8. PEPs IV: Corner Transfer Matrix (CTM)

Goal: more efficient computation of environment tensors; flexible unit cell.

Original idea of "corner transfer matrix": [Baxter 1968], [Baxter 1978]
First application in tensor network context: [Nishino 1996], with Okunishi; "corner transfer renormalization group"
In PEPs context:
[Corboz 2009], with Vidal: proof-of-principle for 2D quantum Ising model
[Corboz 2010], with Jordan, Vidal: spinless fermions with n.n. hub; t-J model
[Corboz 2011], with White, Vidal, Troyer: Stripes in t-J model
[Corboz 2014], with Rice, Troyer: t-J model (competitive results!)

Notation [Brunsgaard 2017]

\[ | \Psi \rangle = \]

\[ \langle \Psi | = \]

Local tensor: \[ M_{\alpha \beta \gamma \rho} \]
Cayley tree tensor: \[ M_{\alpha \beta \gamma \rho}^{\text{CTM}} = \alpha \beta \gamma \rho \]
Contracting out physical indices

Detail relevant for fermions:
Variable unit cell [Corboz 2011]

For systems with spontaneous symmetry breaking, one needs to vary size of unit cell, to accommodate various types of order...

FIG. 4. (Color online) Examples of stripes running in the vertical direction obtained from various iPEPS simulations. Each panel shows one unit cell of the infinite lattice. The diameter of the dots scales with the local hole density with average values given by the (upper) red numbers. The arrows represent the local magnetic moment with average magnitude given by the (lower) black numbers. There is a $\pi$-phase shift in the antiferromagnetic order between adjacent stripes. The width of a bond between two sites scales with the (singlet) pairing amplitude on the bond with a positive (green/dark gray) or negative (cyan/light gray) sign. A pattern with predominantly $d$-wave order is visible, with maximal pairing amplitude 0.01, 0.03, and 0.003 in

Corner transfer matrix approach: main idea

represent environment of $2 \times 2$ "growing window" in terms of 12 tensors:
8 transfer tensors (T’s) & corner tensors (C’s), [here: $2 \times 2$ unit cell]

Bond dimensions:
$\chi \geq D^2$
CTM coarse-graining move: insertion to the left

Instead of \( \lim_{p \to \infty} R^p(\mathcal{E}_z) \): repeatedly

(i) insert unit cell,
(ii) absorb rows,
(iii) renormalize

CTM coarse-graining move: insertion at the top

(i) insertion of extra unit cell

(ii) absorption of first row

(iii) renormalization

...
Renormalization

There is no clear "optimal way" to do this, due to lack of canonical form.

$\Rightarrow$ trial and error / room for creativity! We describe

[Corboz2009], [Corboz2010a], [Corboz2014 Supplement], [Shagano2017]

(i) cut system in half, and do QR-decomposition of up and down parts.

(ii) Construct projectors

Contrast $R_u$ and $R_d$

SVD

Truncate

Inverse

(discard very small singular values, so $S^{-1}$ is well-conditioned)

$1 \approx U S V^T$ represents environment for $R_u$

represents environment for $R_d$
(iii) Insert projectors ( = absorb & renormalize)

\[ \rho_j^x \rightarrow \tilde{\rho}_j^x \]

Renormalized tensors

\[ \tilde{C}_{(x+y)} - \tilde{T}_{(x+y)} \]

- Renormalization encodes information about full environment
- Small singular values can be discarded before inverting $\Sigma$

Larger unit cells

For unit cell of $L_x \times L_y = N$ sites, we need $4C's$, $8T's$ for each (i.e. $4N$, $8W$ in $k$th).

Huge zooming window is not an option, since carb grows rapidly.
(iv) Stepping to find $C_n^*$, $T_s^*$

Initialization:

$C_{id}^*$ = \[ \begin{array}{c}
M_1^* \\
M_2^*
\end{array} \] , \hspace{1cm} T_{id}^* = \begin{array}{c}
M_1^* \\
M_2^*
\end{array}

are $M$'s obtained from previous imaginary time evolution step.

