

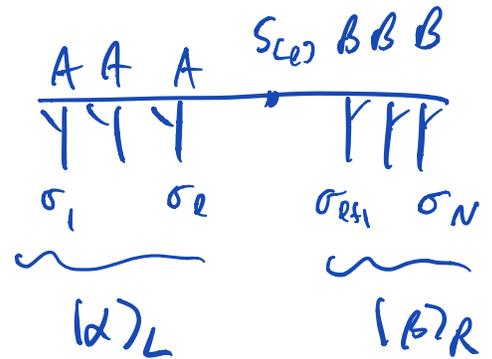
[credit for course materials: Prof. Jan von Delft]

Gamma-Lambda Notation, iTEBD, DMRG II L infiniteVidal's $\Gamma\Lambda$ notation

[Vidal2003, Schollwock2011 Sec 4.6]

We have studied the usual bond-canonical form of MPS:

$$|\psi\rangle = |\beta\rangle_{e,R} |\alpha\rangle_{e,L} S_{(e)}^{\alpha\beta}$$

Choose S diagonal, and call it Λ (following Vidal):

$$|\psi\rangle = |\alpha\rangle_{e,R} |\alpha\rangle_{e,L} \Lambda_{(e)}^{\alpha\alpha}$$

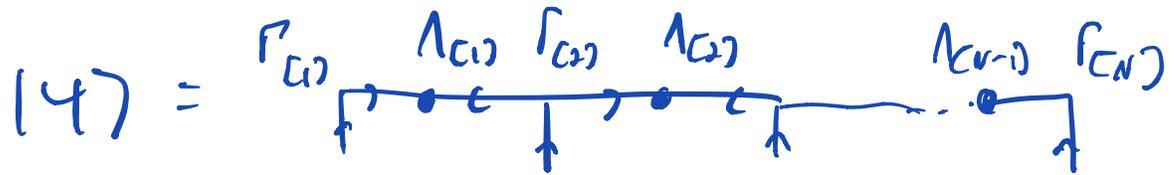
In this format, the reduced density matrices of left and right parts are diagonal, with

eigenvalues $(\Lambda_{(e)}^{\alpha\alpha})^2$:

$$\rho_L = \text{Tr}_R |\psi\rangle\langle\psi| = \sum_{\alpha} |\alpha\rangle_{e,L} (\Lambda_{(e)}^{\alpha\alpha})^2 \langle\alpha|$$

$$\rho_R = \text{Tr}_L |\psi\rangle\langle\psi| = \sum_{\alpha} |\alpha\rangle_{e,R} (\Lambda_{(e)}^{\alpha\alpha})^2 \langle\alpha|$$

Vidal introduced a representation for an MPS in which a Schmidt-decomposition can be directly read off for each bond:



where $\Lambda_{[l]}$ is a diagonal matrix consisting of Schmidt coefficients wrt to bond l
 i.e.

$$|\psi\rangle = |\alpha\rangle_{l,R} |\alpha\rangle_{l,L} \Lambda_{[l]}^{d,d} \quad , \quad \rho \text{ obtained from } (\Lambda_{[l]}^{d,d})^2$$

with orthonormal sets on

$$L: \quad \langle \alpha' | \alpha \rangle_{l,L} = \delta_{\alpha'}^{\alpha}$$

$$R: \quad \langle \alpha' | \alpha \rangle_{l,R} = \delta_{\alpha'}^{\alpha}$$

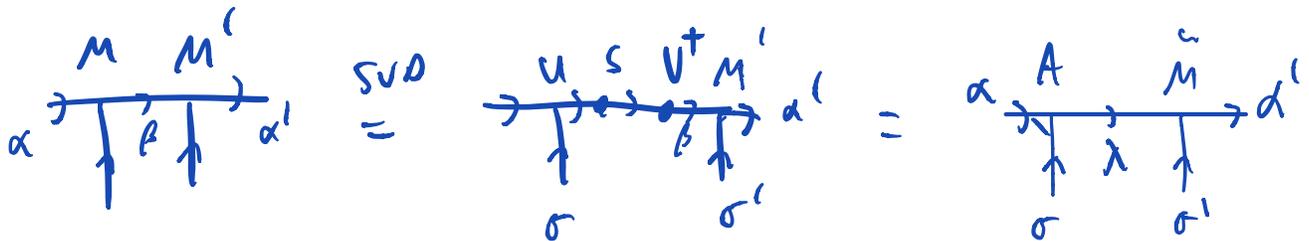
Any MPS can be brought into $\Gamma\Lambda$ form. Proceed in same manner as when left-normalizing:

$$|\psi\rangle = |\bar{\sigma}\rangle_N (M^{\sigma_1} \dots M^{\sigma_N})$$

Successively use SVD on pairs of adjacent tensors:

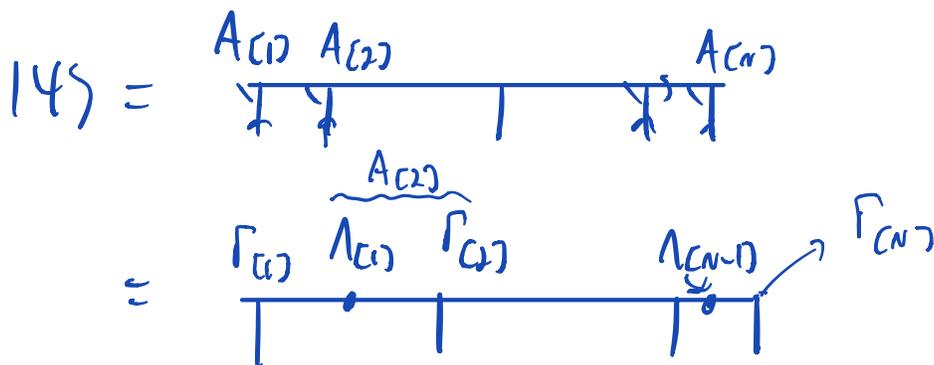
$$MM' = USV^T M' = AM^{\tilde{c}}$$

$$\Rightarrow A=U, \tilde{M} = SV^T M'$$



but store singular values, and at end define:

$$\Gamma_{[1]}^{\sigma_1} = A_{[1]}^{\sigma_1}; \quad \Lambda_{[l-1]} \Gamma_{[l]}^{\sigma_l} \equiv A_{[l]}^{\sigma_l}$$



Note: in numerical practice, this involves dividing by singular values,

$$\Gamma_{[l]}^{\sigma_l} \equiv \Lambda_{[l-1]}^{-1} A_{[l]}^{\sigma_l}$$

So: first truncate states for which

$$\Lambda_{[l-1]}^{\alpha\alpha} = 0$$

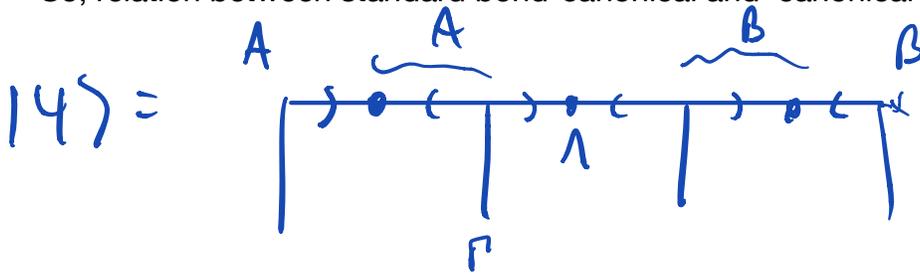
Even then, the procedure can be numerically unstable, since arbitrarily small singular values may arise.

