Simons Summer School Lecture
MPS & DMRG

E.M. Stoudenmire
June 2014

Tensor network wavefunction methods

- Old wisdom: wavefunction too complicated – ED, even with Lanczos/targeted methods scales exponentially
  → map to few/single particle problem: DFT, mean-field
  → parameterize w/ small number of parameters
  variational method – not systematic/unbiased or controlled
  → map to "classical" problem & sample (QMC of path integral)
- Recent wisdom (quantum info.):
  → entanglement: amount of info. needed about subsystem B to know/state of subsystem A

(A|B)

entangled = complicated / less local

But we are discovering that "physical" states esp. ground states are only "lightly" entangled.

Leading to new tools for storing, manipulating, optimizing wavefunctions.

Rather complete knowledge in 1d

- Matrix product states

Efficient, convenient way of representing all 1d wavefunctions having a correlation length

Can represent gapless (power-law correlated) states controllably too

Only bias is low entanglement (very weak for ground states)

Comes with natural optimization algorithm: DMRG

Recommend: Schollwöck, Ann Phys. 326, 96 (2011) [start @ p. 109]

Setting: 1d lattice models

Ex #1: Heisenberg spin chain $S=1/2$ $H = \sum_j \vec{S}_j \cdot \vec{S}_{j+1}$

Local Hilbert space is single spin 1/2 $|s\rangle \in \{|\uparrow\rangle, |\downarrow\rangle\}$

General wavefunction of the form

$$|\psi\rangle = \sum_{s_1 s_2 \cdots s_N} \psi^{s_1 s_2 \cdots s_N} |s_1\rangle \otimes |s_2\rangle \otimes |s_3\rangle \cdots \otimes |s_N\rangle$$

On two sites ($N=2$), ground state is singlet

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} |\uparrow\rangle|\downarrow\rangle - \frac{1}{\sqrt{2}} |\downarrow\rangle|\uparrow\rangle \; ; \; \psi^{s_1 s_2} = \begin{array}{c} \\ \uparrow \\ \downarrow \end{array} \begin{array}{c} \begin{array}{cc} \uparrow & \quad\downarrow \end{array} \\ \begin{bmatrix} 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix} \end{array}$$

For $N>2$, more complicated singlet w/ quasi long range (power law) antiferromagnetic correlations

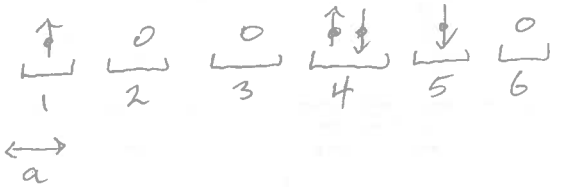
Ex #2: 1d "electronic structure" / electrons in continuum

Can discretize continuum using grid basis

$$\psi(x\sigma, x'\sigma', x''\sigma'') \rightarrow |\psi\rangle = \sum_{s_1 s_2 \cdots} \psi(s_1, s_2, s_3, \ldots, s_N) \, |s_1\rangle|s_2\rangle \cdots |s_N\rangle$$

$$s_j \in \{0, \uparrow, \downarrow, \uparrow\downarrow\}$$

Like binning real space

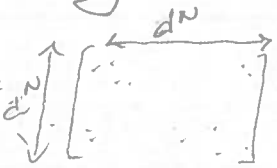


$\rightarrow$ amplitude $\psi(\uparrow, 0, 0, \uparrow\downarrow, \downarrow, 0)$

$$H = -\frac{1}{2}\sum_\sigma \int_x \psi_\sigma^\dagger(x) \partial_x^2 \psi_\sigma(x) \rightarrow H_{grid} = \frac{-1}{2a^2} \sum_{j\sigma} \left(c_{j\sigma}^\dagger c_{j+1\sigma} - 2 c_{j\sigma}^\dagger c_{j\sigma} + c_{j+1\sigma}^\dagger c_{j\sigma} \right)$$

Getting the ground state wavefunction:

① Try putting Hamiltonian as a matrix into a full diagonalization code



$H = \begin{bmatrix} \cdots & \cdots \\ \vdots & \ddots \end{bmatrix} \rightarrow U E U^\dagger$ cost: $(d^N)^3 = d^{3N}$!!

② "Target" low-energy states - Lanczos, Davidson

Work involved is multiplying $H|\psi_n\rangle = |\psi_{n+1}\rangle$

$\{|\psi_1\rangle, H|\psi_1\rangle, H^2|\psi_1\rangle, \ldots, H^M|\psi_1\rangle\}$

↑ M of the order of 10-100

Use sparsity of H, $H|\psi\rangle$ cost $\sim k d^N$, d^{3N}, still $\geq d^N$

Find ground state, low lying excited states in $H^n|\psi\rangle$ basis.

