Matrix Product States and Tensor Networks

2/8/2019
Thomas Quella

Quantum Mechanics has a issue with complexity. At times this is an asset, at others a curse...

Situation: System described by a Hilbert space

$$\mathcal{H} = \bigotimes_{\alpha=1}^{N} \mathcal{H}_\alpha^{LOC} \quad \text{with} \quad \mathcal{H}_\alpha^{LOC} = \mathbb{C}^D \quad [D=2: \text{qubit}]$$

local

$$\{ |i\rangle \} \text{ basis of } \mathbb{C}^D \Rightarrow \{ |i_1 - i_N\rangle = |i_1\rangle \otimes \cdots \otimes |i_N\rangle \} \text{ basis of } \mathcal{H}$$

To describe a general (fixed) state

$$|\psi\rangle = \sum_{i_1, -, i_N} \psi_{i_1 - i_N} |i_1 \cdots i_N\rangle \in \mathcal{H}$$

one thus needs to fix D^N coefficients $\psi_{i_1 - i_N} \in \mathbb{C}$.
For operators on $\mathcal{H}$ there are even $D^N \times D^N = D^{2N}$ coefficients e.g. for the projector $P_{|\psi\rangle} = \frac{1}{\langle\psi|\psi\rangle} |\psi\rangle\langle\psi|$.

On computers: Problems with time and memory...

Note: Many problems in quantum computing (and physics more generally) are of a variational nature

Example: Minimize the energy

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \quad (\text{with } H = \text{Hamiltonian})$$

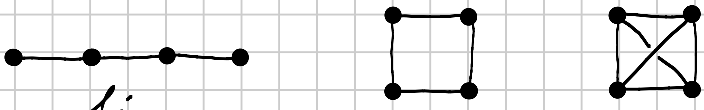
$\Rightarrow$ One would like to keep the search space (number of parameters) as small as possible

Goal: Explore mathematical representation of states (and operators) that are more efficient (in specific circumstances)

[It should be clear that this is not always possible]

Guiding Principle: Make use of the spatial structure and the associated entanglement of the system

Example:

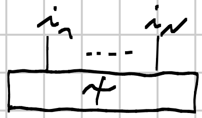
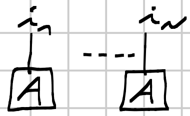
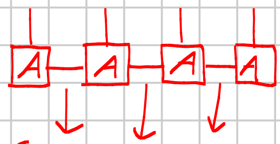


Product states: (Mean field states)

$$|\psi\rangle = |\tilde{\psi}\rangle \otimes \cdots \otimes |\tilde{\psi}\rangle$$

with $|\tilde{\psi}\rangle = \sum_i A^i |i\rangle$ [D coefficients]

$\Rightarrow$ No entanglement

State	General	Product	Matrix Product State
Graphical rep.			

Question: Anything in between?

Entanglement

Matrix Product State (MPS)

Consider matrices $S, A^{i}_{[\alpha]} \in \mathrm{End}(\overbrace{\mathcal{H}^{AUX}}^{\text{auxiliary space, dim}=d,\ \text{basis } \{|\alpha\rangle\}})$ for $\left\{ \begin{matrix} i = 1, \dots, D \\ \alpha = 1, \dots, N \end{matrix} \right\}$. Then

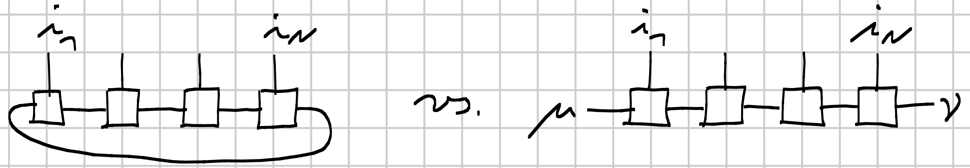
$$|\psi_S\rangle = \sum_{i_1, \dots, i_N} \mathrm{tr}\left(S A^{i_1}_{[1]} \cdots A^{i_N}_{[N]}\right) |i_1 \dots i_N\rangle$$

is called a matrix product state [$L \cdot D \cdot d^2$ coefficients]