So, truncate states for which (say) $\Lambda_{[l-1]}^{\alpha\alpha} < 10^{-8}$ [contributions $\sim \Lambda^2 \rightarrow \Lambda^4 \sim 10^{-16}$ in eps]

Similarly, if we start from the right, SVDs yield right-normalized B tensors, and we can define:

$$\Gamma_{[l]}^{\sigma_l} \Lambda_{[l]} \equiv B_{[l]}^{\sigma_l} \cdot \text{To get } \Gamma, \text{ right-multiply by } \Lambda_{[l]}^{-1}$$

So, relation between standard bond-canonical and 'canonical $\Gamma\Lambda$ form' is:



$$I = A_{[l]\sigma}^{\dagger} A_{[l]}^{\sigma} = \Gamma_{[l]\sigma}^{\dagger} \Lambda_{[l-1]}^{\dagger} \Lambda_{[l-1]} \Gamma_{[l]}^{\sigma}$$

$$= \Gamma_{[l]\sigma}^{\dagger} \rho_{[l-1]R} \Gamma_{[l]}^{\sigma}$$

$$I = B_{[l]}^{\sigma} B_{[l]\sigma}^{\dagger} = \Gamma_{[l]}^{\sigma} \Lambda_{[l]} \Lambda_{[l]}^{\dagger} \Gamma_{[l]\sigma}^{\dagger} = \Gamma_{[l]}^{\sigma} \rho_{[l]L} \Gamma_{[l]\sigma}^{\dagger}$$

Infinite Time-Evolving Block Decimation (iTEBD)

[Vidal2007, Schollwock2011, Sec 10.4]

Goal: ground state search for infinite system while exploiting translational invariance.

We will use Vidal's Λ notation but everything can be translated into AB notation.

Basic idea: 'imaginary time evolution': $\lim_{\beta \rightarrow \infty} e^{-\beta \hat{H}} |4\rangle \propto |g\rangle$

Reason: high-energy states die out quickly (if ground state is gapped):

$$e^{-\beta \hat{H}} = \sum_{\alpha} e^{-\beta E_{\alpha}} |\alpha\rangle \langle \alpha| \xrightarrow{\beta \rightarrow \infty} e^{-\beta E_g} |g\rangle \langle g|$$

$\underbrace{\hspace{10em}}_{\text{ground state}}$

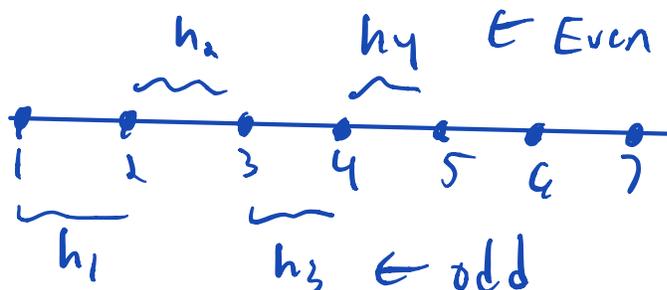
1. Trotter decomposition of time evolution operator

[Schollwock 2011, Sec 7.1.1]

General: write Hamiltonian as

$$\hat{H} = \sum_l \hat{h}_l = \underbrace{\hat{H}_o}_{\text{odd}} + \underbrace{\hat{H}_e}_{\text{even}}$$

\uparrow
connects l + $l+1$



Then all odd terms mutually commute, and all even terms mutually commute:

$$[\hat{h}_l, \hat{h}_{l'}] = 0 \quad \text{if } l, l' \text{ are both odd or even}$$

Divide time interval into L slices:

$$\beta = L \tau \quad \tau \sim \text{time step}$$

$$e^{-\beta H} \stackrel{\text{Trick}}{=} [e^{-\tau H}]^L = [e^{-\tau(\hat{H}_0 + \hat{H}_e)}]^L$$

$$\xrightarrow{L \rightarrow \infty, \tau \rightarrow 0} \approx [e^{-\tau \hat{H}_0} e^{-\tau \hat{H}_e} + o(\tau^2)]^L \quad \text{'first order Trotter approx'}$$

$$\text{or } \approx [e^{-\frac{\tau}{2} \hat{H}_0} e^{-\tau \hat{H}_e} e^{-\frac{\tau}{2} \hat{H}_0} + o(\tau^3)]^L \quad \text{'Second order Trotter approx'}$$

Exploiting commuting properties of odd/even terms, each exponential can be expanded separately without further approximation:

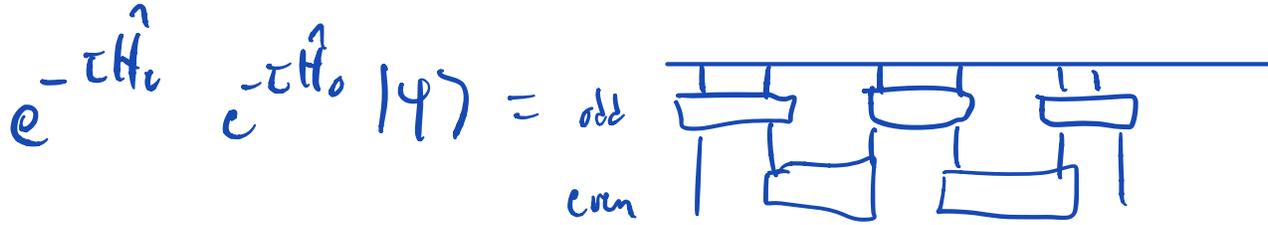
$$e^{-\tau \hat{H}_0} = e^{-\tau \hat{h}_1} e^{-\tau \hat{h}_3} \dots e^{-\tau \hat{h}_{N-1}} \equiv U_{(1)} \dots U_{(N-1)}$$

not unitary!
↓

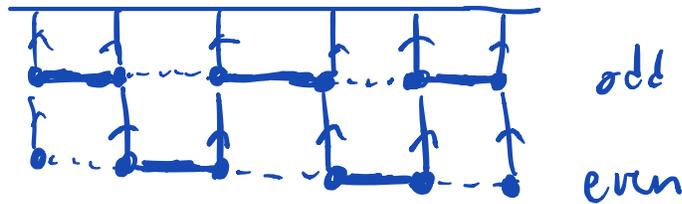
$$e^{-\tau \hat{H}_e} = e^{-\tau \hat{h}_2} e^{-\tau \hat{h}_4} \dots e^{-\tau \hat{h}_N} = U_{(2)} \dots U_{(N)}$$

$$[\text{Assume } N = \text{even}]$$

So, when applying $e^{-\beta \hat{H}}$ to $| \psi \rangle$, we can successively apply all odd terms, then truncate, then all even ones, then truncate, etc.