1 cycle = insert unit cell at left, perform $L_x$ absorption & renormalization.
- top
- right
- bottom

Repeat cycle numerous times until convergence (e.g. of local observables).

Effective contraction pattern

Cost for CIUS scheme on square lattice: $O(D^6 x^3) \sim O(D^6)$.

But only if optimal contraction patterns are used:

$O(D^3 x^2)$

Better: open up in $\mathcal{I}M\mathcal{O}M$:

$O(x^2 D^6)$

(2.107)
Observables (e.g. on two neighboring sites)

\[ \langle \psi | \hat{O}^{(x+1,y)} | \psi \rangle \chi = \]

\[ \begin{array}{c}
\hat{O} \\
\hat{M}_x^+ \\
\hat{M}_x^- \\
\hat{E}_1 \\
\hat{E}_2 \\
\hat{M}_x^+ \\
\hat{E}_3 \\
\hat{M}_y^+ \\
\hat{M}_y^- \\
\hat{E}_4 \\
\hat{E}_5 \\
\hat{M}_y^+ \\
\hat{M}_y^- \\
\hat{E}_6 \\
\end{array} \]

\[ = \begin{array}{c}
\hat{O} \\
\hat{T}_{x+y} \\
\hat{C}_{x+y} \\
\hat{T}_{x+y} \\
\hat{T}_{x+y} \\
\hat{C}_{x+y} \\
\hat{T}_{x+y} \\
\hat{T}_{x+y} \\
\hat{C}_{x+y} \\
\hat{T}_{x+y} \\
\hat{T}_{x+y} \\
\hat{C}_{x+y} \\
\end{array} \]

can be computed efficiently, by systematic blocking...

Repeat for \( \chi = 10, 20, \ldots, 100, 150 \) until convergence...

(which depends on \( D, \text{ and on model} \))

Optimization of \( M \)-tensors

Most PEPS-workers use Trotter-imaginary-time-evolution, as proposed in [Jordan 2009] or PEPS

\[ | \psi_{\text{eq}} \rangle = \lim_{\tau \to \infty} e^{-\tau \mathcal{H}} | \psi \rangle \]

\[ \frac{e^{-\tau \mathcal{H}}}{\| e^{-\tau \mathcal{H}} | \psi \rangle \| ^2} \]

\( e^{-\tau \mathcal{H}} \) : Trotterize into nearest-neighbor two-side gates.

Side remarks:

- Larger-ranged interactions are also possible: [Corboz 2010], [Corboz 2013]

- Alternative approach: variational optimization of
  \[ \frac{\langle \psi_\text{eq} | H | \psi_\text{eq} \rangle}{\langle \psi_\text{eq} | \psi_\text{eq} \rangle} \]
  [Corboz 2016]
  [Vanderstraeten 2016] "gradient method"
Tensor update for N.N. interaction can be done with cost $O(D^6d^3)$ using bond projection: \cite{Li2012} with vDelft, Xiang; \cite{Phien2015}

$$
\exp(-i\delta t |y\rangle \\
= M^e_x \hat{d} M^e_y \\
= D_x^e Y_x^{e+1} Y_y^{e+1} \\
= X_x^e \hat{d} \tilde{v}_x^e D_x^e Y_x^{e+1} Y_y^{e+1} \\
= D_x^e Y_x^{e+1} Y_y^{e+1} \\
= SVD 
$$

(2.113)

Advantage: $X$ and $Y$ are not affected by 2-site gate.

Two reduce bond dimension again, there are two possibilities:

- Simple update (cheap)
- Full update (expensive)

$$
= 14(0,0)
$$

---

Simple update: bond projection for decorated $\Gamma$.

Original idea due to: \cite{Tiang2008}, with Wang, Xiang.

