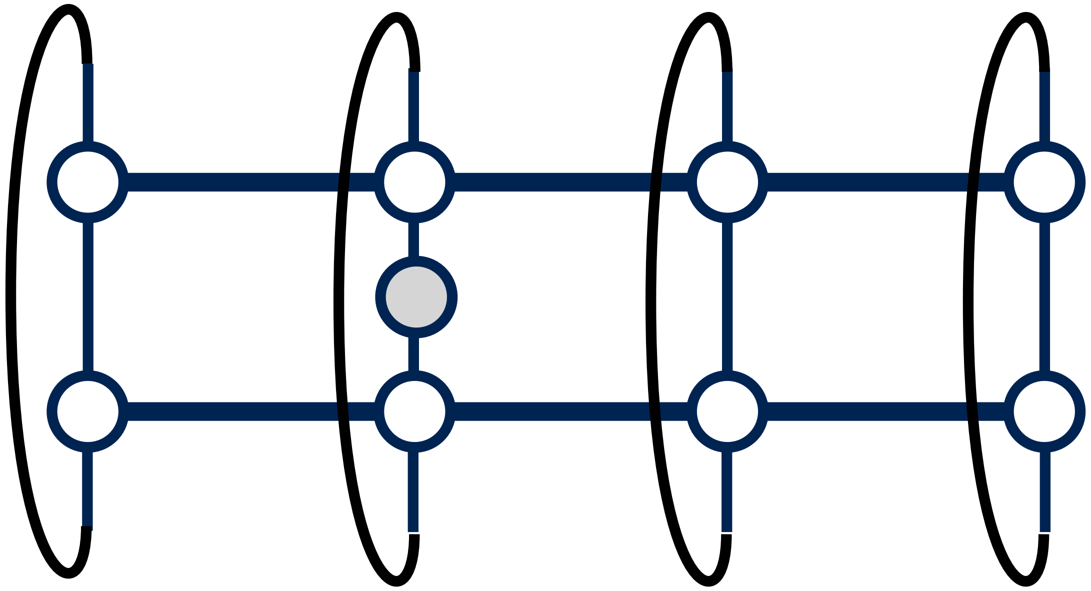
Tensor Networks and Applications



E.M. Stoudenmire

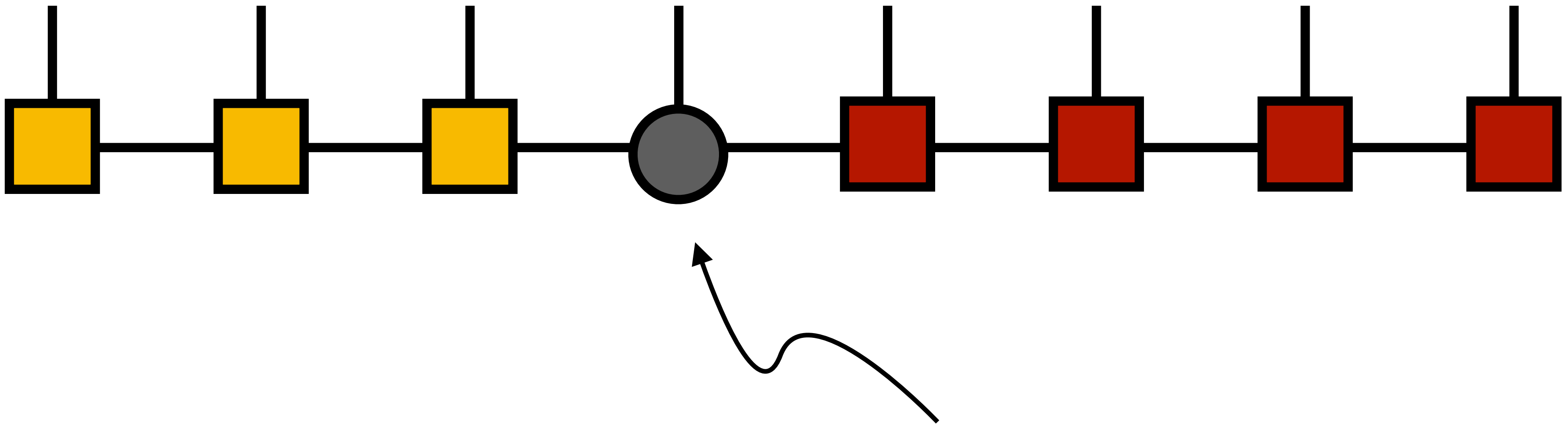
Apr 2018 - Sao Paulo

FLATIRON INSTITUTE

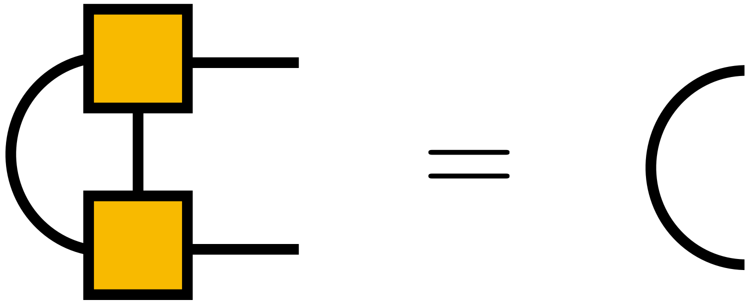
SIMONS FOUNDATION

Review of Previous Lecture

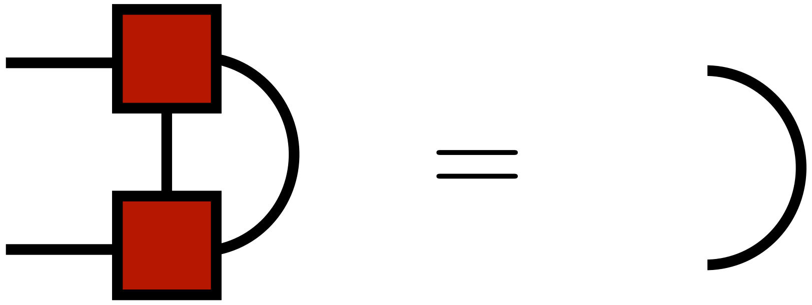
'Gauging' an MPS bestows it with useful properties