In MPO notation:



since \hat{H}_o factorizes, even bonds have dimension $D_{w,e} = 1$
 since \hat{H}_e factorizes, odd bonds have dimension $D_{w,o} = 1$

All of this can be done for finite chain of length N . But a simplification occurs for $N \rightarrow \infty$

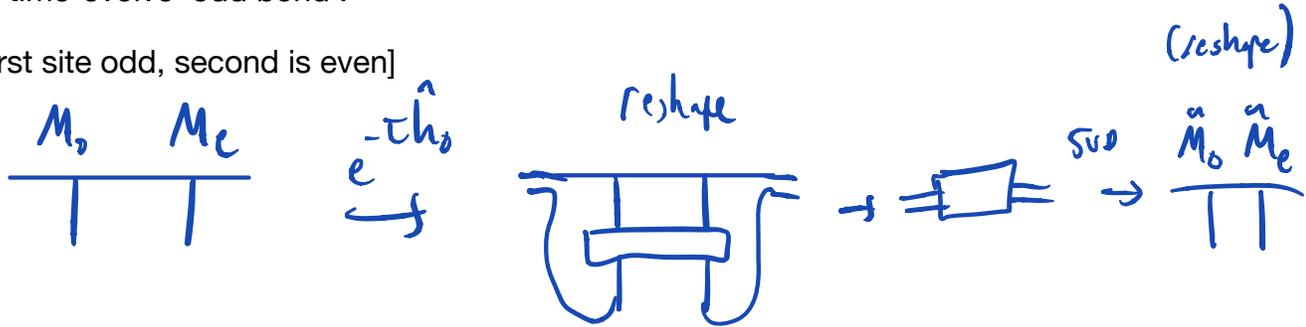
Then we can exploit translational invariance:



Adopt a two-site unit cell (no left- or right- normalization implied).

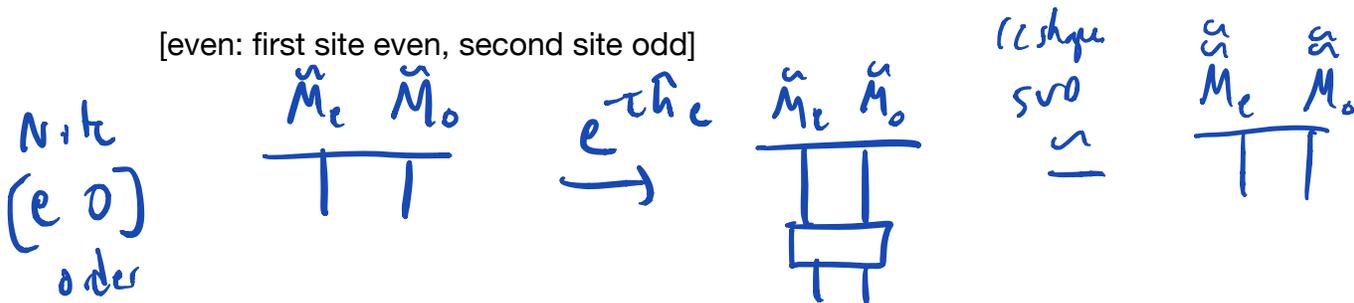
Step 1: time-evolve 'odd bond':

[odd: first site odd, second is even]



Step 2: time-evolve (updated!) even bond:

[even: first site even, second site odd]



Iterate until convergence! (To discuss details, we will use $\Gamma \Lambda$ notation.)

iTEBD is a 'power method': the projector to the ground state is constructed as an

increasing number of powers of $e^{-\tau \hat{H}_e} e^{-\tau \hat{H}_o}$

This is to be contrasted to DMRG ground state search, which is a variational method.

Main advantage of iTEBD: costs not proportional to system size, hence cheaper.

Main disadvantage: loss of orthogonality due to projection without explicit reorthogonalization.

2. iTEBD: Explicit formulation

[Vidal 2007, Schollwock2011, Sec 10.4]

Use two-site unit cell, $\begin{array}{|c|c|} \hline A_0 & A_e \\ \hline \end{array}$, repeated periodically,

$$(4) = \left(\begin{array}{|c|c|} \hline A_0 & A_e \\ \hline \end{array} \right) \begin{array}{|c|c|} \hline A_0 & A_e \\ \hline \end{array} \begin{array}{|c|c|} \hline A_0 & A_e \\ \hline \end{array} \begin{array}{|c|c|} \hline A_0 & A_e \\ \hline \end{array} \dots$$

$\sigma_0 \quad \sigma_e$

and express it in Λ notation: $A_0 = \Lambda_e \sigma_0$, $A_e = \Lambda_0 \sigma_e$

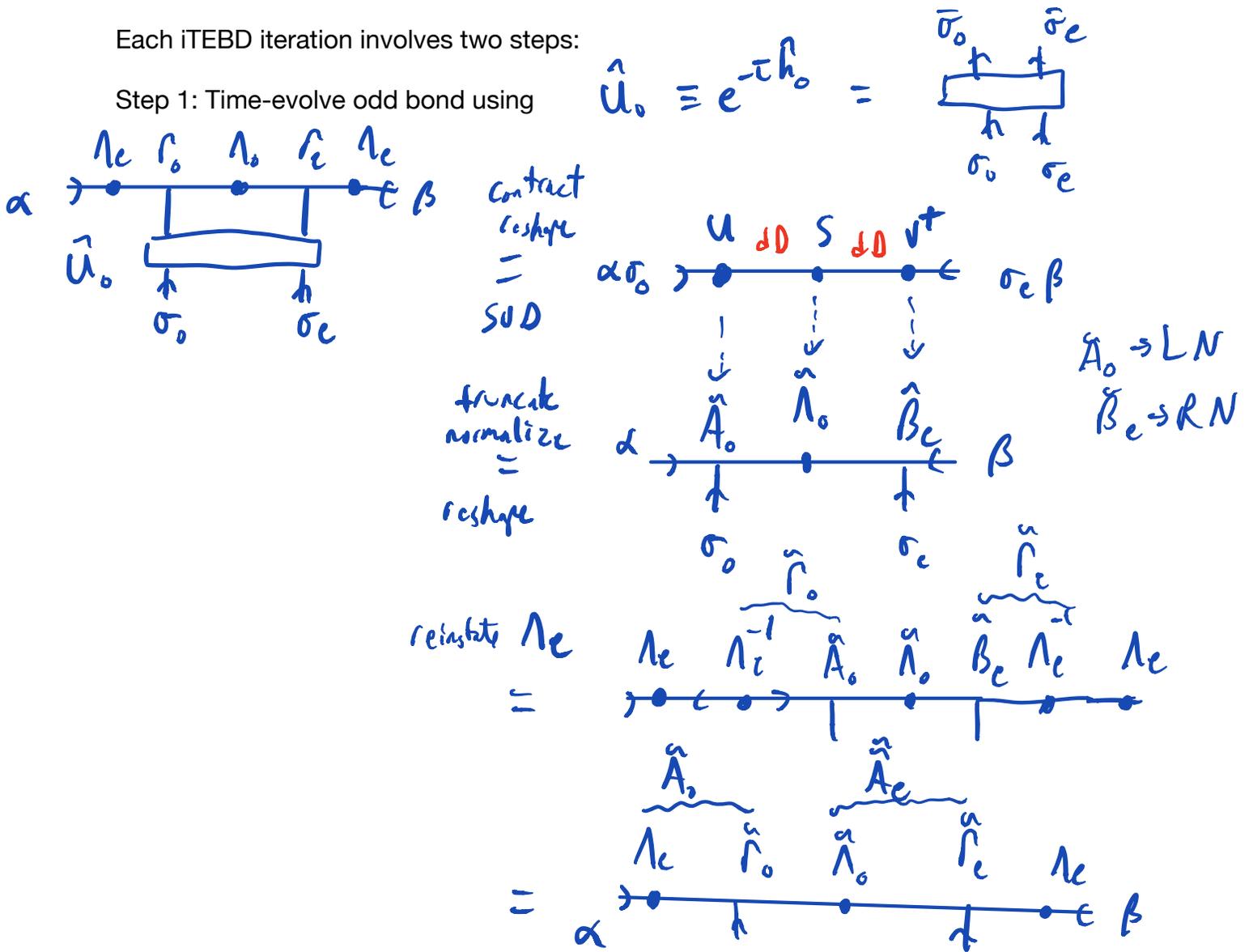
$$(4) = \begin{array}{|c|c|} \hline \Lambda_e \sigma_0 & \Lambda_0 \sigma_e \\ \hline \end{array} \begin{array}{|c|c|} \hline \Lambda_0 \sigma_e & \Lambda_e \sigma_0 \\ \hline \end{array} \begin{array}{|c|c|} \hline \Lambda_e \sigma_0 & \Lambda_0 \sigma_e \\ \hline \end{array} \begin{array}{|c|c|} \hline \Lambda_0 \sigma_e & \Lambda_e \sigma_0 \\ \hline \end{array} \dots$$