Go to a Gamma representation (as $\pi_1$ in TEBD):

Consider "decorated $\Gamma"$: $\lambda_i$'s contain info about environment.
Simple update (continued): apply gate

\begin{equation}
\Gamma'_{x} = \gamma^{-1} \hat{\Gamma}'_{x} \gamma \gamma^{-1} = \hat{\Gamma}'_{x} \gamma^{-1} \gamma = \hat{\Gamma}'_{x} \gamma^{-1} \gamma
\end{equation}

Comments on simple update

- Cost cheap: $O(D^3d^6)$
- Update would be optimal if environment were in canonical form.
  But it is not, due to loops! (Would be optimal on Bethe lattice) [Li 2012].
- Do simple update generally along all bonds in unit cell.
- 2nd-order Trotter possible: reverse order of gates every 2nd step.
- To achieve normalizaiton of PEPs: normalize $\hat{\Gamma}'_s$ such that $\sum_i \xi_i = 1$.
- Iteration scheme:
  - Start from random PEPs
  - Initially large time steps, small bond dimensions.
  - Gradually reduce time steps when SVD spectra converge, down to $\delta t \sim 10^{-5}$. 
Better: Variationally optimize using same environment for both terms:

\[ d(\tilde{v}, \tilde{w}, v', w') = \| |\psi(v', w')\rangle - |\psi(\tilde{v}, \tilde{w})\rangle\|^2 \]

Environment including back projection:

\[ d(\tilde{v}, \tilde{w}, v', w') = \| |\psi(v', w')\rangle - |\psi(\tilde{v}, \tilde{w})\rangle\|^2 \]

\[ = \langle \psi(v', w')|\psi(v', w')\rangle + \langle \psi(\tilde{v}, \tilde{w})|\psi(\tilde{v}, \tilde{w})\rangle - \langle \psi(v', w')|\psi(\tilde{v}, \tilde{w})\rangle - \langle \psi(\tilde{v}, \tilde{w})|\psi(v', w')\rangle \]
Optimization conditions

$$\frac{\partial}{\partial v} d(\tilde{v}, \tilde{w}, \tilde{v}', \tilde{w}') = 0 \Rightarrow \begin{array}{c}
\begin{array}{c}
\tilde{v}^x = \tilde{w}^x \\
\tilde{v}'^x = \tilde{w}'^x
\end{array}
\end{array}
= \begin{array}{c}
\begin{array}{c}
R \ E_{\text{full}} \ S \\
R \ v' = S
\end{array}
\end{array}
\Rightarrow R \ v' = S \Rightarrow \text{solve for } \tilde{v}'
$$

Cost: $O(d^3 D^4)$ [without bandwidth projection: $O(D^3)$, exactly: $O(D^2)$]

Next:

$$\frac{\partial}{\partial w} d(\tilde{v}, \tilde{w}, \tilde{v}', \tilde{w}') = 0 \Rightarrow \begin{array}{c}
\begin{array}{c}
\tilde{w}^x = \tilde{v}^x \\
\tilde{w}'^x = \tilde{v}'^x
\end{array}
\end{array}
= \begin{array}{c}
\begin{array}{c}
R \ E_{\text{full}} \ S \\
R \ w' = S
\end{array}
\end{array}
\Rightarrow R \ w' = S \Rightarrow \text{solve for } \tilde{w}'
$$

Steady, until convergence: $fd = |\delta t_{i+1} - \delta t_i| \leq 10^{-6}$

Before original form:

$$M^{v,x} \ = \ X^{v,x} \tilde{v}'^x, \ M^{w,x} \ = \ Y^{w,x} \tilde{w}'^x$$

Comments on full update

- Overall cost: $O(D^4 D^4)$ [same as simple update] -
  
  [but needs to be repeated over and over again ($\geq 10^4$ times)]

- "Fast full update": update environment only after full unit cell has been involved (not after every Zitter-burst) [Phien 2015]
  [Use with care: can cause instabilities...]

- Further improvement: gauge fixing! [Lubash 2014], [Phien 2015]

- Improve stability: maximize all masses such that largest element in $O(1)$.

- Restore hermiticity to environment: use $E_{\text{full}} \equiv \frac{1}{2}(E_{\text{full}} + E_{\text{full}}^*)$. 

PEPS 8.15

PEPS 8.15