"orthogonality center" site

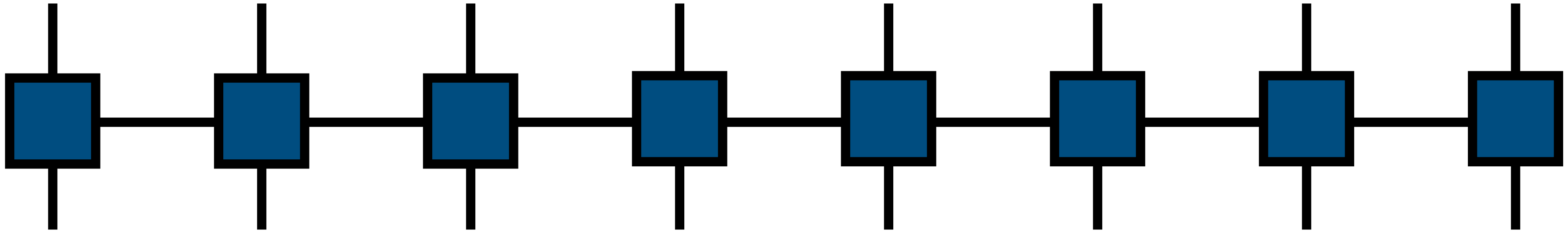


"left orthogonal"



"right orthogonal"

MPOs can compactly and generically represent Hamiltonians



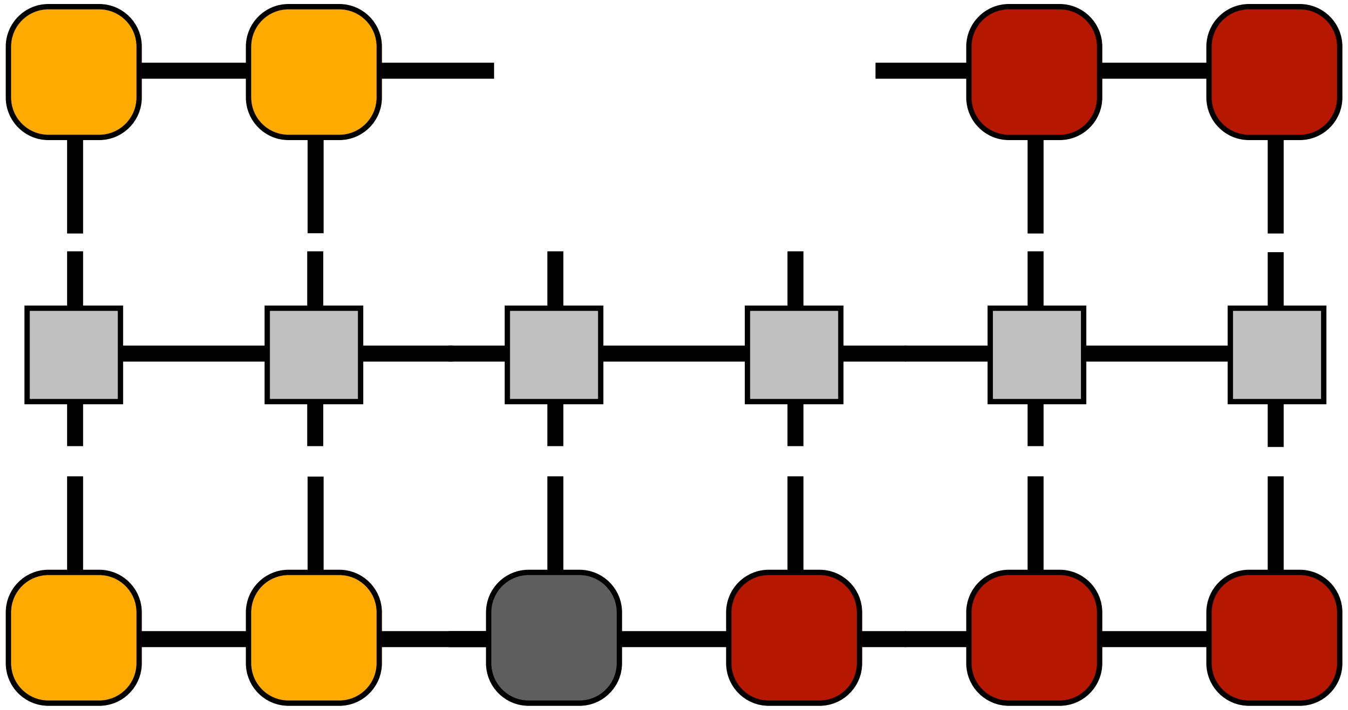
$$\begin{bmatrix} I_j & & & & \\ S_j^+ & & & & \\ S_j^- & & & & \\ S_j^z & & & & \\ 0 & \frac{1}{2}S_j^- & \frac{1}{2}S_j^+ & S_j^z & I_j \end{bmatrix}$$

