In MPS, the vector components are given by
  \[ x_i = x_{i_1,...,i_N} = \text{trace} \left( A_1^{(i_1)} \cdot A_2^{(i_2)} \cdots A_N^{(i_N)} \right) \]
  with matrices \( A_i^{(0)}, A_i^{(1)} \in \mathbb{C}^{D_i \times D_{i+1}} \)
- Thus, the vector \( x \) may be written as
  \[ x = \sum_{i=1}^{2^N} x_i e_i = \sum_{i_1,...,i_p} \text{trace} \left( A_1^{(i_1)} \cdot A_2^{(i_2)} \cdots A_p^{(i_N)} \right) e_{i_1,...,i_N} = x_{i_1,...,i_N} e_i \]
- The MPS formalism leads to
  \[ x = \sum_{m_1,...,m_p} a_{1,m_1,m_2} \otimes \cdots \otimes a_{p,m_p,m_1} \]

\[ \Rightarrow \text{Efficient computation of } Hx = \left( \sum_{k=1}^{M} \alpha_k Q_1^{(k)} \otimes \cdots \otimes Q_N^{(k)} \right) x. \]
How to compute the inner product efficiently?

\[ x_{MPS}^H y_{MPS} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \bar{A}_{i_1}^{(i_1)} \cdots \bar{A}_{i_N}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B_{i_1}^{(i_1)} \cdots B_{i_p}^{(i_p)} \right) \right) \]

\[ = \sum_{i_1, \ldots, i_N} \sum_{k_1, \ldots, k_N} \bar{a}_{i_1, k_1, k_2}^{(i_1)} \cdots \bar{a}_{i_N, k_N, k_1}^{(i_N)} \sum_{m_1, \ldots, m_N} b_{i_1, m_1, m_2}^{(i_1)} \cdots b_{i_N, m_N, m_1}^{(i_N)} \]
How to compute the inner product efficiently?

\[ x^H_{\text{MPS}} y_{\text{MPS}} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \bar{A}^{(i_1)} \cdots \bar{A}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B_{i_1}^{(i_1)} \cdots B_{i_P}^{(i_P)} \right) \right) \]

\[ = \sum_{i_1, \ldots, i_N} \sum_{k_1, \ldots, k_N} \bar{a}_{i_1,k_1,k_2}^{(i_1)} \cdots \bar{a}_{i_N,k_N,k_1}^{(i_N)} \sum_{m_1, \ldots, m_N} b_{i_1,m_1,m_2}^{(i_1)} \cdots b_{i_P,m_N,m_1}^{(i_P)} \]

Find an efficient ordering of the summations:
How to compute the inner product efficiently?

\[ x_{\text{MPS}}^H y_{\text{MPS}} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \tilde{A}^{(i_1)} \cdots \tilde{A}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B^{(i_1)} \cdots B^{(i_p)} \right) \right) \]

\[ = \sum_{i_1, \ldots, i_N} \sum_{k_1, \ldots, k_N} \tilde{a}^{(i_1)}_{1,k_1,k_2} \cdots \tilde{a}^{(i_N)}_{N,k_N,k_1} \sum_{m_1, \ldots, m_N} b^{(i_1)}_{1,m_1,m_2} \cdots b^{(i_N)}_{p,m_N,m_1} \]

Find an efficient ordering of the summations:
How to compute the inner product efficiently?

\[ x_{MPS}^H y_{MPS} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \bar{A}_{i_1}^{(i_1)} \cdots \bar{A}_{i_N}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B_{i_1}^{(i_1)} \cdots B_{i_p}^{(i_p)} \right) \right) \]

\[ = \sum_{i_1, \ldots, i_N} \sum_{k_1, \ldots, k_N} \bar{a}_{i_1,k_1,k_2}^{(i_1)} \cdots \bar{a}_{i_N,k_N,k_1}^{(i_N)} \sum_{m_1, \ldots, m_N} b_{i_1,m_1,m_2}^{(i_1)} \cdots b_{i_p,m_N,m_1}^{(i_p)} \]

Find an efficient ordering of the summations:
How to compute the inner product efficiently?

