Efficient Tensor Decomposition and Its Application

Naoki KAWASHIMA (ISSP)
Dec. 3, 2018
Occam's Razor

"Pluralitas non est ponenda sine necessitate."
(We should not make more assumptions, if not necessary.)

Be stingy with model parameters!

William of Occam
ca. 1285-1349
from Wikipedia
Tensor Network State

$$\left| \psi \left( \{ T_\alpha \} \right) \right\rangle = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \cdots \sum_{S_N=\pm 1} \text{Cont} \left( \bigotimes T_\alpha \right)_{S_1,S_2,\ldots,S_N} \left| S_1, S_2, \ldots, S_N \right\rangle$$

Parametrized by only $O(N)$ tensors.

Traditional model $O(1)$ << TN model $O(N)$ << Exact model $O(e^N)$
Majority of low-T condensed matter physics problems satisfy the "Area Law"

PEPS satisfies the area law by definition.
Ising Model is a TN
Real Space RG with TN

Occam's Razor in TRG --- SVD

Singular Value Decomposition (SVD) with Truncation

\[ T = USV = \hat{U}\hat{S}\hat{V} = (\hat{U}\sqrt{\hat{S}})(\sqrt{\hat{S}}\hat{V}) = T_1T_2 \]
How good is it?

2D Potts model ($L \leq 1,048,576$) HOTRG calculation with $\chi \sim 50$

S. Morita and NK: arXiv:1806.10275

See Morita's talk on Wednesday

Finite Size Scaling

1st order nature of 5-Potts confirmed
Tensor Network Calculation of ab-initio model for Na$_2$IrO$_3$

Experimental observation (zigzag state) is reproduced.

T. Okubo (U. Tokyo)

ab-initio model for Na$_2$IrO$_3$


\[ \hat{H} = \sum_{\Gamma=X,Y,Z} \sum_{(\ell,m) \in \Gamma} \vec{S}_\ell^T \mathcal{J}_\Gamma \vec{S}_m, \]

\[ \mathcal{J}_Z = \begin{bmatrix} J & I_1 & I_2 \\ I_1 & J & I_2 \\ I_2 & I_1 & K \end{bmatrix}, \mathcal{J}_X = \begin{bmatrix} K' & I''_1 & I'_2 \\ I''_2 & J'' & I'_1 \\ I'_2 & I'_1 & J' \end{bmatrix}, \mathcal{J}_Y = \begin{bmatrix} J'' & I''_2 & I'_1 \\ I''_1 & K' & I'_2 \\ I'_1 & I'_2 & J' \end{bmatrix} \]
S=1 Bilinear-Biquadratic Model


$$H = \sum_{\langle i, j \rangle} \left[ \left( \cos \phi - \frac{\sin \phi}{2} \right) S_i \cdot S_j + \frac{\sin \phi}{2} Q_i \cdot Q_j \right]$$
Improvement of TRG


Optimization condition for $u$, $v$ and $w$

RG transformation:

- can get rid of local entanglement
- converges faster when $D$ increased
What has been improved?
--- Corner Double Line (CDL) Tensor ---

It also appears as the fixed point tensor of the TRG procedure in the disordered phase.
The fate of a local entanglement loop

Suppose each tensor is a CDL
The fate of a local entanglement loop

Focus on a plaquette
The fate of a local entanglement loop

The 1st SVD

The entanglement loop is deformed.
The fate of a local entanglement loop

The network after contraction of small squares.

The green loop is still there.
The fate of a local entanglement loop

The 2nd SVD
Some of the ent. loops have been removed. But at each generation the influence of the original ent. loop remains.

The expressive capacity of the network is wasted.

The fate of a local entanglement loop

After the 2nd SVD

The green loop still survives.
Removal of Ent. Loops

By pinching the "information path", we can split the remaining loop, and remove them at the next contraction.