$\sigma_0 \quad \sigma_e$

[to avoid cluttering, σ indices on Λ, A are not displayed but implicit]

Each iTEBD iteration involves two steps:

Step 1: Time-evolve odd bond using



Now, \hat{U}_0 is projector (not unitary) hence reduces norm. Thus, $\hat{\Lambda}_0$ is normalized to unity by hand:

$$\hat{\Lambda}_0 \equiv \frac{S_{\text{frunc}}}{\sqrt{\text{Tr}(S_{\text{frunc}}^\dagger S_{\text{frunc}})}}$$

$$= \frac{\hat{\Lambda}_0 \hat{\Lambda}_0 \hat{\beta}_e}{\hat{\Lambda}_0^\dagger \hat{\Lambda}_0^\dagger \hat{\beta}_e^\dagger} = \frac{\hat{\Lambda}_0}{\hat{\Lambda}_0^\dagger} = \text{Tr} \hat{\Lambda}_0^\dagger \hat{\Lambda}_0 = 1$$

This is update of odd bond. The updated MPS now has the form

$$|\psi\rangle = \prod_{\bar{\sigma}} \bar{A}_0 \hat{\tilde{A}}_e \tilde{A}_0 \hat{\tilde{A}}_e \dots$$

Updated bond energy:

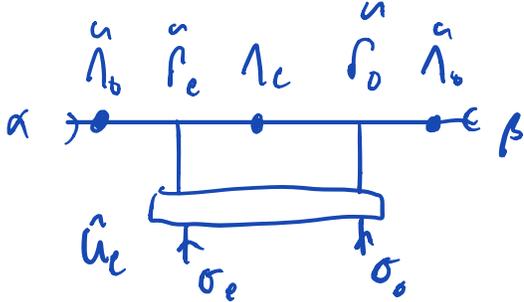
$$\bar{h}_{\text{bond}} = \frac{1}{2} (\bar{h}_0 + \bar{h}_e) \equiv \frac{(h_0) + (h_e)}{2}$$

ignoring tensors for the rest of chain, $\hat{\tilde{A}}_e$ is not LN.

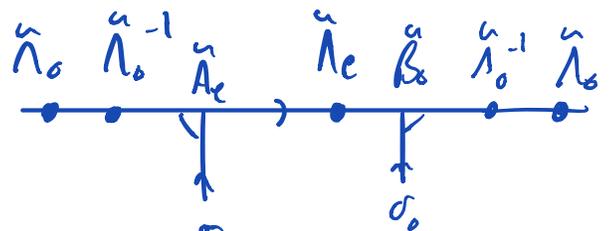
Updating odd bonds lower \bar{h}_0 , slightly raises \bar{h}_e
 ('odd bond happy, even bond slightly unhappy')

Step 2: Time-evolve even bond, using

$$\hat{U}_e \equiv e^{-i\hat{h}_e} = \begin{matrix} \bar{\sigma}_e & \uparrow & \uparrow & \bar{\sigma}_0 \\ \sigma_e & \uparrow & \uparrow & \sigma_0 \end{matrix}$$



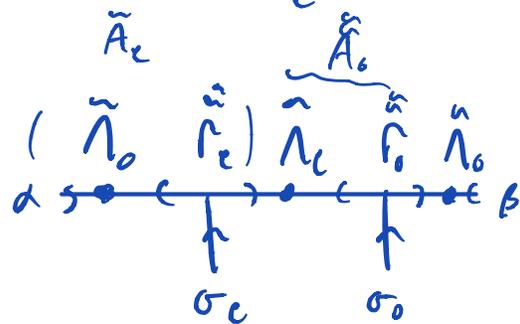
Contract
 reshape, SVD
 =
 reshape
 truncate
 normalize
 reinstate Λ_0



Define

$$\tilde{A}_e^u \equiv \tilde{\Lambda}_0^{u-1} \hat{A}_e^u$$

$$\tilde{A}_0^u \equiv \tilde{b}_0^u \tilde{\Lambda}_0^u$$



This completes update of even bond. Updated MPS now has the form

$$|\tilde{\psi}\rangle = \prod_{\tilde{\sigma}} |\tilde{\sigma}\rangle \tilde{A}_e \tilde{A}_o \tilde{A}_e \tilde{A}_o \dots$$

Compute updated bond energy using same equation as before but with o <-> e.

As before, updating even bond lowers \tilde{h}_e , slightly raises \tilde{h}_o .

Now iterate (apply \hat{U}_o , then \hat{U}_e , etc) until convergence (monitor ground state energy).

Remarks:

1. Computation of $\tilde{\Lambda}_o, \tilde{\Lambda}_e$ can become unstable because singular values can be small. Thus: truncate by discarding smallest singular values $\leq 10^{-8}$, then invert.