③ DMRG (White)

Problem with ED, even with targeting, is exponentially large Hilbert space basis

Can one identify a smaller basis sufficient for targeting low-energy states?

The "job" of the ground state is to minimize the energy; Hamiltonian is sum of local operators

Which part of the ground state wavefunction do local operators (e.g. terms in H) see?

Consider expectation value of operator $\hat{n}_2$ (for concreteness) measuring density on site 2.

$$\langle\psi|\hat{n}_2|\psi\rangle = \sum_{s_1 s_2 s_2' s_3 \cdots} \bar{\psi}_{s_1 s_2 s_3 \cdots s_N} n^{s_2}{}_{s_2'} \psi^{s_1 s_2' s_3 \cdots s_N}$$

$$= \sum_{s_1 s_2 s_2'} n^{s_2}{}_{s_2'} \rho^{s_1 s_2'}{}_{s_1 s_2}$$

where $\rho^{s_1' s_2'}{}_{s_1 s_2} = \sum_{s_3 s_4 \cdots s_N} \bar{\psi}_{s_1 s_2 s_3 s_4 \cdots s_N} \psi^{s_1' s_2' s_3 s_4 \cdots s_N}$

ρ is called a reduced density matrix

only information about $|\psi\rangle$ we need to understand properties on sites $1 \neq 2$.

Reduced density matrices Hermitian & positive semi-definite

Can diagonalize as $\rho^{s_1 s_2'}{}_{s_1' s_2'} = U^{s_1 s_2}{}_{\alpha} p^{\alpha}{}_{\alpha'} U^{\dagger \alpha'}{}_{s_1' s_2'}$

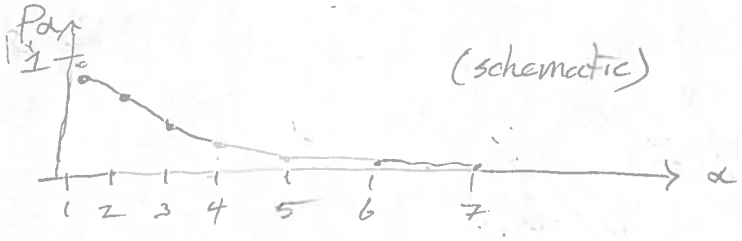
$$p^{\alpha}{}_{\alpha'} = \begin{bmatrix} p_1 & & \\ & p_2 & \\ & & p_3 \cdots \end{bmatrix}; \sum p_\alpha = 1 \quad ; \quad U = \begin{bmatrix} | & | & \\ u_1^{s_1 s_2} & u_2^{s_1 s_2} & \cdots \\ | & | & \end{bmatrix}$$

Can interpret columns of U as wavefunctions for the s_1-s_2 subsystem $|\alpha\rangle = u_\alpha^{s_1 s_2} |s_1 s_2\rangle$ and p_α as probabilities for finding these sites in the state $|\alpha\rangle$.

In the density-matrix eigenbasis ("Schmidt basis") expectation values of local operators take the form

$$\langle\psi|\hat{n}_2|\psi\rangle = \sum_\alpha \langle\alpha|\hat{n}_2|\alpha\rangle\, p_\alpha$$

For gapped 1d systems, the p_α fall very fast (# number of large p_α independent of system size ⇒ Hastings)



⟷ (jump to examples - next page)

Suggests an approximation: truncate (project out) states $|\alpha\rangle$ having p_α less than some cutoff

Guaranteed to preserve all local properties up to an error proportional to the sum of discarded probabilities.

Origin of the term "density matrix renormalization group"

Practitioners often talk about the "number of states kept" in DMRG - refers to keeping first m states, discarding (projecting out) the rest.

Basic strategy: apply such an approximation for <u>all</u> bipartitions of the system, not just $s_1 s_2 | s_3 \ldots$

Result will be a matrix product state.

(From this point of view, wavefunction is a chain of projectors.)

Gives highly efficient (optimal in some cases) compression scheme for 1d many-body states.