Heisenberg model

$$H = \sum_j \mathbf{S}_j \cdot \mathbf{S}_{j+1}$$

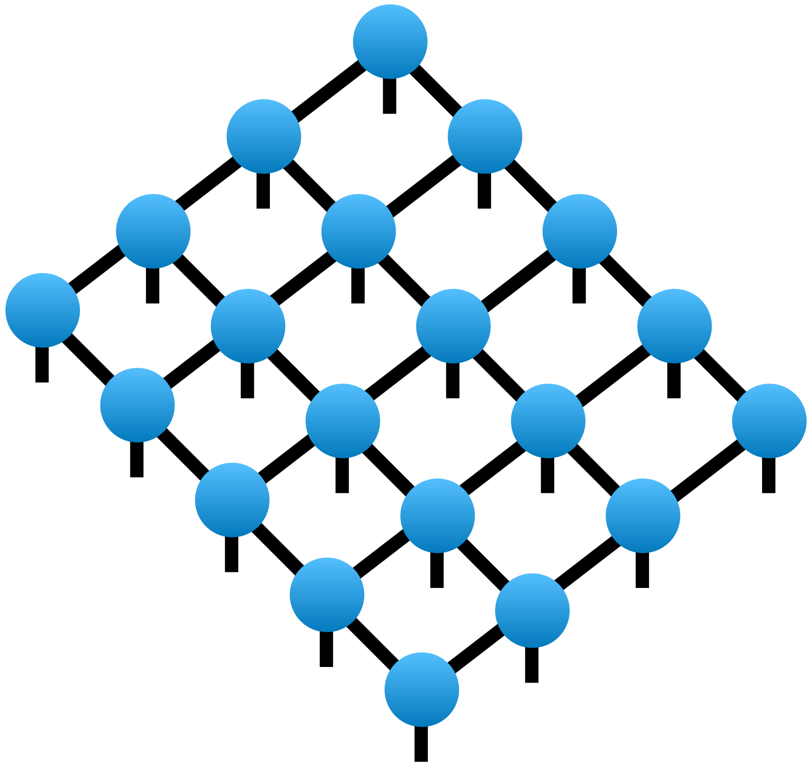
DMRG algorithm for optimizing MPS

$$\tilde{H}|\tilde{\Psi}\rangle = E|\tilde{\Psi}\rangle$$



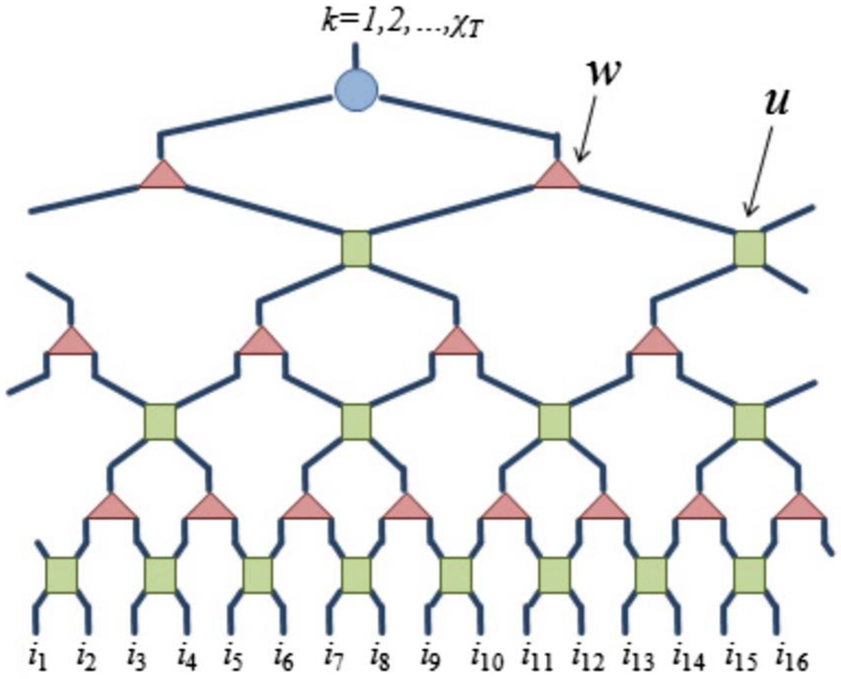
Beyond Matrix Product States

Besides matrix product state network, other very interesting networks are PEPS and MERA



PEPS

(2D systems)



MERA

(critical systems)

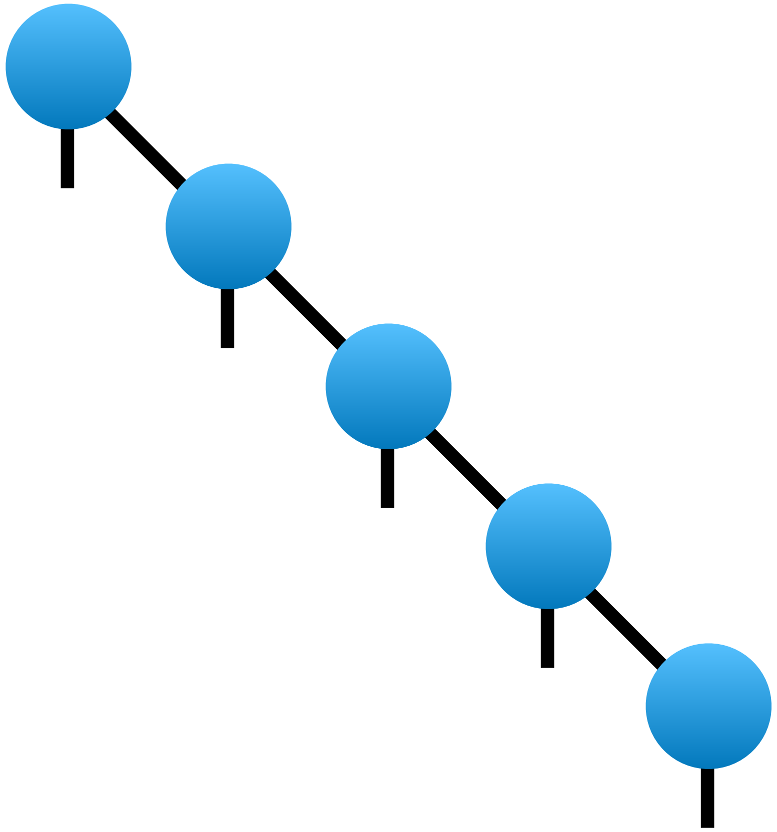
Evenbly, Vidal, PRB **79**, 144108 (2009)

Verstraete, Cirac, cond-mat/0407066 (2004)

Orus, Ann. Phys. **349**, 117 (2014)