The dimension d of the auxiliary space is a measure for the entanglement in $|\psi\rangle$

Periodic BC: $\quad S = \mathbb{1} \quad \longrightarrow \quad \mathrm{tr}\left(A^{i_1}_{[1]} \cdots A^{i_N}_{[N]}\right)$

Open BC: $\quad S = |\nu\rangle\langle\mu| \longrightarrow \left(A^{i_1}_{[1]} \cdots A^{i_N}_{[N]}\right)_{\mu\nu}$



[Open BC: Two perspectives:
- Vary $\mu, \nu \longrightarrow d^2$ states
- Fix $\mu, \nu \longrightarrow A_{[1]}, A_{[N]}$ are effectively row / column vectors, not matrices]

<u>Thm:</u> Every state $|\psi\rangle \in \mathcal{H}$ has a representation as an MPS [see below]

<u>Remark:</u> The MPS representation is most efficient if the entanglement is distributed linearly. In other cases more general "tensor network states" such as PEPS (in 2D) may be more appropriate.

Thm: Singular Value Decomposition (SVD)

Let M be $N_A \times N_B$ matrix and $N = \min\{N_A, N_B\}$. Then there exists a decomposition

$$M = U S V^{+}$$

where:
- S is diagonal with entries $\underbrace{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r}_{\text{singular values}} > 0$
  r = (Schmidt) rank of M
  [The remaining entries are zero]
- U has orthonormal columns (left singular vectors), i.e. $U^+ U = 1\!\!1$
- V^+ has orthonormal rows (right singular vectors), i.e. $V^+ V = 1\!\!1$

Note:
- If $N_A = N$, U is unitary and one also has $UU^+ = 1\!\!1$
- If $N_B = N$, V^+ is unitary and one also has $VV^+ = 1\!\!1$
- Truncation in singular values (only keeping $r' < r$) leads to an optimal approximation w.r.t. the Frobenius norm $\|M\|_F^2 = \sum_{i,j} |M_{ij}|^2 = \mathrm{tr}(M^+ M)$

  in terms of a matrix of rank r'!

$$\mathrm{tr}(M^+ M) = \mathrm{tr}\left[(USV^+)^+ (USV^+)\right]$$
$$= \mathrm{tr}\left[V S^+ \underbrace{U^+ U}_{1\!\!1} S V^+\right]$$
$$= \mathrm{tr}\left[S^+ S \underbrace{V^+ V}_{1\!\!1}\right]$$
$$= \mathrm{tr}(S^+ S) = \sum_{a=1}^{r} \lambda_a$$

Thm: Schmidt decomposition

Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and $\{|i\rangle_A\}$, $\{|j\rangle_B\}$ orthonormal bases with N_A, N_B elements. Any state

$$|\psi\rangle = \sum_{i,j} \underbrace{\Psi_{ij}}_{\hookrightarrow \text{ rectangular matrix}} |i\rangle_A \otimes |j\rangle_B \in \mathcal{H}$$

can be written in the form

$$|\psi\rangle = \sum_{a=1}^{r} \lambda_a \, |a\rangle_A \otimes |a\rangle_B \qquad \text{[Schmidt decomposition]}$$

where $\{|a\rangle_A\}$ and $\{|a\rangle_B\}$ are orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ and λ_a are the (non-zero) singular values of Ψ.

Proof: $\Psi = U S V^\dagger$ (SVD)

$$\Rightarrow |\psi\rangle = \sum_{i,j} \sum_{a=1}^{r} U_{ia} \lambda_a \overline{V_{ja}} \, |i\rangle_A \otimes |j\rangle_B$$

$$= \sum_{a=1}^{r} \lambda_a \underbrace{\Big(\sum_i U_{ia} |i\rangle_A\Big)}_{|a\rangle_A} \otimes \underbrace{\Big(\sum_j \overline{V_{ja}} |j\rangle_B\Big)}_{|a\rangle_B}$$

Orthonormality: ${}_A\langle a|b\rangle_A = \sum_{i,j} {}_A\langle i| \overline{U_{ia}} U_{jb} |j\rangle_B$

$$= \sum_i (U^\dagger)_{ai} U_{ib}$$

$$= (\underbrace{U^\dagger U}_{\mathbb{1}})_{ab} = \delta_{ab}$$

Similarly for ${}_B\langle a|b\rangle_B$ ...