Density matrix examples

Ex #1: singlet

$$\psi^{s_1 s_2} = \begin{matrix} \uparrow \\ \downarrow \end{matrix}\begin{bmatrix} \uparrow & \downarrow \\ 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix} \rightarrow \rho^{s_1}{}_{s_1'} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

Ex #2:

$$\psi^{s_1 s_2} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}(\uparrow\uparrow + \uparrow\downarrow + \downarrow\uparrow + \downarrow\downarrow) \qquad |\rightarrow\rangle\,|\rightarrow\rangle$$

$$\rho^{s_1 s_1'} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}$$

To carry out the compression, first truncate U:

$$U = \begin{bmatrix} | & | & | & | \\ U_1^{s_1 s_2} & U_2^{s_1 s_2} & \cancel{U_3^{s_1 s_2}} & \cancel{U_4^{s_1 s_2}} \\ | & | & | & | \end{bmatrix} \rightarrow \begin{bmatrix} | & | \\ U_1 & U_2 \\ | & | \end{bmatrix}$$

Rewrite wavefunction in (truncated) Schmidt basis

$$|\psi\rangle = \sum_{s_1 \cdots s_N} \psi^{s_1 s_2 s_3 \cdots s_N} |s_1 s_2 s_3 \cdots s_N\rangle \simeq \sum_{\alpha s_3 \cdots s_N} \psi^{\alpha_2 s_3 s_4 \cdots s_N} |\alpha_2\rangle\,|s_3 s_4 \cdots s_N\rangle$$

$$\psi^{\alpha_2 s_3 \cdots s_N} = \sum_{s_1 s_2} \psi^{s_1 s_2 s_3 \cdots s_N} U^{\dagger \alpha_2}_{s_1 s_2}$$

For later purpose of optimizing wavefunction to search for ground state, transform Hamiltonian terms into this basis too

$$\langle s_1 s_2 | \hat{h}_{12} | s_1' s_2' \rangle \rightarrow \underbrace{\langle \alpha | \hat{h}_{12} | \alpha' \rangle}_{\text{save in memory for later}} = \sum_{\substack{s_1 s_2 \\ s_1' s_2'}} U^{\dagger \alpha}_{s_1 s_2} \langle s_1 s_2 | \hat{h}_{12} | s_1' s_2' \rangle U^{s_1' s_2'}_{\alpha}$$

At this point, useful to introduce diagrammatic notation for tensors

Rank r tensor is a blob with r lines emanating from it.

E.g. vector - rank 1 v^i = 

matrix - rank 2 M^{ij} = 

many-body wavefunction - rank N

$\psi^{s_1 s_2 \cdots s_N}$ = 

Similar to Einstein summation, connecting two index lines implies the indices are summed or contracted

E.g. matrix - vector multiplication

$\sum_j M^i_j v^j$ $\leftrightarrow$ 

Trace of product of matrices

$Tr[AB] = \sum_{ij} A^i_j B^j_i$ $\leftrightarrow$ 

Can observe rank of the result from number of open lines remaining.

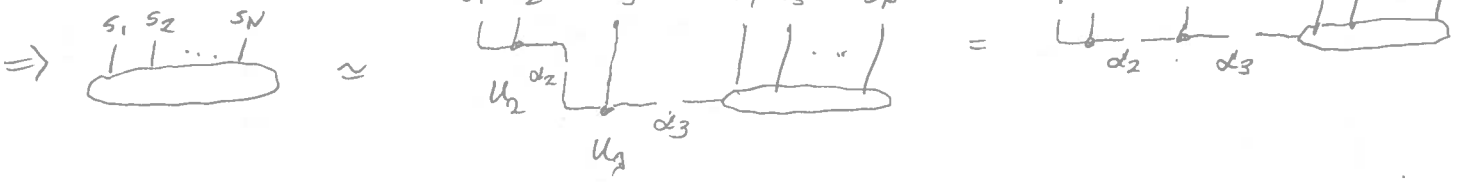
Now let's use this notation to picture the procedure of projecting out the unimportant states of the density matrix of the first two sites.

Recall $|\psi\rangle = \sum_{s_1 s_2 \cdots s_N} \psi^{s_1 s_2 \cdots s_N} |s_1 s_2\rangle |s_3 \cdots s_N\rangle \simeq \sum_{\alpha_2 s_3 \cdots s_N} \psi^{\alpha_2 s_3 \cdots s_N} |\alpha_2\rangle |s_3 \cdots s_N\rangle$

where $\psi^{\alpha_2 s_3 \cdots s_N} = \sum_{s_1 s_2} \psi^{s_1 s_2 s_3 \cdots} U^{\dagger \alpha_2}_{s_1 s_2}$