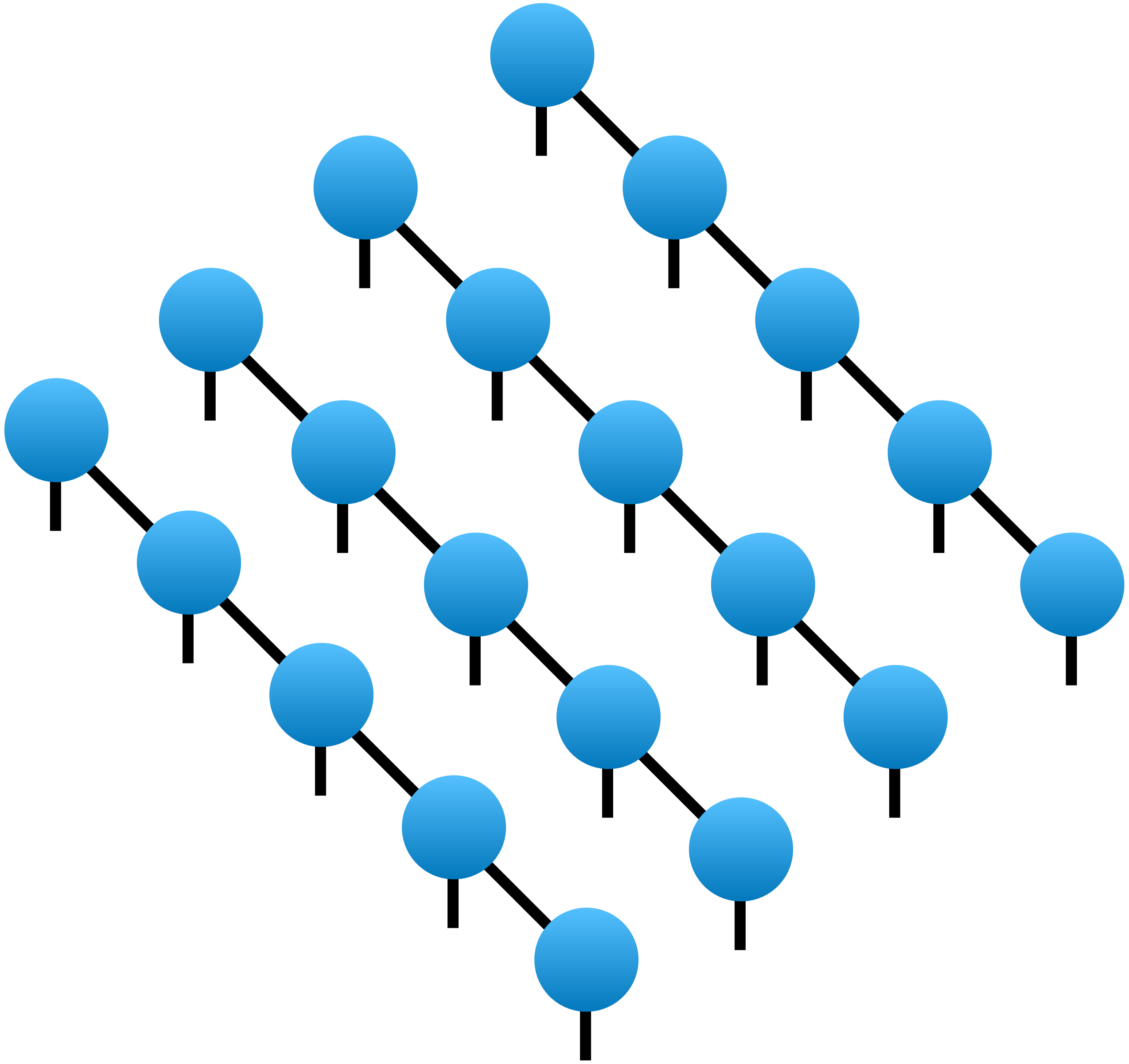
PEPS Tensor Network

Most straightforward extension of matrix product states to two-dimensional lattices



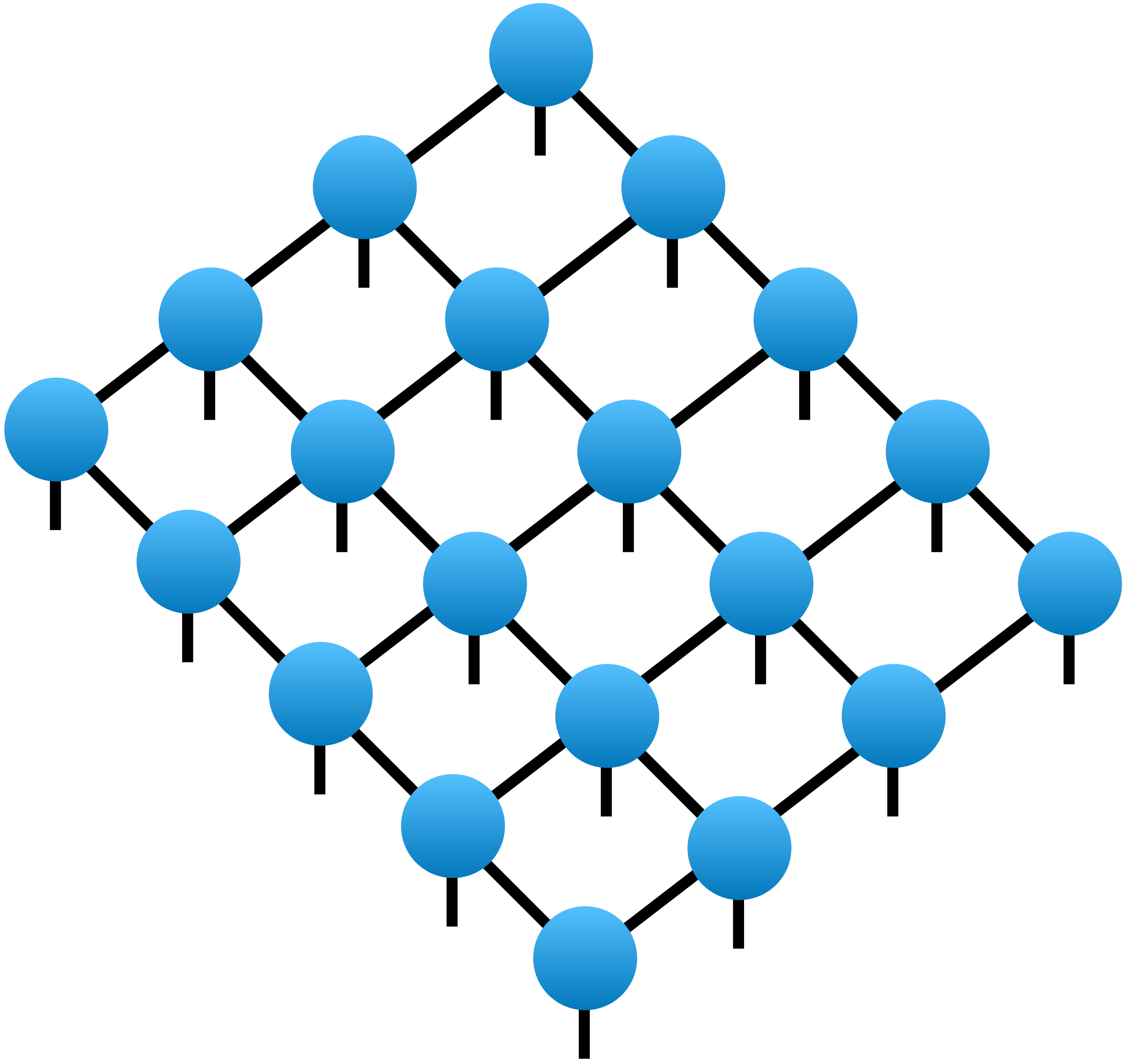
PEPS Tensor Network

Most straightforward extension of matrix product states to two-dimensional lattices