The Schmidt decomposition captures the entanglement of state $|\psi\rangle$ (wrt. the decomposition $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$).

The entanglement (or von Neumann) entropy of $|\psi\rangle$ with respect to the decomposition $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$) is given by

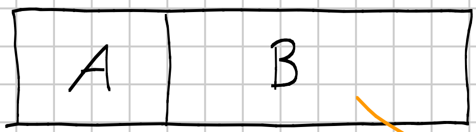
$$S_{A|B}(|\psi\rangle) := -\operatorname{tr}[\rho_A \log \rho_A] = -\operatorname{tr}[\rho_B \log \rho_B]$$

where
$$\left.\begin{aligned} \rho_A &= \operatorname{tr}_B(|\psi\rangle\langle\psi|) = \Psi\Psi^\dagger \\ \rho_B &= \operatorname{tr}_A(|\psi\rangle\langle\psi|) = \Psi^\dagger\Psi \end{aligned}\right\} \text{reduced density matrices}$$

A simple calculation shows

$$S_{A|B}(|\psi\rangle) = -\sum_{\alpha=1}^{r} \lambda_\alpha^2 \log \lambda_\alpha^2$$

<u>Picture:</u>



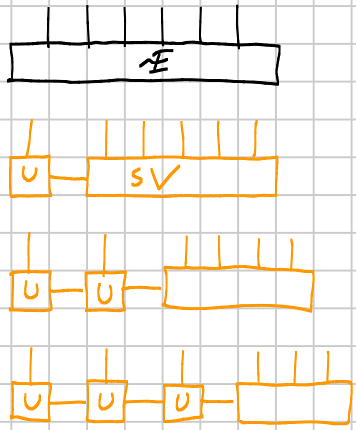
Full system

trace out B (ignore information about B)

From Schmidt decomposition to MPS

Consider a general (normalized) state

$$|\psi\rangle = \sum_{i_1,\dots,i_N} \underbrace{\Psi_{i_1\dots i_N}}_{\Psi_{i_1,(i_2\dots i_N)}} \underbrace{|i_1 i_2 \dots i_N\rangle}_{|i_1\rangle \otimes |i_2 \dots i_N\rangle}$$



An SDV gives

$$\Psi_{i_1\dots i_N} = \Psi_{i_1,(i_2\dots i_N)}$$

$$= \sum_{a_1=1}^{r_1} \underbrace{U_{i_1 a_1}}_{(A^{i_1}_{[1]})_{a_1}} \underbrace{\lambda_{a_1} (V^\dagger)_{a_1,(i_2\dots i_N)}}_{\Psi_{(a_1 i_2),(i_3\dots i_N)}}$$

Iterating this procedure we obtain

$$\Psi_{(a_1 i_2)(i_3\dots i_N)} = \sum_{a_2=1}^{r_2} \underbrace{U_{(a_1 i_2),a_2}}_{(A^{i_2}_{[2]})_{a_1 a_2}} \underbrace{\lambda_{a_2} (V^\dagger)_{a_2,(i_3\dots i_N)}}_{\Psi_{(a_2 i_3),(i_4\dots i_N)}}$$

and eventually

$$\Psi_{i_1\dots i_N} = \sum_{a_1\dots a_{N-1}} (A^{i_1}_{[1]})_{a_1} (A^{i_2}_{[2]})_{a_1 a_2} \dots (A^{i_N}_{[N]})_{a_{N-1}}$$

$$= A^{i_1}_{[1]} \dots A^{i_N}_{[N]}$$

Note: - The first and last A^i are row/column vectors, the remaining ones (potentially rectangular) matrices