$\Rightarrow \psi^{s_1 s_2 s_3 \cdots s_N} \simeq \sum_{\alpha_2} U^{s_1 s_2}_{\alpha_2} \psi^{\alpha_2 s_3 \cdots s_N}$

MPS idea is to repeat this iteratively thinking of α_2 as a "site"



Computed U_3 by diagonalizing density matrix

$$\rho^{\alpha_2 s_3}_{\alpha_2' s_3'} = \sum_{s_4 s_5 \cdots s_N} \psi^{\alpha_2 s_3 s_4 s_5 \cdots s_N} \bar{\psi}_{\alpha_2' s_3' s_4 s_5 \cdots s_N}$$

$$= \sum_{\alpha_3 \alpha_3'} U^{\alpha_2 s_3}_{\alpha_3} \rho^{\alpha_3}_{\alpha_3'} U^{\dagger \alpha_3'}_{\alpha_2' s_3'}$$

At this point it's useful to observe we could have obtained these decompositions using the singular value decomposition (SVD) instead.

Recall SVD algorithm computes for any possibly rectangular matrix

$$M^{i\bar{j}} = U^{i}_{\alpha} D^{\alpha\alpha'} V^{\bar{j}}_{\alpha'}$$

such that

$$D^{\alpha\alpha'} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix} \qquad \lambda_\alpha\text{'s real, } \geq 0, \text{ decr. order}$$

and $U^\dagger U = I \Rightarrow \sum_i \bar{U}^{\alpha'}_{i} U^{i}_{\alpha} = \delta^{\alpha'}_{\alpha}$ ($UU^\dagger \neq I$!)

$$V V^\dagger = I \Rightarrow \sum_j V_{\alpha'}{}^{j} \bar{V}^{\alpha}_{j} = \delta_{\alpha'}{}^{\alpha} \qquad (V^\dagger V \neq I!)$$

Diagrammatically, 

and 

Importantly, provides controlled approx for M

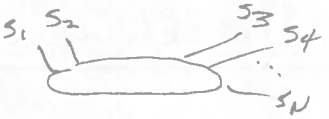
Define $\tilde{M}_n$ as product $U \tilde{D}_n V$ where

$$\tilde{D}_n = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_n & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \leftarrow \begin{array}{l} \text{set all} \\ \text{but first} \\ n \text{ to zero} \end{array}$$

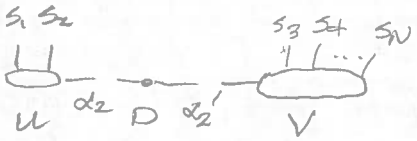
then $$\| \tilde{M}_n - M \|^2 = \sum_{\alpha=n+1}^{n_{max}} (\lambda_\alpha)^2 \quad \leftarrow \text{"discarded weight"}$$

Returning to many-body wavefunction, think of as (highly rectangular) matrix

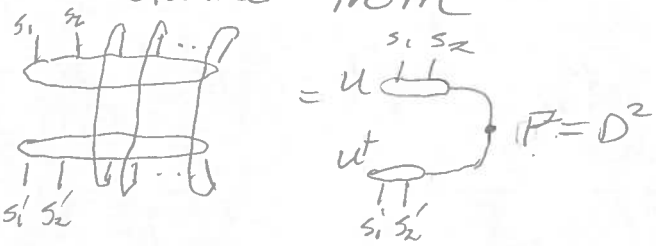
$\psi^{(s_1 s_2)(s_3 s_4 \cdots s_N)}$ =



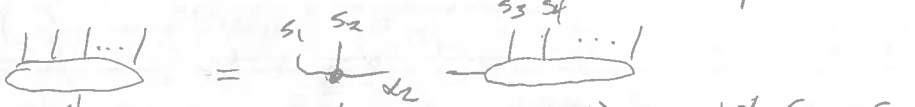
SVD =



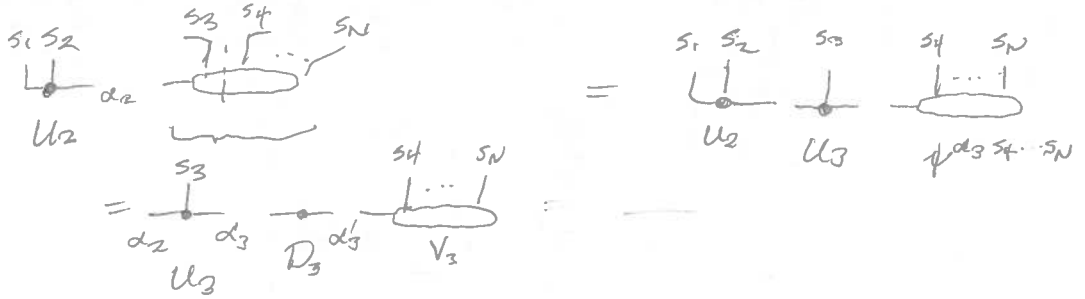
Can show that u is same as obtained from density matrix $\rho^{s_1 s_2}_{s_1' s_2'}$ =