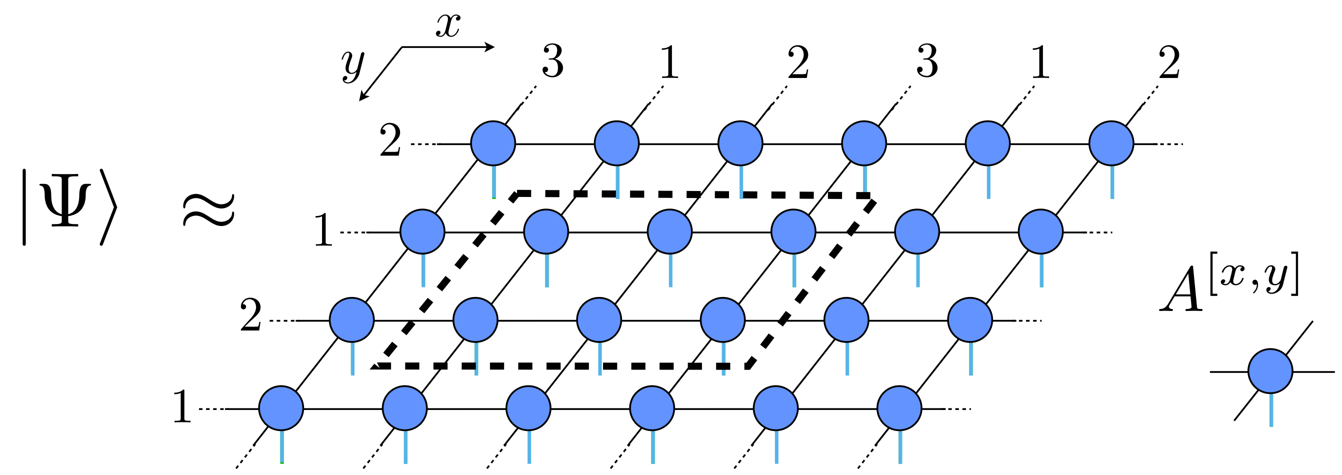
PEPS Tensor Network

Most straightforward extension of matrix product states to two-dimensional lattices



PEPS Tensor Network

Powerful algorithms to address *infinite* 2D systems



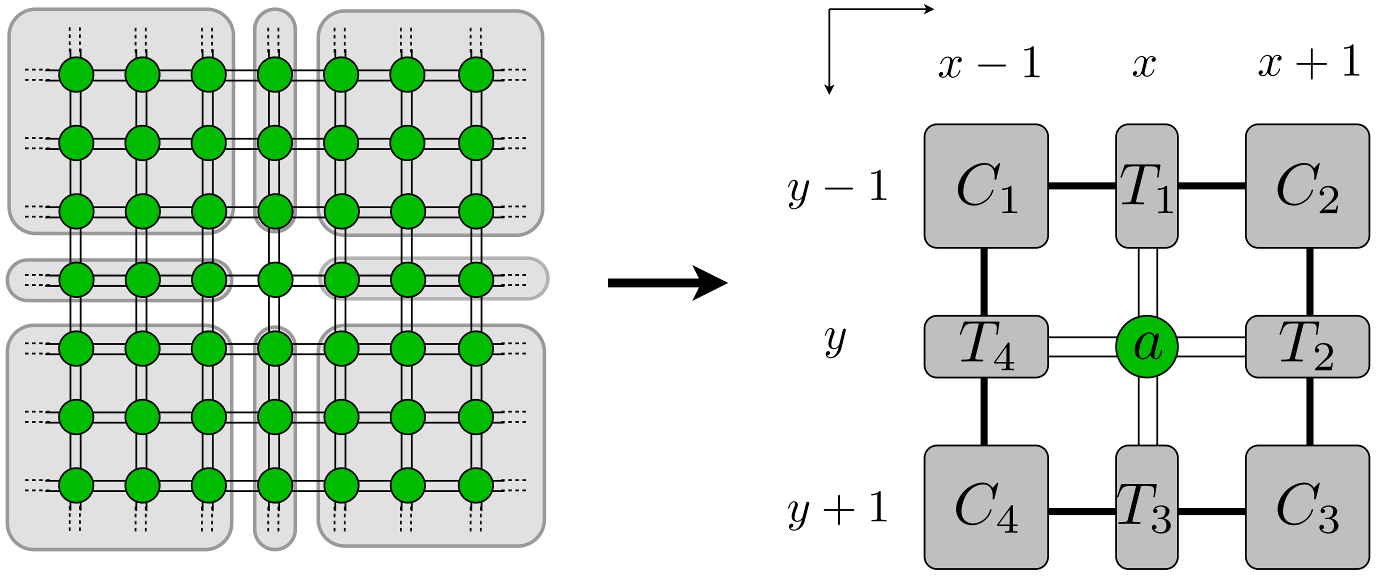
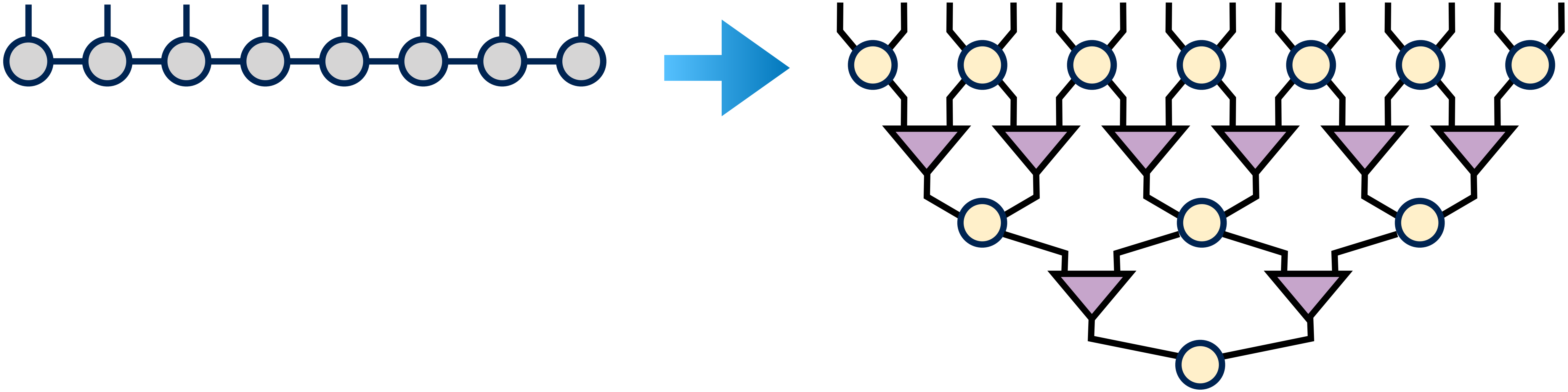


Figure from: Corboz, PRB 94, 035133 (2016)