Another example of loops: Tensor Ring Decomposition (TRD)
TRD in Informatics --- Images

Zhao, Cichocki ら arXiv:1606.05535

Columbia Object Image Libraries (COIL)-100 dataset

data set = 100 object image sets
1 object image set = 72 images
1 image = 128 x 128 dots
1 dot = 3 colors

\[ T_{xyci} \]

\[ x = 1, \ldots, 128 \quad \text{... x-coordinate} \]
\[ y = 1, \ldots, 128 \quad \text{... y-coordinate} \]
\[ c = 1, \ldots, 3 \quad \text{... color} \]
\[ i = 1, \ldots, 7200 \quad \text{... image ID} \]
COIL100 2D image classification task
128 x 128 x 3 x 7200 bits

KNN classifier (K=1) applied to the image specifier core ($Z_4$).

<table>
<thead>
<tr>
<th>Method</th>
<th>tolerated error</th>
<th>maximum bond dim.</th>
<th>average bond dim.</th>
<th>score (%) (large training)</th>
<th>score (%) (small training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT-SVD &quot;open chain&quot;</td>
<td>0.19</td>
<td>67</td>
<td>47.3</td>
<td>99.05</td>
<td>89.11</td>
</tr>
<tr>
<td></td>
<td>0.28</td>
<td>23</td>
<td>16.3</td>
<td>98.99</td>
<td>88.45</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
<td>8</td>
<td>6.3</td>
<td>96.29</td>
<td>86.02</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
<td>3</td>
<td>2.7</td>
<td>47.78</td>
<td>44.00</td>
</tr>
<tr>
<td>TR-SVD &quot;ring&quot;</td>
<td>0.19</td>
<td>23</td>
<td>12.0</td>
<td>99.14</td>
<td>89.29</td>
</tr>
<tr>
<td></td>
<td>0.28</td>
<td>10</td>
<td>6.0</td>
<td>99.19</td>
<td>89.89</td>
</tr>
<tr>
<td></td>
<td>0.36</td>
<td>5</td>
<td>3.5</td>
<td>98.51</td>
<td>88.10</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
<td>3</td>
<td>2.3</td>
<td>83.43</td>
<td>73.20</td>
</tr>
</tbody>
</table>

Ring decomposition shows better performance.

... but can we do that easily?
Alternating Least Square (ALS)

(1) random initial tensors $Z_i$

(2) for $i=1,2,3,4$, update $Z_i$ by

$$\min_{Z_i} T_{Z_1} Z_2 Z_3 Z_4$$

(3) repeat until the error converges

However, ALS is trapped by local loops.
ALS on CDL

\[ \delta \psi = \frac{\| T - t \text{Tr}\{Z^a\} \|_F}{\| T \|_F} \]

ALS on CDL is either unstable or stuck with a local minimum.
(At least partially, due to the local entanglement loops.)

H.-Y. Lee and N.K.
arXiv:1807.03862

sALS ... The initial condition obtained by sequential (open chain) SVD
Redundant entanglement loops causes various problems. (reduction of expressive power, and obstacle in optimization.)

Any direct method for removing them?
If we knew $U$, $V$, $W$ and $x$, $y$, $z$ explicitly, we can find $Z_1$, $Z_2$, $Z_3$ of the TRD very easily.

... but how do we know them?
Ring Decomposition by Index Splitting

When the given tensor $T$ is a CDL, i.e., it must have the following form:

$$\Delta_{pq}^i = \delta_{i,I(p,q)}$$

$e^{i\phi}$

... then, we can find $U$, $V$ and $W$ by HOSVD
Index Splitting


\[ I(p, q) \]

... injection from \((p, q)\) to \(i\) such as e.g.,

\[
\begin{align*}
I(0, 0) &= 0 \\
I(0, 1) &= 1 \\
I(1, 0) &= 2 \\
I(1, 1) &= 3
\end{align*}
\]
Uniqueness of HOSVD


If $T$ is expressed as a core tensor $t$ and unitaries $U$, $V$, and $W$, where any matrix slices of $t$ are mutually orthogonal, such an expression is unique up to the permutation within each index and the phase factors.

CDL satisfies mutual orthogonality

$\Rightarrow$ $U, V, W$ can be obtained by HOSVD
By working directly with the inner structure, we can avoid the difficulty of the local minima in the optimization.
2D Ising Model above \( T_c \)

Summary

- TN representation makes it possible to handle extremely large systems, frustrated systems, etc.
- CDL-like structure typical in TN-based RG often cause serious difficulty.
- Index-splitting based on HOSVD may be useful in overcoming the difficulty.