Multiplying D into V, get same decomposition as before

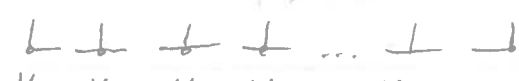


Armed w/ simpler SVD approach, apply recursively



etc.

$=$



(MPS) matrix product state in "right canonical form"

What have we accomplished?

Say when doing SVD's we truncate to at most m singular values at each step.

Obtain variational w.f. ansatz / description of particular w.f. using only $\sim N m^2$ parameters.

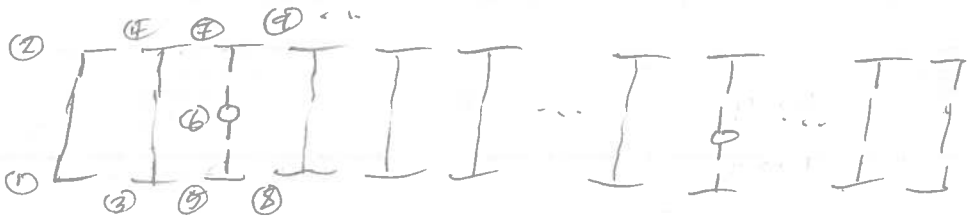
Compare to d^N for ED.

Of course, only useful if accurate for some moderate sized m.

Hastings has shown [J Stat Mech P08024 (2007)] there exists for gapped, finite range 1d Hamiltonians, some m independent of system size such that ground state well approximated by an MPS of "bond dimension" m

MPS also extremely efficient for computing properties of wf (or operations on wf)

Say we want expectation value of some product of local operators e.g. a correlator:

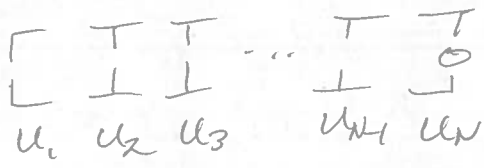


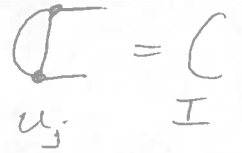
By contracting piece-by-piece can compute with $\sim N m^3$ cost.

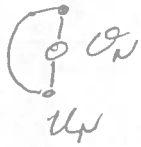
Again, compare to d^N for ED.

Can do even better for single local operator

Consider $\langle \mathcal{O}_N \rangle$ acting on site N.



All U_j have property  from SVD

$\Rightarrow \langle \psi | \mathcal{O}_N | \psi \rangle =$ 

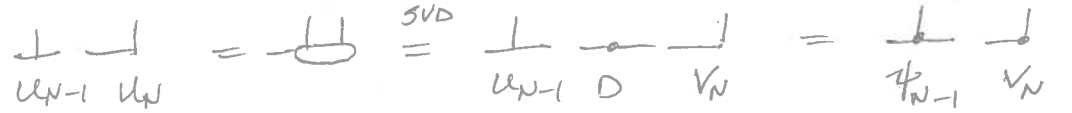
Only m cost! No factor of N.

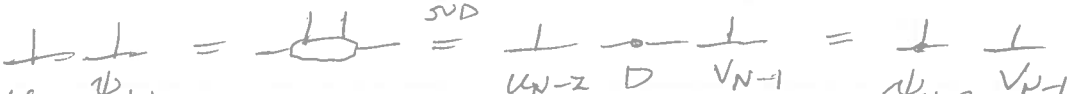
We have exploited MPS gauge

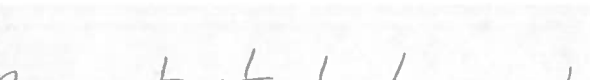
Many possible gauges $(\perp \overset{M}{\bullet})(\overset{M^{-1}}{\bullet} \perp) = \perp\ \perp$

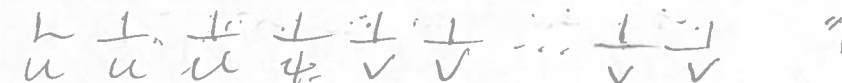
Handful of important useful ones.