MERA Tensor Network

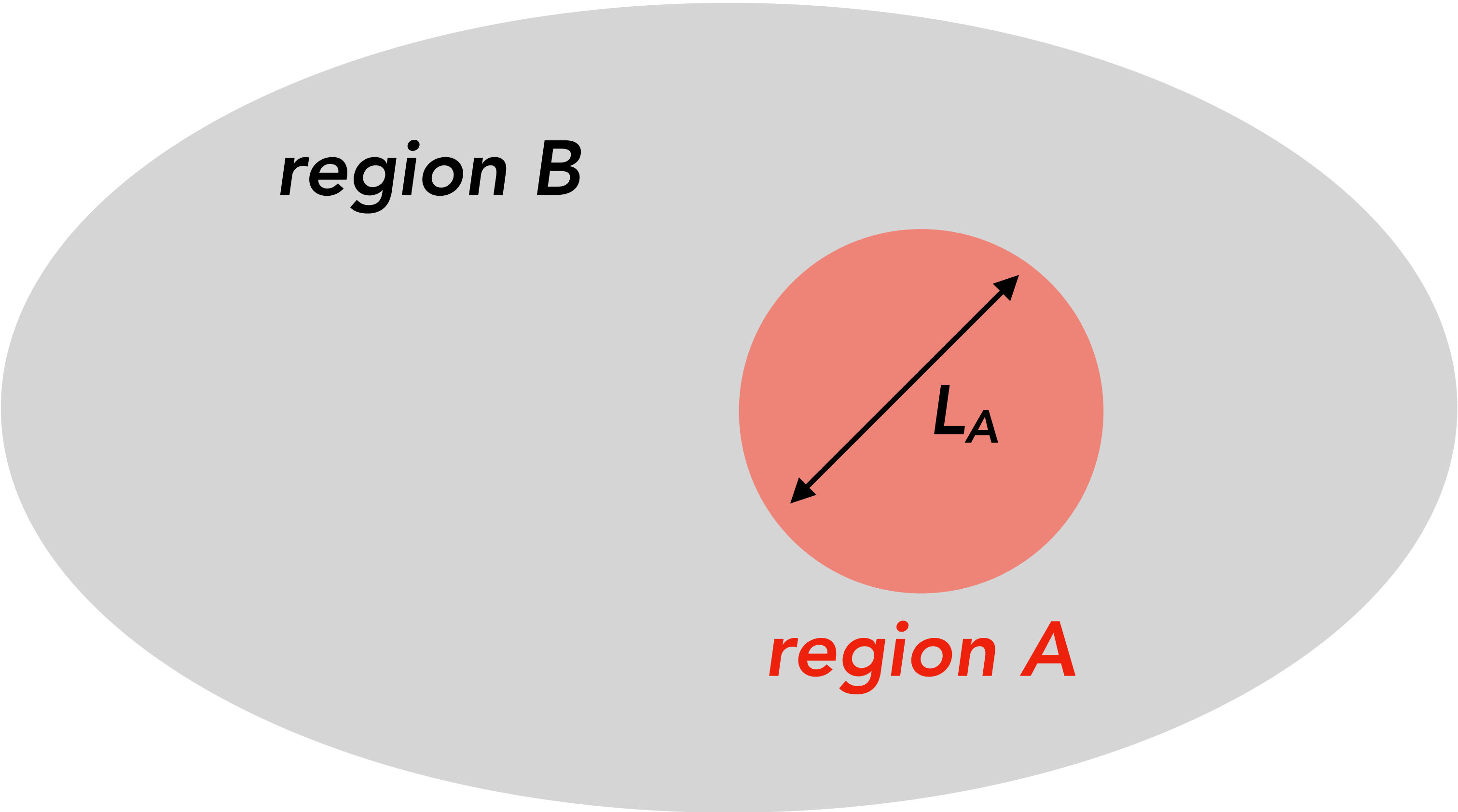
The MERA tensor network generalizes matrix product state to a layered structure



Entanglement Scaling

Key motivation of PEPS and MERA is capturing scaling of entanglement beyond 1D, gapped case

Two basic types of behavior (dimensions d=1,2):