2. Note that \tilde{A}_o is left-normalized but $\tilde{A}_e = \tilde{\Lambda}_o \tilde{B}_e \tilde{\Lambda}_e^{-1}$ is not!
 $\tilde{A}_e = \tilde{\Lambda}_e \tilde{B}_o \tilde{\Lambda}_o^{-1}$

'Loss of orthogonality'. This causes problems when computing expectation values. For example, odd bond energy, given by:



does not reduce to earlier expression because zippers cannot be closed from left and right. Hence our evaluation for energy involves an approximation.

Summary remarks on iTEBD:

Main advantage of iTEBD: costs not proportional to system size, hence comparatively cheap. Main disadvantage: loss of orthogonality due to projection, without explicit reorthogonalization.

3. Improvements

1. Hasting's trick

Performing iTEBD involves inverting a singular value matrix, which could lead to the numerically unstable process of dividing by small singular values (even after truncation).

Hastings [Hastings2009 Sec IIA, Schollwock2011 Sec 7.3.2] reported a method to avoid this division by a series of contractions and SVDs. For this class I just want to make you aware of it; due to time constraints I will not go through it.

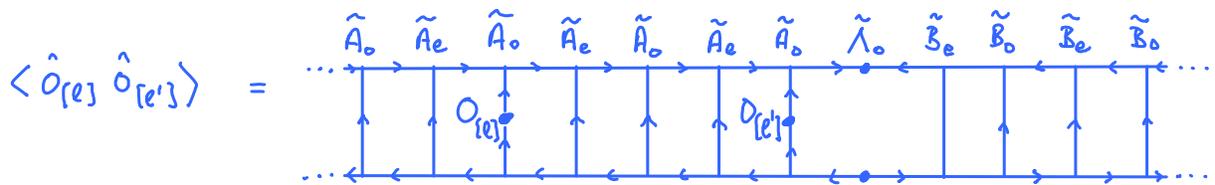
2. Orthonormalization

Correlators via transfer matrix [Schollwock2011, Sec 10.5.1]

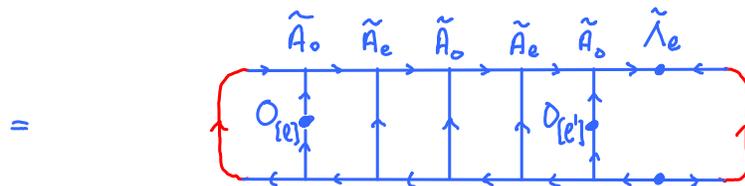
Recall that an infinite, translationally invariant MPS with two-site unit cell, expressed in the form

is called 'canonical' if Γ are left-normalized and $\tilde{\Gamma}$ are right-normalized.

Correlators can then be computed using transfer matrix methods:

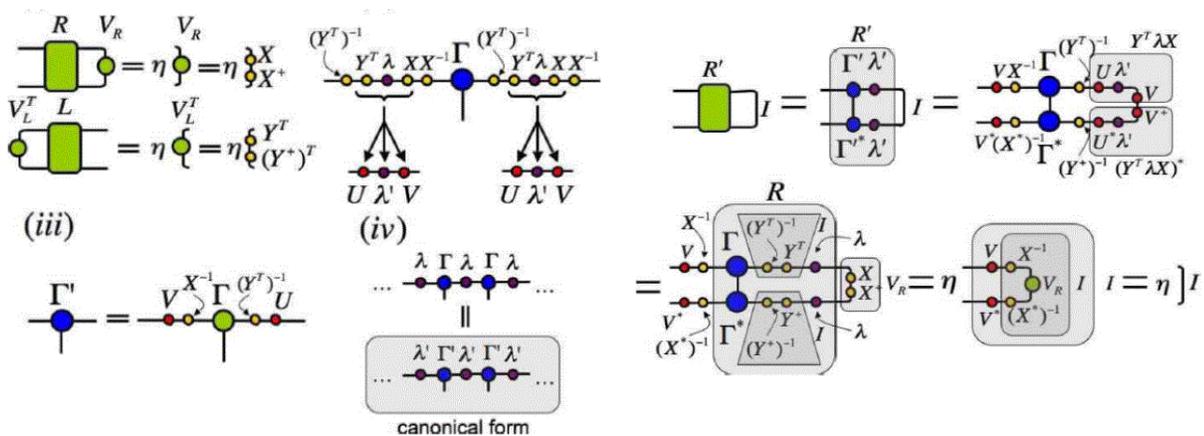


close zippers



Problem: iTEBD (including Hastings' version) yields infinite MPS that are not in canonical form, due to loss of orthogonality. It is possible to restore orthogonality (albeit at the cost of inverting singular value matrices).

[Orus2008, Schollwock2011 Sec 10.5]



DMRG II: tDMRG, purification for finite temperature

1. Time-dependent DMRG (tDMRG)

[Daley2004, White2004]

Invented in 2004 by Daley, Kollath, Schollwöck, Vidal, and independently by White,

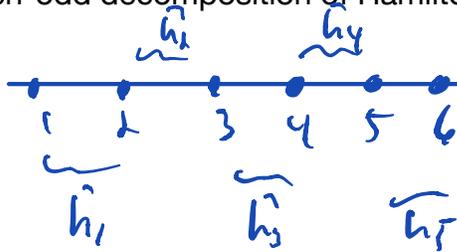
Feiguin. *Precedes iTEBD (2007)*

Goal: to compute

$$|\psi(t)\rangle = e^{-i\hat{H}t} |\psi\rangle$$

Time-evolution operator for nearest-neighbor interactions

Even-odd decomposition of Hamiltonian:



$$\hat{H} = \sum_l \hat{h}_l = \hat{H}_o + \hat{H}_e$$

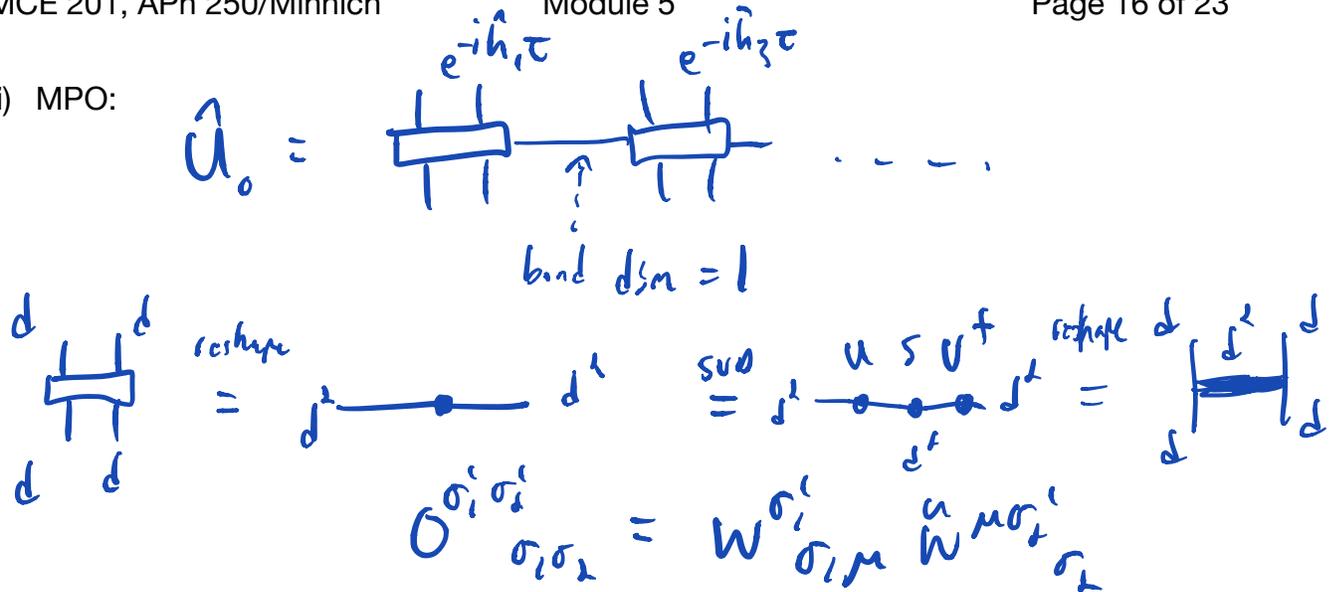
Trotterize: $t = \tau N_t$

$$\hat{U}(t) = e^{-i\hat{H}t} = \left(e^{-i\tau(\hat{H}_o + \hat{H}_e)} \right)^{N_t} \approx \left(e^{-i\tau\hat{H}_e} e^{-i\tau\hat{H}_o} + o(\tau^2) \right)^{N_t}$$

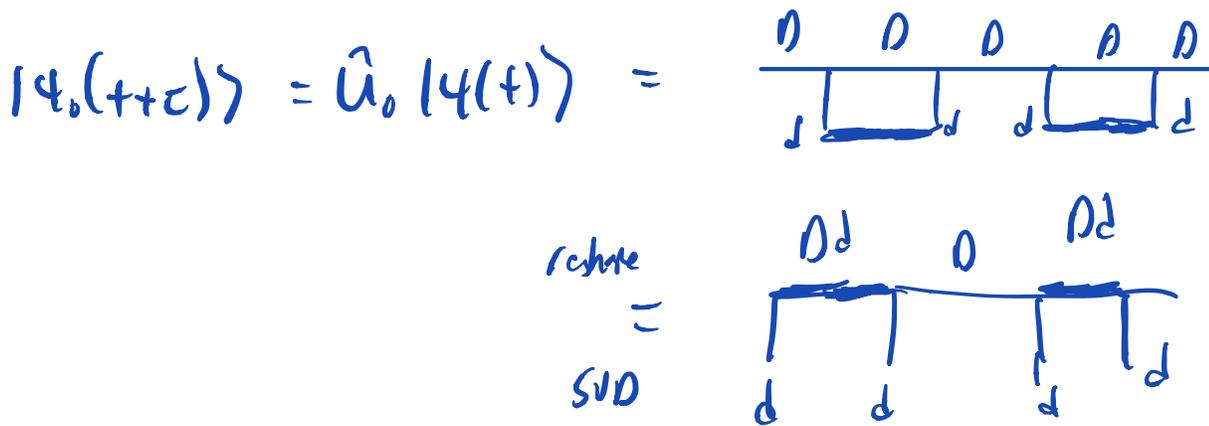
Time-evolution protocol [Schollwöck2011, Sec 7.1-7.3]

Construct MPO representations for $\hat{U}_o + \hat{U}_e$, compute $|\psi(t+\tau)\rangle = \hat{U}_e \hat{U}_o |\psi(t)\rangle$

(i) MPO:



(ii) Evolve



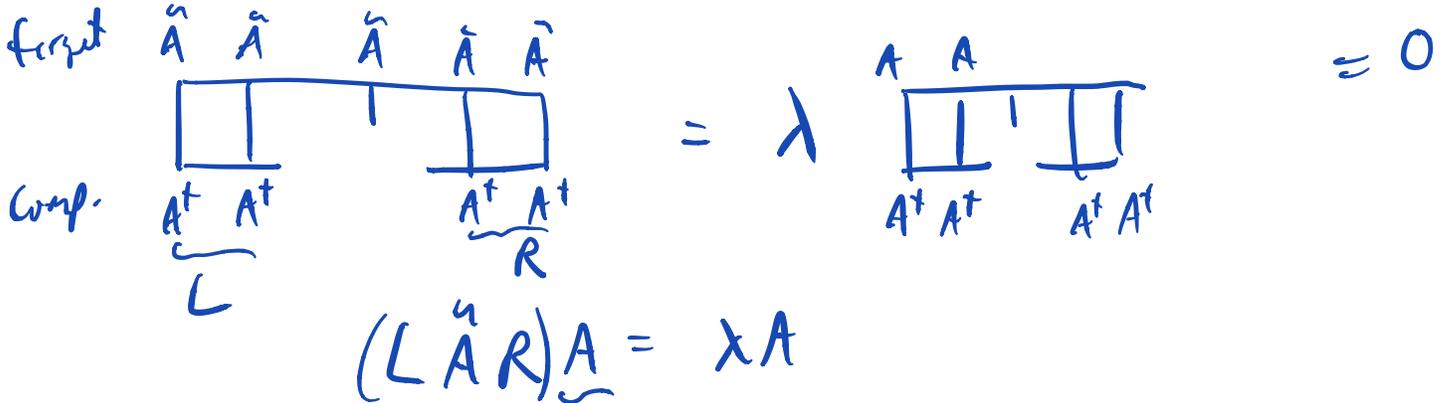
(iii) Compress: either 'variationally' (global) or 'bond by bond' (local)

Variational compression: First apply full MPO for \hat{U}_0 to entire chain. Then variationally minimize

$$\| \underbrace{|\psi(t+\tau)\rangle}_{Dd} - \underbrace{|\psi_{\text{compressed}}\rangle}_D \|^2$$

This yields optimal (in variational sense) way to compress $|\psi_{\text{full}}\rangle$ to $|\psi_{\text{compressed}}\rangle$ with given bond dimension.

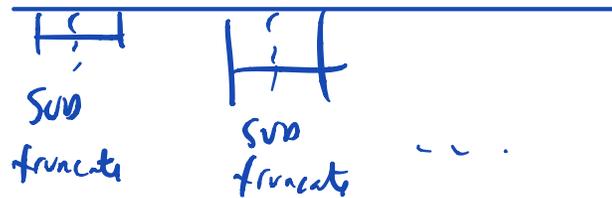
Explicitly: $\frac{\partial}{\partial A^T_{(c)}} \left[\langle \psi_{\text{compressed}} | \psi_{\text{target}} \rangle - \lambda \langle \psi_{\text{comp}} | \psi_{\text{comp}} \rangle \right]$



Sweep back and forth until overlap $\langle \psi_c | \psi_t \rangle$ no longer changes. Then apply \hat{U}_c .

Bond by bond compression

Apply \hat{U}_b to bond 1-2,



then reshape, SVD, truncate,

repeat for bond 3-4, 5-6, etc

This approach keeps bond dimensions low throughout, hence is cheaper. However, some interdependence of successive truncations may enter in, hence variational compression is cleaner.