Shift gauge by contracting pairs of site tensors and computing SVD:



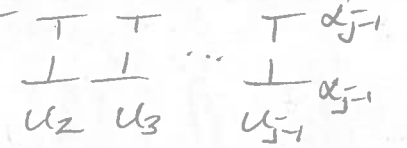


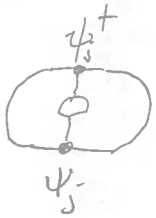
Can go from 

to  "mixed canonical"

Site j is said to be the "orthogonality center" (OC) because all its indices label orthonormal states

$$|\psi\rangle = \sum_{\alpha_{j-1} s_j \alpha_j} \psi_j^{\alpha_{j-1} s_j \alpha_j} |\alpha_{j-1}\rangle |s_j\rangle |\alpha_j\rangle$$

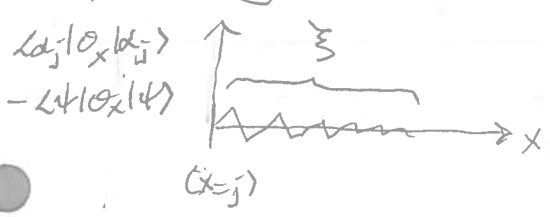
$\langle \alpha'_{j-1} | \alpha_{j-1} \rangle =$  $= \delta^{\alpha'_{j-1}}_{\alpha_{j-1}}$, similarly for $\langle \alpha_j | \alpha'_j \rangle$

Measuring expectation value $\langle \psi | O_j | \psi \rangle =$  $\sim d^2 m^2$

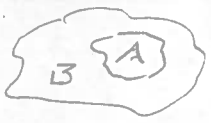
Intuitively, what are the $|\alpha_j\rangle$ states?
(Called "Schmidt states"; density matrix eigenstates.)

They are precisely the states needed to describe the many-body w.f.

For a ground state of a gapped Hamiltonian w/ correlation length ξ, can show that $|\alpha_j\rangle$ differ from ground state only within ξ of the cut



Entanglement entropy



Defined as $S = -\mathrm{Tr}[\rho_A \ln \rho_A]$

where ρ_A is the reduced density matrix describing some subsystem A.

In terms of eigenvalues $\{p_\alpha\}$ of ρ_A, $S = -\sum_\alpha p_\alpha \ln p_\alpha$

Ex 1 singlet $\psi^{s_1 s_2} = \begin{bmatrix} 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix} \Rightarrow \rho_1^{s_1}{}_{s_1'} = \begin{bmatrix} 1/2 & \\ & 1/2 \end{bmatrix} \Rightarrow \boxed{S = \ln 2}$

Ex 2 $|\rightarrow\rangle |\rightarrow\rangle$ $\psi^{s_1 s_2} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow \rho_1^{s_1}{}_{s_1'} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}$

$\Rightarrow \boxed{S = 0}$

Measures degree of correlation (statistical) between systems A & B.

For an MPS in mixed-canonical gauge, easily access entropy of left $(1, 2, \ldots, j)$ with right $(j+1, \ldots, N)$

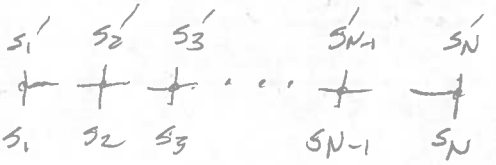
Can show eigenvalues of ρ_L (= evals of ρ_R) are squares of singular values $\lambda_\alpha^2 = p_\alpha$ from diagonal of D_j.

If bond dimension of MPS is m, this implies a maximum entropy $\underline{\log(m)}$

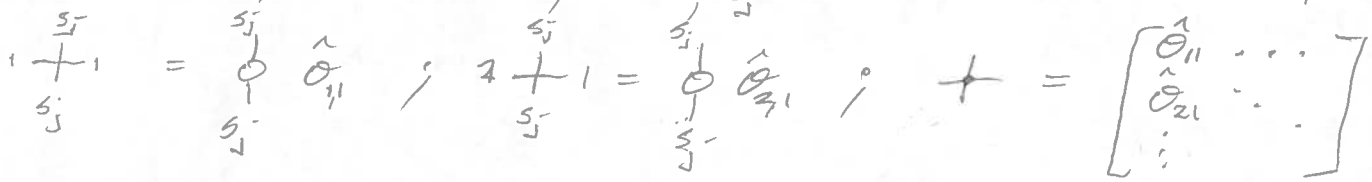
MPS naturally obey the boundary law ("area law")

Before discussing DMRG algorithm, useful to construct matrix product operator (MPO) for the Hamiltonian.