\[ X_{MPS}^H Y_{MPS} = \sum_{i_1,\ldots,i_N} \left( \text{trace} \left( \tilde{A}_1^{(i_1)} \cdots \tilde{A}_N^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B_1^{(i_1)} \cdots B_{p}^{(i_p)} \right) \right) \]

\[ = \sum_{i_1,\ldots,i_N} \sum_{k_1,\ldots,k_N} \tilde{a}_1^{(i_1)} \cdots \tilde{a}_N^{(i_N)} \sum_{m_1,\ldots,m_N} b_1^{(i_1)} \cdots b_{p}^{(i_p)} \]

Find an efficient ordering of the summations:
How to compute the inner product efficiently?

\[ x_{MPS}^H y_{MPS} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \bar{A}^{(i_1)} \ldots \bar{A}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B^{(i_1)} \ldots B^{(i_p)} \right) \right) = \sum_{i_1, \ldots, i_N} \sum_{k_1, \ldots, k_N} \bar{a}^{(i_1)}_{k_1, k_2} \ldots \bar{a}^{(i_N)}_{N, k_N, k_1} \sum_{m_1, \ldots, m_N} b^{(i_1)}_{1, m_1, m_2} \ldots b^{(i_N)}_{p, m_N, m_1} \]

Find an efficient ordering of the summations:
How to compute the inner product efficiently?

\[ x_{MPS}^H y_{MPS} = \sum_{i_1, \ldots, i_N} \left( \text{trace} \left( \tilde{A}_{i_1}^{(i_1)} \cdots \tilde{A}_{i_N}^{(i_N)} \right) \right) \cdot \left( \text{trace} \left( B_1^{(i_1)} \cdots B_p^{(i_p)} \right) \right) \]

Find an efficient ordering of the summations:
Uniqueness of MPS

• The MPS formalism

\[ x = \sum_{i_1, \ldots, i_N} \text{trace} \left( A_{i_1}^{(i_1)} \cdot A_{i_2}^{(i_2)} \cdots A_{i_N}^{(i_N)} \right) e_{i_1, \ldots, i_N} \]

is not unique:

\[ A_j^{(i_j)} \rightarrow M_j^{-1} A_j^{(i_j)} M_{j+1} \]
Uniqueness of MPS

- The MPS formalism

\[ x = \sum_{i_1, \ldots, i_N} \text{trace} \left( A_{i_1}^{(1)} \cdot A_{i_2}^{(2)} \cdots A_{i_N}^{(N)} \right) e_{i_1, \ldots, i_N} \]

is not unique:

\[ A_j^{(i_j)} \rightarrow M^{-1}_j A_j^{(i_j)} M_{j+1} \]

- Replace the matrices \( A_j^{(0)}, A_j^{(1)} \) by parts of unitary matrices using the SVD:

\[
\begin{pmatrix}
A_j^{(0)} \\
A_j^{(1)}
\end{pmatrix} =
\begin{pmatrix}
U_j^{(0)} \\
U_j^{(1)}
\end{pmatrix} \Sigma_j V_j.
\]
Uniqueness of MPS

- The MPS formalism

\[ x = \sum_{i_1, \ldots, i_N} \text{trace} \left( A_{i_1}^{(i_1)} \cdot A_{i_2}^{(i_2)} \cdots A_{i_N}^{(i_N)} \right) e_{i_1, \ldots, i_N} \]

is not unique:

\[ A_{j}^{(i_j)} \rightarrow M_j^{-1} A_{j}^{(i_j)} M_{j+1} \]

- Replace the matrices \( A_j^{(0)}, A_j^{(1)} \) by parts of unitary matrices using the SVD:

\[
\begin{pmatrix}
A_j^{(0)} \\
A_j^{(1)}
\end{pmatrix} = 
\begin{pmatrix}
U_j^{(0)} \\
U_j^{(1)}
\end{pmatrix} \Sigma_j V_j.
\]

- Multiply the \( \Sigma_j V_j \) parts on the right neighbour.
Uniqueness of MPS

• The MPS formalism

\[ x = \sum_{i_1, \ldots, i_N} \text{trace} \left( A_{i_1}^{(i_1)} \cdot A_{i_2}^{(i_2)} \cdots A_{i_N}^{(i_N)} \right) e_{i_1, \ldots, i_N} \]

is not unique:

\[ A_{ij}^{(i_j)} \longrightarrow M_j^{-1} A_{ij}^{(i_j)} M_{j+1} \]

• Replace the matrices \( A_j^{(0)}, A_j^{(1)} \) by parts of unitary matrices using the SVD:

\[
\begin{pmatrix}
A_j^{(0)} \\
A_j^{(1)}
\end{pmatrix} =
\begin{pmatrix}
U_j^{(0)} \\
U_j^{(1)}
\end{pmatrix}
\Sigma_j V_j.
\]