The difference between variational and bond-by-bond compression becomes negligible for sufficiently small τ , because then the state does not change much during a time step anyway, thus the effect of truncation is less.

With bond-to-bond compression, there is no need to split $\hat{H} = \hat{H}_s + \hat{H}_e$, $\hat{U} = \hat{U}_e \hat{U}_s$

Instead, Trotterize as follows:

$$e^{-i\hat{H}\tau} = e^{-i\hat{h}_{N-1}\tau} \dots e^{-i\hat{h}_1\tau} + O(\tau^2)$$

[First order Trotter]

$$e^{-i\hat{H}\tau} = \left(e^{-i\hat{h}_1\tau/2} \dots e^{-i\hat{h}_{N-2}\tau/2} \right) e^{-i\hat{h}_{N-1}\tau} \left(e^{-i\hat{h}_{N-2}\tau/2} \dots e^{-i\hat{h}_1\tau/2} \right) + O(\tau^3)$$

or [Second order Trotter]

Error analysis

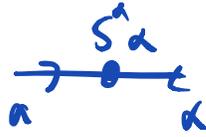
$\epsilon_{\text{Trotter}}$ (error per step) \times (# of steps) = $\frac{1}{\tau^{n+1}} \cdot \frac{t}{\tau} = \tau^n \cdot t$

linear in time; controllable by reducing τ

Truncation error due to truncation of bond dimensions:

$$\epsilon_{\text{trunc}} \sim e^t, \text{ grows exponentially!}$$

Reason: under time evolution, state becomes increasingly more entangled, on a bond entanglement entropy is



$$S_E = -\sum_{\alpha} (S_\alpha)^k \ln(S_\alpha^2)$$

This is maximal if all singular values on bond are equal,

$$(S_\alpha)^k = \frac{1}{D} \Rightarrow S_E \leq \ln_k D$$

If Hamiltonian $H(t)$ is changed abruptly (quench) such that global energy changes extensively, then

$$S(t) \leq S(0) + ct$$

[For less dramatic changes (e.g. local perturbation), entanglement growth is slower but still significant.]

Bond dimension needed to encode entanglement entropy S_E is given by $D(t) \geq 2^{S(t)}$

If, however, bond dimension D is held fixed during time evolution, errors will grow exponentially.

A quantitative error analysis has been performed by [Gobert2005] on the exactly solvable XX model:

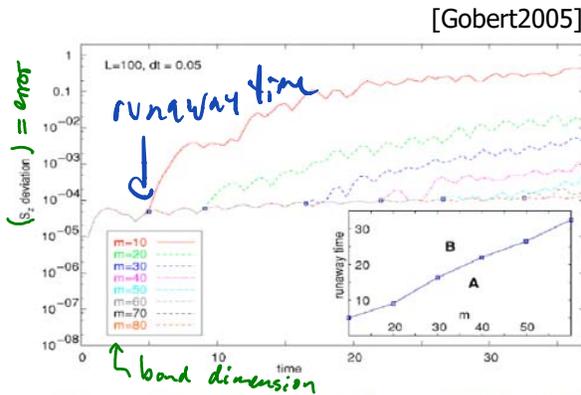


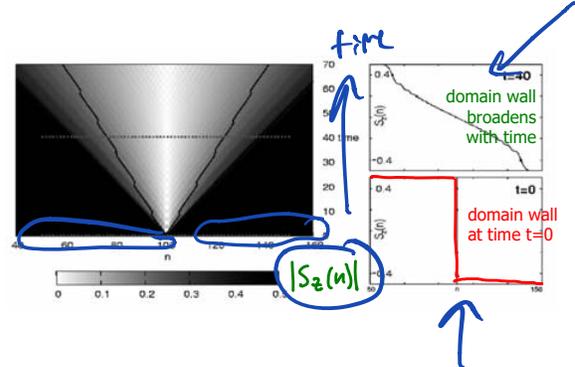
FIG. 6. Magnetization deviation $\Delta M(t)$ as a function of time for different numbers m of DMRG states. The Trotter time interval is fixed at $dt=0.05$. Again, two regimes can be distinguished: For early times, for which the Trotter error dominates, the error is slowly growing (essentially linearly) and independent of m (regime A); for later times, the error is entirely given by the truncation error, which is m -dependent and growing fast (almost exponential up to some saturation: regime B). The transition between the two regimes occurs at a well-defined "runaway time" t_R (small squares). The inset shows a monotonic, roughly linear dependence of t_R on m .

$$H_{XX} = J \sum_l S_{[l]}^x S_{[l+1]}^x + S_{[l]}^y S_{[l+1]}^y \quad (22)$$

They performed quench, with initial state

$$|4\rangle_{J=0} = \uparrow \uparrow \uparrow \uparrow \downarrow \downarrow \downarrow \downarrow$$

For $t > 0$: $J \neq 0$, domain wall widens...



2. Finite temperature: purification

[Verstraete2004, Schollwöck2011 Sec 7.2.1]

General quantum-mechanical density matrix for a mixed state,

$$\hat{\rho} = \sum_{\mu, \nu} |\mu\rangle_\rho \rho_{\mu\nu} \langle \nu|$$

↑ physical

has three defining properties:

(1) Hermiticity: $\hat{\rho}^\dagger = \hat{\rho}$

(2) Positivity: Eigenvalues are non-zero ^{negative}

$$\hat{\rho}_{\text{diagonalized}} = \sum_\alpha |\alpha\rangle_\rho \rho_\alpha \langle \alpha|$$

$\rho_\alpha \geq 0$

(3) Normalized $\text{Tr} \hat{\rho} = 1 \Leftrightarrow \sum_{\alpha} p_{\alpha} = 1$

Expectation values: $\langle \hat{O} \rangle = \text{Tr}(\hat{\rho} \hat{O}) \left[\frac{\text{Tr}(\hat{\rho} \hat{O})}{\text{Tr}(\hat{\rho})} \right]$
 if $\hat{\rho}$ not normalized

Purification

Can we represent $\hat{\rho}$ in terms of a pure state?

Yes: double Hilbert space by introducing an 'auxiliary' state for each physical state, and define 'purified state':

$$|4\rangle = \sum_{\alpha} |\alpha\rangle_a |\alpha\rangle_p \sqrt{p_{\alpha}} \in \mathcal{H}_a \otimes \mathcal{H}_p$$

↑
auxiliary

This can be viewed as a Schmidt decomposition of a pure state in doubled Hilbert space.