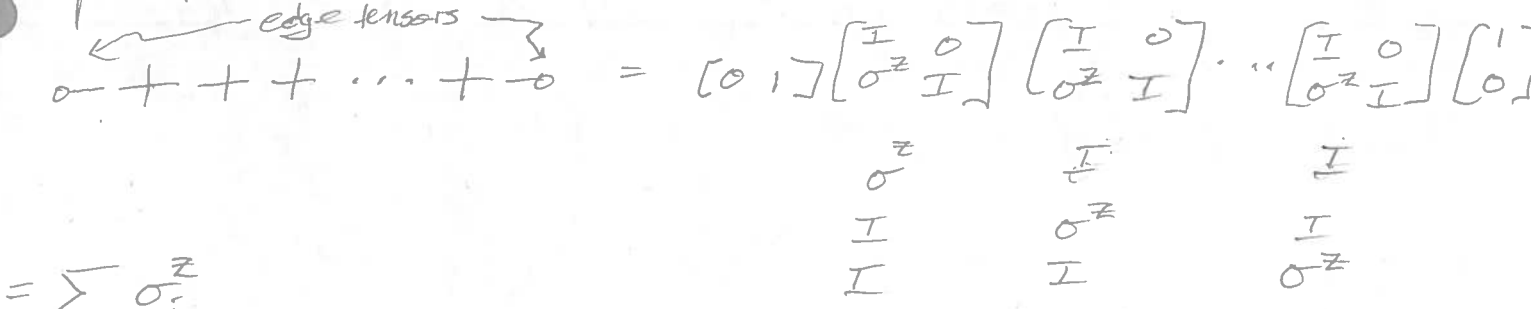
Tensor network which mimics MPS but for operato



Can view each tensor  as matrix of operators



Simple to construct sums of finite-range operators in 1d



$$= \sum_j \sigma_j^z$$

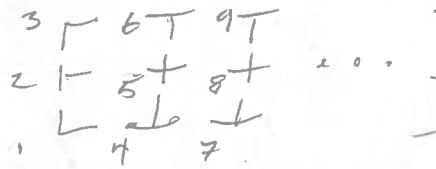
$$[0\ 0\ 1]\begin{bmatrix} I & 0 & 0 \\ \sigma^z & 0 & 0 \\ -h\sigma^x & \sigma^z & I \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ \sigma^z & 0 & 0 \\ -h\sigma^x & \sigma^z & I \end{bmatrix}\cdots\begin{bmatrix} I & 0 & 0 \\ \sigma^z & 0 & 0 \\ -h\sigma^x & \sigma^z & I \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

σ^z σ^z I
I σ^z I
$-h\sigma^x$ I I

σ^z

$$= \sum_j \sigma_j^z \sigma_{j+1}^z - h\sum_j \sigma^x$$

MPO makes it convenient & efficient to compute e.g. total energy



$\sim m^3 N$

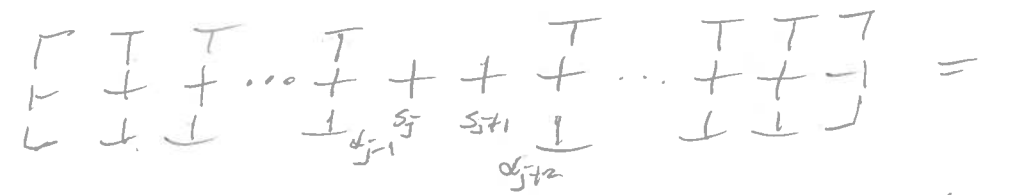
Also convenient for formulating DMRG algorithm

Recall if OC'd site j, $\langle u_1 u_2 u_3 \cdots u_{j-1} | \psi \rangle = C$

$\langle v_{j+1} v_{j+2} \cdots v_N | \psi \rangle = \Psi$

So $\langle u_1 u_2 \cdots u_{j-1} \alpha_j |$ changes from lattice basis to $|\alpha_j\rangle$ basis & projects into first m α_j's

Hamiltonian in $|\alpha_{j-1}\rangle |s_j\rangle |s_{j+1}\rangle |\alpha_{j+2}\rangle$ basis:

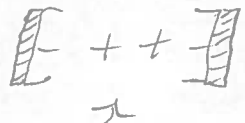


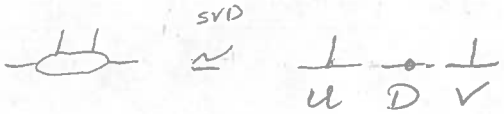
$\sim H^{\alpha_{j-1} s_j' s_{j+1}' \alpha_{j+2}'}_{\alpha_{j-1} s_j s_{j+1} \alpha_{j+2}}$

Idea of ground-state DMRG is to compute ground state of $\tilde{H}$ at each bond, use to update wavefunction

Importantly, if MPS already approx's ground state well, is fixed point of this procedure. If not, hope is that updating w.f. this way enriches matrix product basis i.e. obtain better set of states $|\alpha_j\rangle$.