• Multiply the \( \Sigma_j V_j \) parts on the right neighbour.
• Replace all \( A \) matrices by \( U \) matrices, up to one.
• This remaining term will be optimized via ALS.
Minimization Algorithm in terms of MPS

- Start with initial guesses for the $A_j^{(ij)}$ matrices.
- Replace $A_j^{ij}$ by $U_j^{(ij)}$ up to $A_1^{(i_1)}$.
- Consider all $U$ matrices as fixed and optimize $X_1 = A_1$.
- Put this ansatz in the Rayleigh quotient

$$\min_{X_1} \left( \sum_{{ij}} \text{trace} \left( X_1^{(i_1)} \cdot A_2^{(i_2)} \cdots A_N^{(i_N)} \right) e_{i_1}, \ldots, e_{i_N} \right) \quad \text{with effective Hamiltonian } R_1$$

$$= \min_{X_1} \frac{X_1 \dagger R_1 X_1}{X_1 \dagger X_1}$$

with effective Hamiltonian $R_1$

- Leads to a dense eigenvalue problem for the $(2D^2 \times 2D^2)$ matrix $R_1$.
- Repeat in the ALS sense.
Decompositions of a tensor

- Tucker decomposition:

\[ A = U S V \]

Not advantageous in our case \( (n_i = 2) \)
Decompositions of a tensor

- Tucker decomposition:

\[ A = UV \]

Not advantageous in our case \((n_i = 2)\)

- **Canonical decomposition (Parallel Factorization):**

\[ A = \sum_{i=1}^{r} U_i V_i^T + \ldots \]

- Sum of rank-one tensors
- The following vector ansatz is based on this canonical decomposition scheme.
Tensor Product Ansatz for the Eigenvector

- The tensor product structure of $H$, i.e.

$$H = \sum_{k=1}^{M} \alpha_k Q_1^{(k)} \otimes \cdots \otimes Q_N^k$$

suggests some kind of tensor product ansatz for eigenvector $x$:

$$x = x_1 \otimes x_2 \otimes \cdots \otimes x_q$$

- This ansatz allows easy computation of $Hx$ and $y^\dagger x$
- Use Rayleigh quotient minimization

$$\min_{x_1,\ldots,x_q} \frac{(x_1 \otimes x_2 \otimes \cdots \otimes x_q)^\dagger H (x_1 \otimes x_2 \otimes \cdots \otimes x_q)}{(x_1 \otimes x_2 \otimes \cdots \otimes x_q)^\dagger (x_1 \otimes x_2 \otimes \cdots \otimes x_q)}.$$

- Use ALS for finding optimal blocks $x_i$. 
Minimizing the Rayleigh Quotient Using ALS

- Suppose $x_L, x_R$ to be fixed and optimize $x_i$:

$$
\min_{x_i} \frac{(x_L \otimes x_i \otimes x_R)^\dagger \left( \sum_{k=1}^M \alpha_k H_L^{(k)} \otimes H_i^{(k)} \otimes H_R^{(k)} \right) (x_L \otimes x_i \otimes x_R)}{(x_L \otimes x_i \otimes x_R)^\dagger (x_L \otimes x_i \otimes x_R)}
$$

$$
= \min_{x_i} \frac{\sum_{k=1}^M \alpha_k \left( (x_L^\dagger H_L^{(k)} x_L) (x_R^\dagger H_R^{(k)} x_R) x_i^\dagger H_i^{(k)} x_i \right)}{\left( (x_L^\dagger x_L) (x_R^\dagger x_R) x_i^\dagger x_i \right)}
$$

$$
= \min_{x_i} \frac{\sum_{k=1}^M \alpha_k \left( \beta_k x_i^\dagger H_i^{(k)} x_i \right)}{\gamma x_i^\dagger x_i} = \min_{x_i} \frac{x_i^\dagger \left( \sum_{k=1}^M \frac{\alpha_k \beta_k}{\gamma} H_i^{(k)} \right) x_i}{x_i^\dagger x_i}.
$$

- Leads to a standard eigenvalue problem $\min_{x_i} \frac{x_i^\dagger H_i x_i}{x_i^\dagger x_i}$. 