Norm yields trace:

$$\text{Tr} \hat{\rho}_p \stackrel{?}{=} \langle 4|4\rangle = \sum_{\alpha, \alpha'} \sqrt{p_{\alpha'}} \langle \alpha' | \underbrace{\langle \alpha' | \alpha \rangle_a}_{\delta^{\alpha' \alpha}} |\alpha\rangle_p \sqrt{p_{\alpha}} = \sum_{\alpha} p_{\alpha} = \text{Tr} \rho_p$$

Tracing out auxiliary state space from $|4\rangle\langle 4|$ (a pure DM in doubled Hilbert space) yields physical density matrix $\hat{\rho}_p$ (a mixed DM in physical Hilbert space)

$$\begin{aligned} \text{Tr}_a |4\rangle\langle 4| &= \sum_{\beta} \sum_{\alpha, \alpha'} \underbrace{\langle \beta | \alpha' \rangle_a}_{\delta^{\beta \alpha'}} |\alpha'\rangle_p \sqrt{p_{\alpha'}} \sqrt{p_{\alpha}} \underbrace{\langle \alpha | \beta \rangle_a}_{\delta^{\alpha \beta}} \\ &= \sum_{\alpha} |\alpha\rangle_p p_{\alpha} \langle \alpha| = \hat{\rho}_p \end{aligned}$$

Purified-state expectation values in doubled Hilbert space yield thermal averages in physical space

$$\langle \Psi | I_a \otimes \hat{O}_p | \Psi \rangle = \sum_{\alpha' \alpha} \sqrt{p_{\alpha'}} \langle \alpha' | \langle \alpha' | I_a \otimes \hat{O}_p | \alpha \rangle_a | \alpha \rangle_p \sqrt{p_{\alpha}}$$

$$= \sum_{\alpha} \langle \alpha | \hat{O} | \alpha \rangle_p p_{\alpha} = \text{Tr}_p \hat{\rho}_p \hat{O}_p = \langle \hat{O}_p \rangle$$

If $\hat{\rho}$ is not normalized, use

$$\frac{\langle \Psi | I_a \otimes \hat{O}_p | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\text{Tr} \hat{\rho}_p \hat{O}_p}{\text{Tr} \hat{\rho}_p} = \langle \hat{O}_p \rangle$$

Thermal density matrix

Thermal density matrix is described by

$$\hat{\rho}_\beta = e^{-\beta \hat{H}_p} = \sum_{\alpha} |\alpha\rangle_p e^{-\beta E_{\alpha}} \langle \alpha|_p$$

Not normalized:

$$\text{Tr}_p \hat{\rho}_\beta = \sum_{\alpha} e^{-\beta E_{\alpha}} = Z(\beta) = \text{partition function}$$

Purified version:

$$|\Psi\rangle_p = \sum_{\alpha} |\alpha\rangle_a |\alpha\rangle_p \frac{e^{-\beta E_{\alpha}/2}}{\sqrt{p_{\alpha}}} = e^{-\beta \hat{H}_p/2} \sum_{\bar{\sigma}} |\bar{\sigma}\rangle_a |\bar{\sigma}\rangle_p$$

$$|\Psi_0\rangle = \sum_{\bar{\sigma}} |\bar{\sigma}\rangle_a |\bar{\sigma}\rangle_p = \sum_{\bar{\sigma}} |\sigma_N\rangle_a |\sigma_N\rangle_p \dots |\sigma_1\rangle_a |\sigma_1\rangle_p$$

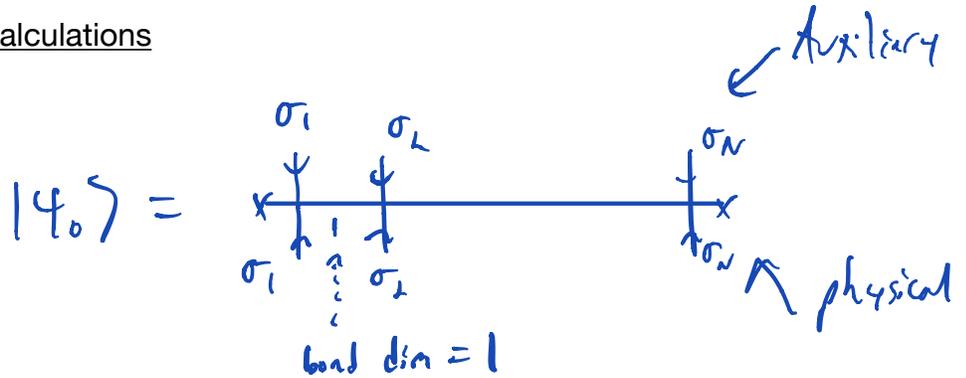
= product state, with each factor describing maximal a-p entangled at site l

$$= \prod_{l=1}^N \left(\sum_{\sigma_l} |\sigma_l\rangle_a |\sigma_l\rangle_p \right)$$

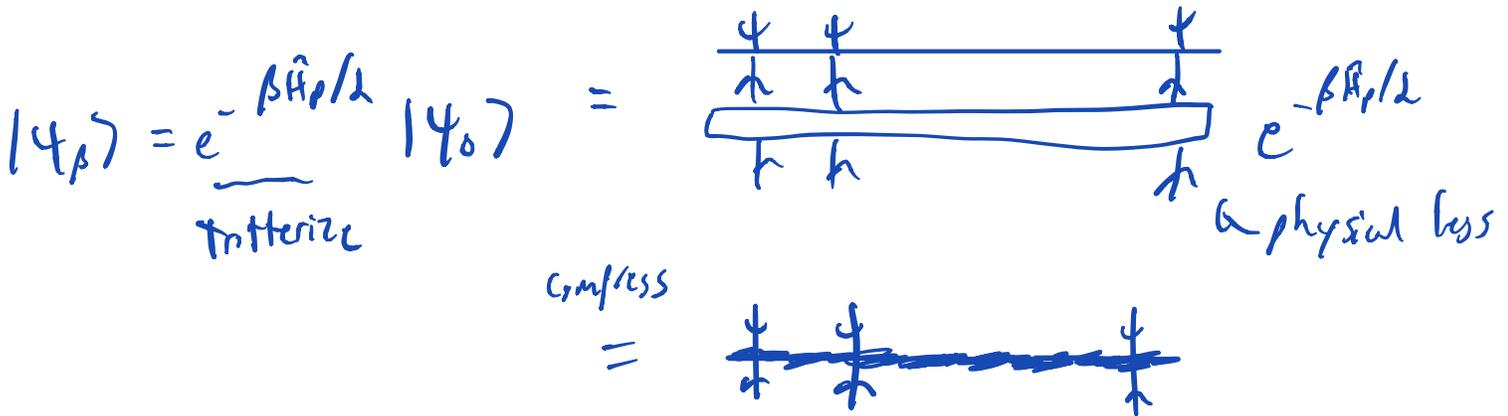
Note: at $T=0$, i.e. $\beta=0$ we have $|\psi\rangle=|\psi_0\rangle$ (all states $|\sigma\rangle$ are equally likely)

Protocol for finite-T DMRG calculations

Start from pure
product state in
doubled Hilbert space:



Perform imaginary-time evolution over a 'time' $\beta/2$, acting only on physical space:



For thermal averages, trace out auxiliary space:

$$\langle \hat{O}_{(p)} \rangle = \frac{\langle \Psi_\beta | I_A \otimes \hat{O}_p | \Psi_\beta \rangle}{\langle \Psi_\beta | \Psi_\beta \rangle}$$