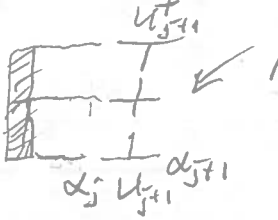
Steps of DMRG

1. Given "projected Hamiltonian" $\tilde{H}$ =  and initial wavefunction 
find grand state (approximately) of $\tilde{H}$ * using iterative exact diag. code (Lanczos, Davidson,

2. SVD new approx ground state to restore MPS form


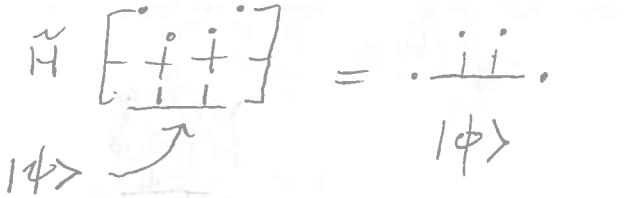
Truncate to only m states (m singular values, set rest to zero)
to control cost of calculation
[This is the reason to expose two sites.]

3. If next two sites are to the right, say, contract (multiply) D with V to shift gauge to the right.

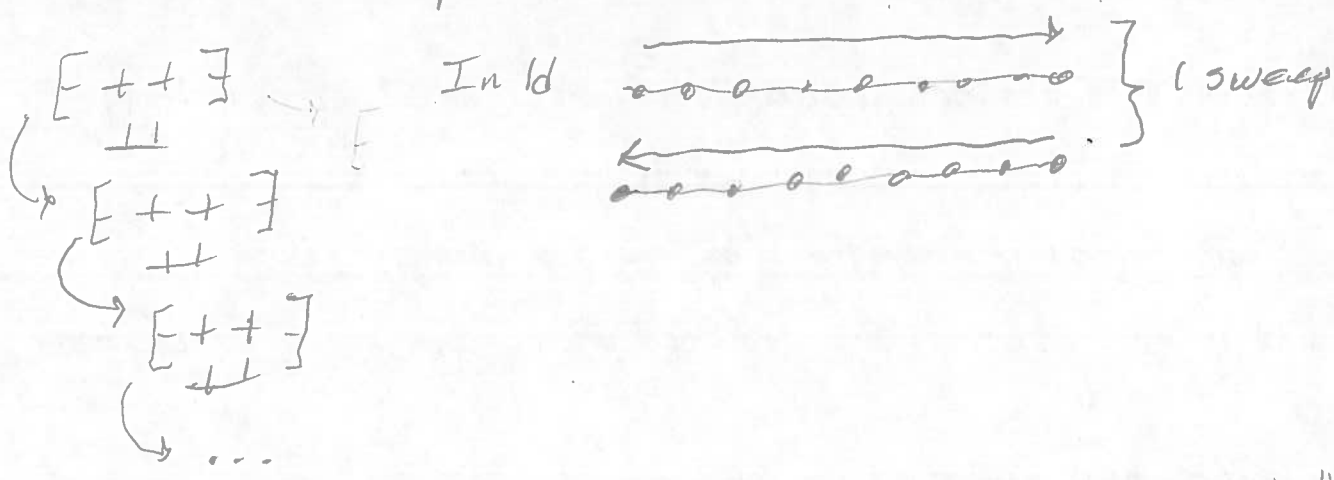

4. Build $\tilde{H}_{\alpha_{j+1} s_{j+1} s_{j+2} \alpha_{j+3}}$ (or pieces of it) by "growing" left edge tensor (if exposing next two sites to the right)


Recover previous right edge tensor saved in memory


---

* Diagrammatically, $\tilde{H}|\psi\rangle = |\phi\rangle$ looks like


Leads to "sweeping" algorithm



For most "normal" problems (1d, short-range lattice models) convergence is exponentially fast in number of sweeps over system.

Other topics

- MPO's for long-range
- collapsing an MPS: perfect sampling (& METTS)
- using DMRG for quasi-2d