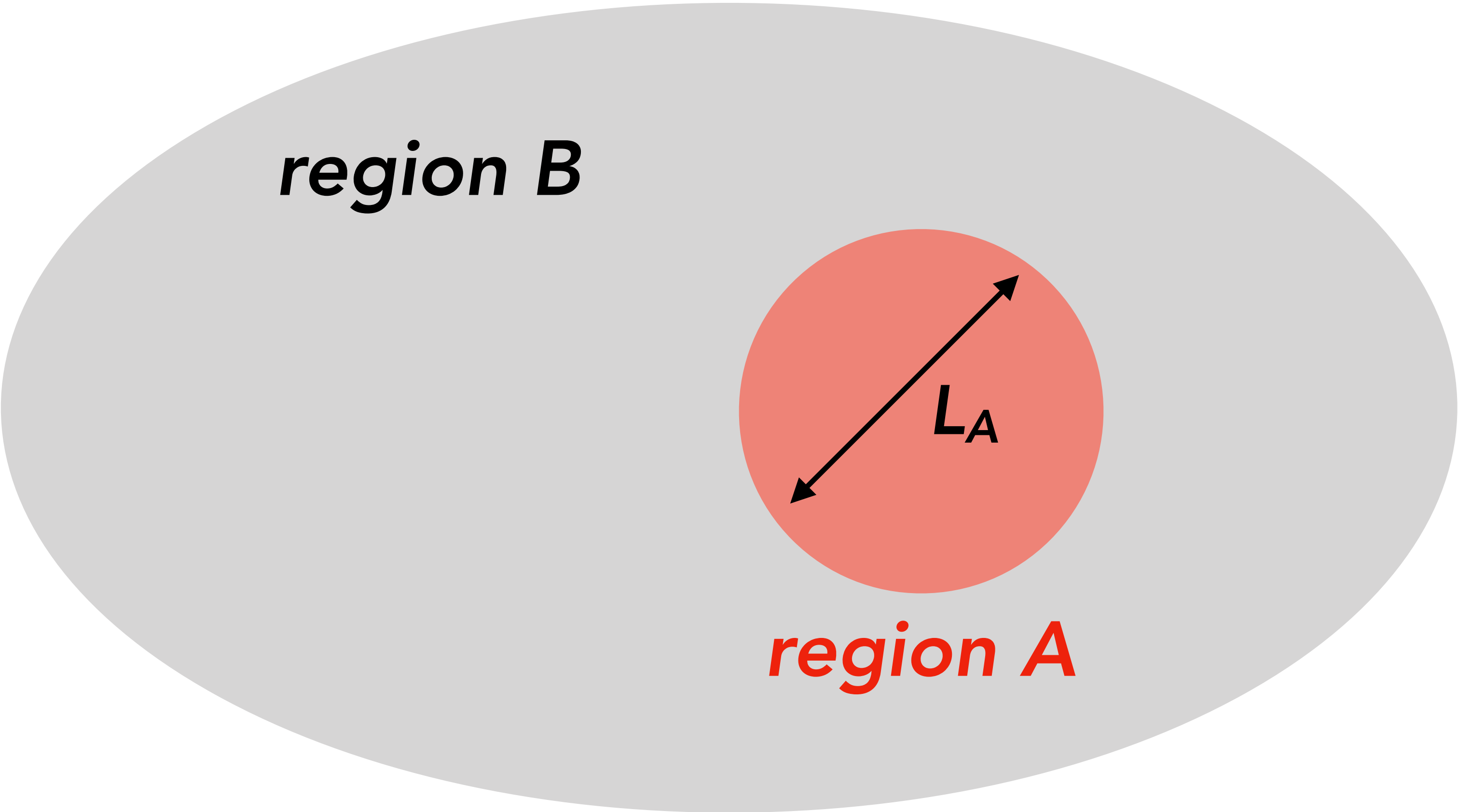


Entanglement Scaling

Key motivation of PEPS and MERA is capturing scaling of entanglement beyond 1D, gapped case

Two basic types of behavior (dimensions d=1,2):

Boundary law: $S \sim L_A^{d-1}$



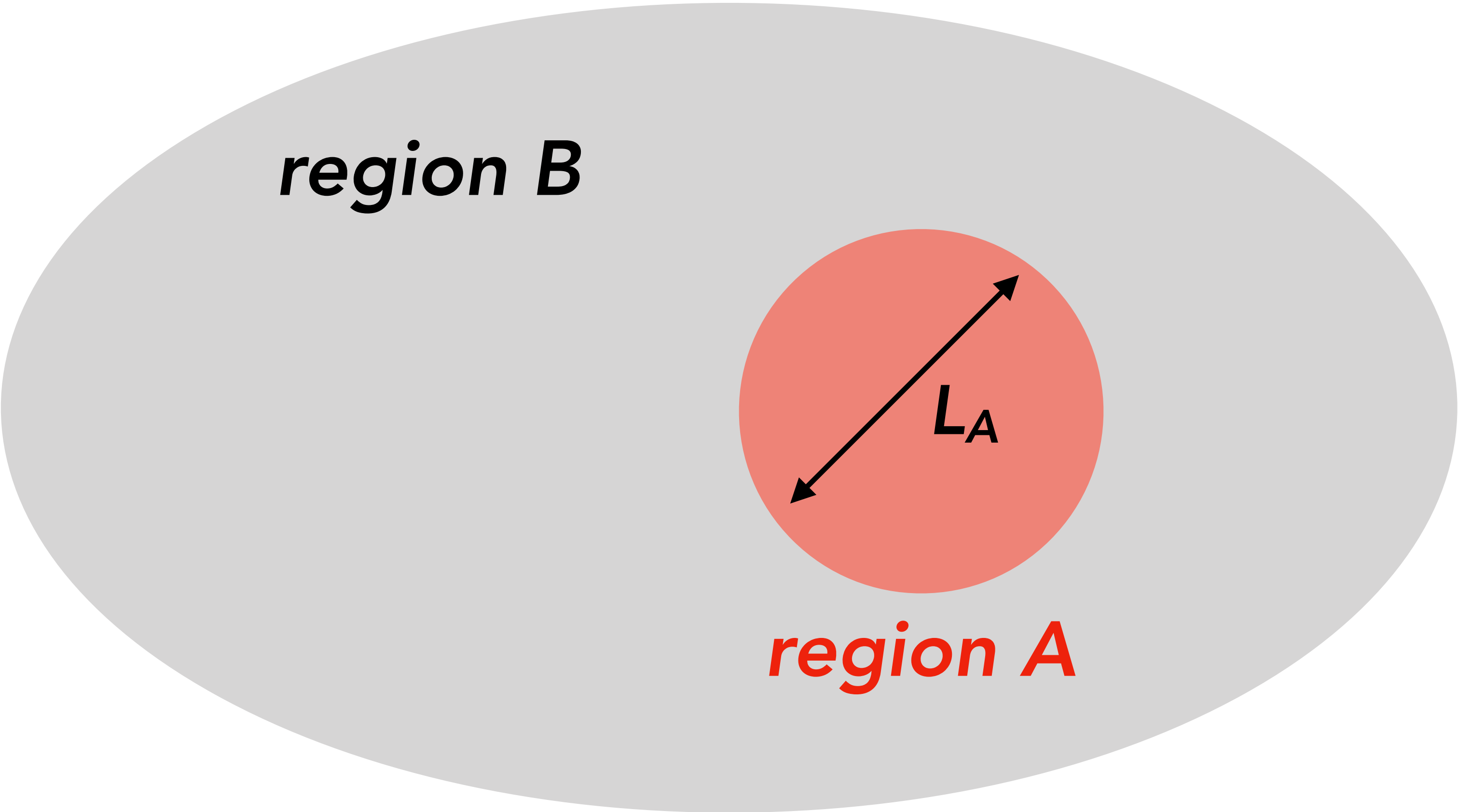
Entanglement Scaling

Key motivation of PEPS and MERA is capturing scaling of entanglement beyond 1D, gapped case

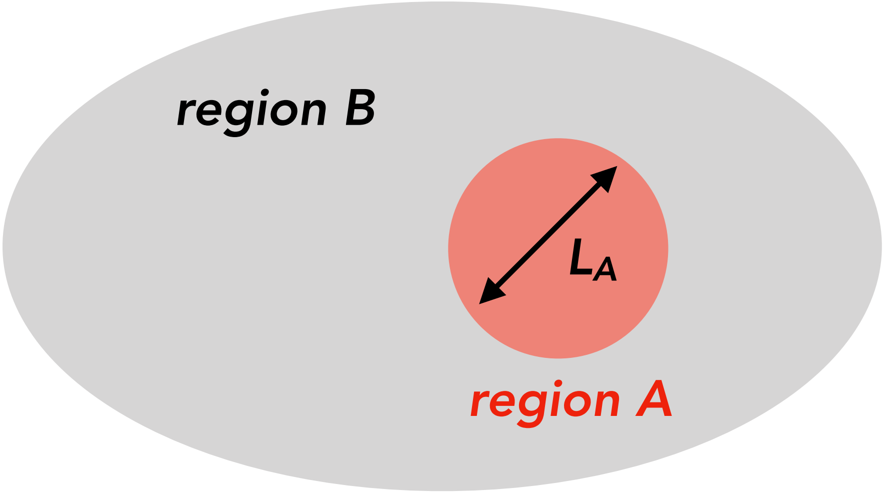
Two basic types of behavior (dimensions d=1,2):