K. Waldherr: Large Eigenvalue Problems: Computation of Ground States
PMAA 2010, Basel, June 2010
Tensor Product Ansatz for Eigenvector

- Consider $x = x_1, \ldots, x_q + \sum_{j=1}^J y_1^{(j)} \otimes \cdots \otimes y_q^{(j)}$
- Determine optimal $x_1, \ldots, x_q$ (all $y_i^{(j)}$ assumed to be fixed)
- Apply ALS technique:

$$\min_{x_i \neq 0} \frac{\left( x_L \otimes x_i \otimes x_R + \sum_{j=1}^J y_L^{(j)} \otimes y_i^{(j)} \otimes y_R^{(j)} \right)^\dagger H \left( x_L \otimes x_i \otimes x_R + \sum_{j=1}^J y_L^{(j)} \otimes y_i^{(j)} \otimes y_R^{(j)} \right)}{\left( x_L \otimes x_i \otimes x_R + \sum_{j=1}^J y_L^{(j)} \otimes y_i^{(j)} \otimes y_R^{(j)} \right)^\dagger} \left( x_L \otimes x_i \otimes x_R + \sum_{j=1}^J y_L^{(j)} \otimes y_i^{(j)} \otimes y_R^{(j)} \right)$$

$$= \min_{x_i \neq 0} \frac{x_i^\dagger H_i x_i + u^\dagger x_i + x_i^\dagger u + \beta}{x_i^\dagger x_i + v^\dagger x_i + x_i^\dagger v + \rho} = \min_{x_i \neq 0} \frac{\begin{pmatrix} x_i^\dagger & 1 \end{pmatrix} \begin{pmatrix} H_i & u \\ u^\dagger & \beta \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix}}{\begin{pmatrix} x_i^\dagger & 1 \end{pmatrix} \begin{pmatrix} I & v \\ v^\dagger & \rho \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix}}$$

- Leads to generalized eigenvalue problem

$$\begin{pmatrix} H_i & u \\ u^\dagger & \beta \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \begin{pmatrix} I & v \\ v^\dagger & \rho \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$
Tensor product ansatz for vectors

- The matrix on the right can be factorized:

$$
\begin{pmatrix}
I & v \\
v^\dagger & \rho
\end{pmatrix}
= \begin{pmatrix}
I & 0 \\
v^\dagger & \sqrt{\rho - v^\dagger v}
\end{pmatrix}
\begin{pmatrix}
I & v \\
0 & \sqrt{\rho - v^\dagger v}
\end{pmatrix}
$$

- Ansatz leads to a standard eigenvalue problem:

$$
\begin{pmatrix}
I & 0 \\
-v^\dagger & 1
\end{pmatrix}
\begin{pmatrix}
H_i & u \\
u^\dagger & \beta
\end{pmatrix}
\begin{pmatrix}
I & -v \\
0 & \sqrt{\rho - v^\dagger v}
\end{pmatrix}
y = \lambda y,
$$

$$
X = \begin{pmatrix}
I & -v \\
0 & \sqrt{\rho - v^\dagger v}
\end{pmatrix}
y.
$$

- $u, v, \beta, \rho$ are computed explicitly, the matrix $H_i$ is only applied implicitly.
Numerical results: MPS vs. ParaFac

Figure: Approximation error in the computation of the smallest eigenvalue of an Ising-type Hamiltonian of size $2^9 \times 2^9$.  

Figure: Approximation error in the computation of the smallest eigenvalue of an Ising-type Hamiltonian of size $2^{10} \times 2^{10}$.  

K. Waldherr: Large Eigenvalue Problems: Computation of Ground States
Problems with ALS

- $N = 100$ particles, Ising-type Hamiltonian
- Pattern $[15,15,15,15,15,15,10]$
- $J = 30$ addends
Problems with ALS