- By construction for every factor

$$\sum_i (A^i)^\dagger A^i = 1 \qquad \text{[left-normalized]}$$

We note that an MPS representation is by no means unique:

- By starting the procedure from the right one obtains an MPS with $\sum_i A^i (A^i)^\dagger = \mathbb{1}$ [right-normalized]

  [Mixtures are possible as well]

- Even for left-normalized MPS one still has the gauge symmetry $A_{[i]} \longmapsto U_{[i]} A_{[i]} V_{[i]}^{-1}$

  as long as
$$\left.\begin{aligned} U_{[1]} &= \mathbb{1} \\ V_{[N]} &= \mathbb{1} \\ V_{[i]} &= U_{[i+1]} \end{aligned}\right\}$$
This has very interesting physical consequences...

Why MPS?

In many cases one can approximate the desired state really well (error 10^{-10}) with relatively small matrices A.
[This can be shown to be rigorously true for ground states of gapped 1D quantum systems in view of what is called the "area law" → Hastings]

Also, the calculation of (local) expectation values and overlaps simplifies tremendeously...

Let $|\psi\rangle, |\tilde{\psi}\rangle$ be states with MPS matrices $A, \tilde{A}$.

$$\Rightarrow \langle\psi|\tilde{\psi}\rangle = \sum_{i_1, \dots, i_N} \underbrace{\left(\bar{A}^{i_1}_{[1]} \cdots \bar{A}^{i_N}_{[N]}\right)}_{\text{matrix product}} \underbrace{\left(\tilde{A}^{i_1}_{[1]} \cdots \tilde{A}^{i_N}_{[N]}\right)}_{\text{matrix product}}$$

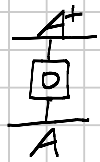
Inefficient: – Calculate matrix product for each choice of i's
– Sum up all contributions

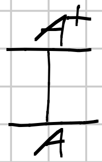
Efficient: – Rewrite as

$$\langle\psi|\tilde{\psi}\rangle = \sum_{i_N} (\tilde{A}^{i_N})^{+} \left[\cdots \left[\sum_{i_1} (\tilde{A}^{i_1})^{+} A^{i_1} \right] \cdots \right] A^{i_N}$$

– Sum according to bracketing (inside out)

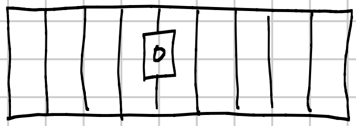
For the expectation value $\langle O_\alpha \rangle$ define

$$E_O = \sum_{i,j} (A^i)^\dagger O_{ij} A^j =$$ 

with special case $E = E_1 =$ 

$d^2 \times d^2$ matrices

Then the expectation values can be calculated as follows:

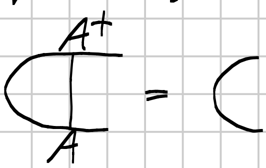
$\langle \psi | O_\alpha | \psi \rangle =$ 

$= E_1 E_2 \cdots (E_O)_\alpha \cdots E_N$

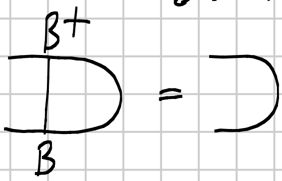
Note that high powers of E can be calculated efficiently (or even analytically) by diagonalizing it

[→ transfer matrix method]

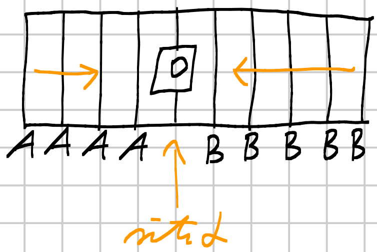
For a mixed left/right normalized MPS with

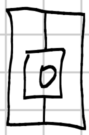


and



one obtains

$\langle \psi | O_\alpha | \psi \rangle =$ 

$=$ 

References

quant-ph/0608197: Matrix Product State Representations

→ Quantum many-body focus

1008.3477: The density-matrix renormalization group in the age of matrix product states

→ Numerical focus

1708.00006: Quantum Tensor Networks in a Nutshell

→ QC focus

Mainly used for this talk