- Problem: Iteration gets stuck into local minima
- Resorts:
  - Better choice for initial guesses
  - Simultaneous optimization

$$H_i = \sum_{k=1}^{M} \alpha_k H^i(k).$$
Problems with ALS

- Problem: Iteration gets stuck into local minima
- Resorts:
  - Better choice for initial guesses
  - Simultaneous optimization
- Improvement of the initial guesses
  - Apply optimization routines (e.g. linesearch)
  - Choose orthogonal basis given by Arnoldi vectors of the simplified effective Hamiltonian

\[ H_i = \sum_{k=1}^{M} \alpha_k H_i^{(k)}. \]
Simultaneous Optimization

- Consider \( x = (x_1 \otimes x_2 \otimes \cdots \otimes x_q) + (y_1 \otimes y_2 \otimes \cdots \otimes y_q) \)
- Optimize \( x_i \) and \( y_i \) simultaneously, assume \( x_L, y_L, x_R, y_R \) to be fixed

\[
\min_{x_i, y_i \neq 0} \quad \frac{(x_L \otimes x_i \otimes x_R + y_L \otimes y_i \otimes y_R)^\dagger H (x_L \otimes x_i \otimes x_R + y_L \otimes y_i \otimes y_R)}{(x_L \otimes x_i \otimes x_R + y_L \otimes y_i \otimes y_R)^\dagger (x_L \otimes x_i \otimes x_R + y_L \otimes y_i \otimes y_R)}
\]

\[
= \min_{x_i, y_i \neq 0} \quad \frac{x_i^\dagger H_{i}^{xx} x_i + x_i^\dagger H_{i}^{xy} y_i + y_i^\dagger H_{i}^{yx} x_i + y_i^\dagger H_{i}^{yy} y_i}{\beta_{xx} x_i^\dagger x_i + \beta_{xy} x_i^\dagger y_i + \beta_{yx} y_i^\dagger x_i + \beta_{yy} y_i^\dagger y_i}
\]

\[
= \min_{x_i, y_i \neq 0} \quad \frac{(x_i^\dagger \quad y_i^\dagger)}{(\beta_{xx} I \quad \beta_{xy} I \quad \beta_{yx} I \quad \beta_{yy} I)} \begin{pmatrix} H_{i}^{xx} & H_{i}^{xy} \\ H_{i}^{yx} & H_{i}^{yy} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix}
\]

- Arnoldi vectors as possible initial guesses
Numerical Results: Parallelization

Test setup:

- Shared-memory parallelization
- $N = 60$ particles, Ising-type Hamiltonian
- Pattern: [15 15 15 15]
- $J = 10$ addends, 10 iterations per addend
- Intel Xeon L5520 (8 cores, 2.26 GHz)
Numerical Results: Parallelization

Test setup:

- Shared-memory parallelization
- \( N = 60 \) particles, Ising-type Hamiltonian
- Pattern: [15 15 15 15]
- \( J = 10 \) addends, 10 iterations per addend
- Intel Xeon L5520 (8 cores, 2.26 GHz)

Performance result:

<table>
<thead>
<tr>
<th>Threads</th>
<th>Time [min]</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83</td>
<td>1x</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>1.64x</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>2.39x</td>
</tr>
<tr>
<td>8</td>
<td>26</td>
<td>3.18x</td>
</tr>
</tbody>
</table>
Conclusions / Outlook

• Conclusions:
  • Computation of ground states corresponds to the solution of eigenvalue problems
  • Several decomposition schemes (MPS, ParaFac)
  • ALS technique leads to eigenvalue problems of smaller size

Outlook:

• Improve the ALS (e.g. linesearch methods)
• Allow different blockings and permutations in the decomposition schemes
• Parallelization of the simultaneous optimization

Thank you very much for your attention.
Conclusions / Outlook

• Conclusions:
  • Computation of ground states corresponds to the solution of eigenvalue problems
  • Several decomposition schemes (MPS, ParaFac)
  • ALS technique leads to eigenvalue problems of smaller size

• Outlook:
  • Improve the ALS (e.g. linesearch methods)
  • Allow different blockings and permutations in the decomposition schemes
  • Parallelization of the simultaneous optimization
Conclusions / Outlook

• Conclusions:
  • Computation of ground states corresponds to the solution of eigenvalue problems
  • Several decomposition schemes (MPS, ParaFac)
  • ALS technique leads to eigenvalue problems of smaller size

• Outlook:
  • Improve the ALS (e.g. linesearch methods)
  • Allow different blockings and permutations in the decomposition schemes
  • Parallelization of the simultaneous optimization
Conclusions / Outlook

• Conclusions:
  • Computation of ground states corresponds to the solution of eigenvalue problems
  • Several decomposition schemes (MPS, ParaFac)
  • ALS technique leads to eigenvalue problems of smaller size

• Outlook:
  • Improve the ALS (e.g. linesearch methods)
  • Allow different blockings and permutations in the decomposition schemes
  • Parallelization of the simultaneous optimization

Thank you very much for your attention.