Boundary law: $S \sim L_A^{d-1}$

Boundary law times log correction: $S \sim \log(L_A)\ L_A^{d-1}$



Entanglement Scaling



1. *Boundary law:* $S \sim L_A^{d-1}$
2. *Boundary law times log correction:* $S \sim \log(L_A)\ L_A^{d-1}$

Intuition behind these two behaviors:

1. Product state + local fluctuations: all entanglement near boundary between A and B

2. Scale invariant. Boundary law contribution at each scale, log(L_A) rescaling transformations until region A shrinks to a point

Entanglement Scaling

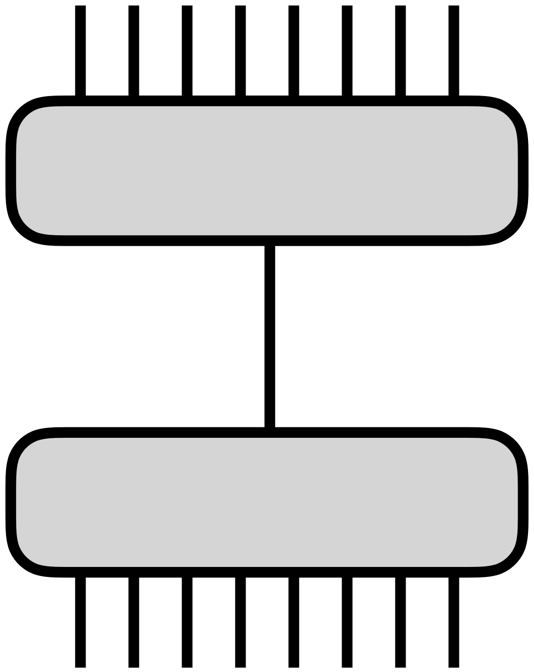
Based on numerical & field theory evidence, can make following table of expected scaling:

Dimension	Gap ?	Scaling S with L_A
1	gapped	const.
1	critical	$\log(L_A)$
2	gapped	L_A
2	gapless, Fermi points	L_A
2	gapless, Fermi surface	$L_A \log(L_A)$

Min-cut Argument

To motivate why PEPS and MERA are a good idea make following observation:

Consider a matrix (= density matrix) having internal 'tensor network' structure

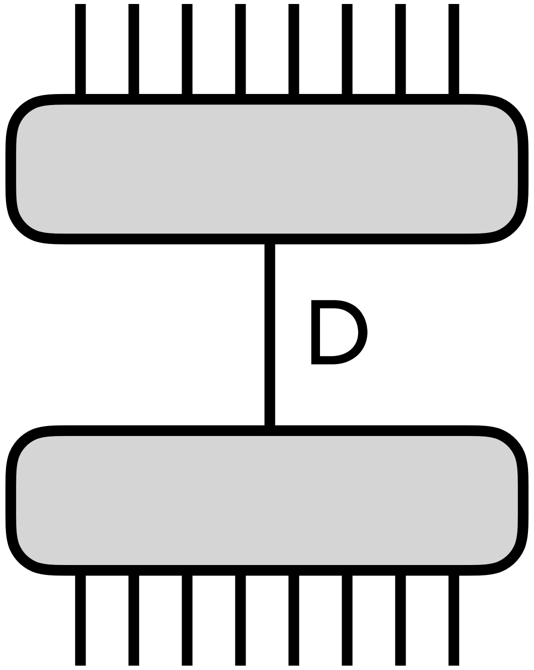
$\rho =$ 

Min-cut Argument

For simplicity, assume all index lines of size D

Then ρ is D^8 x D^8 matrix

But can be written as sum of outer products of only D vectors (each of dimension D^8)

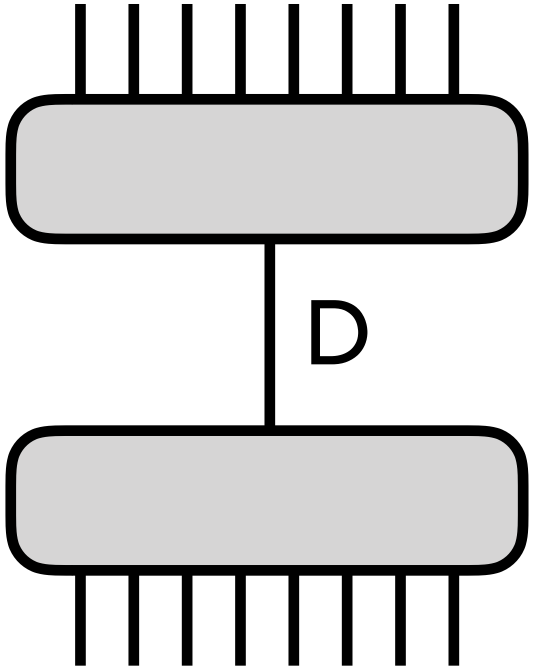


Min-cut Argument

Call linear dimension of ρ "R" (= D^8 in example)

Then ρ has a kernel (null space) of dimension R-D

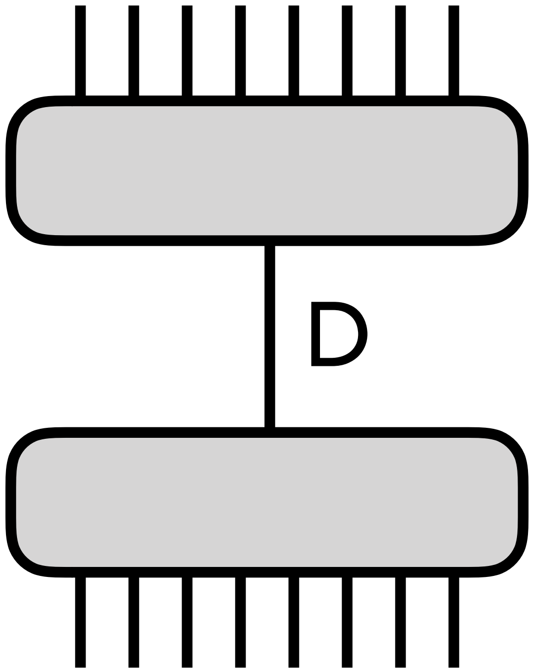
At most D of the vectors forming ρ can be linearly independent



Min-cut Argument

So rank of ρ is R-(dimension of kernel) = R-(R-D) = D

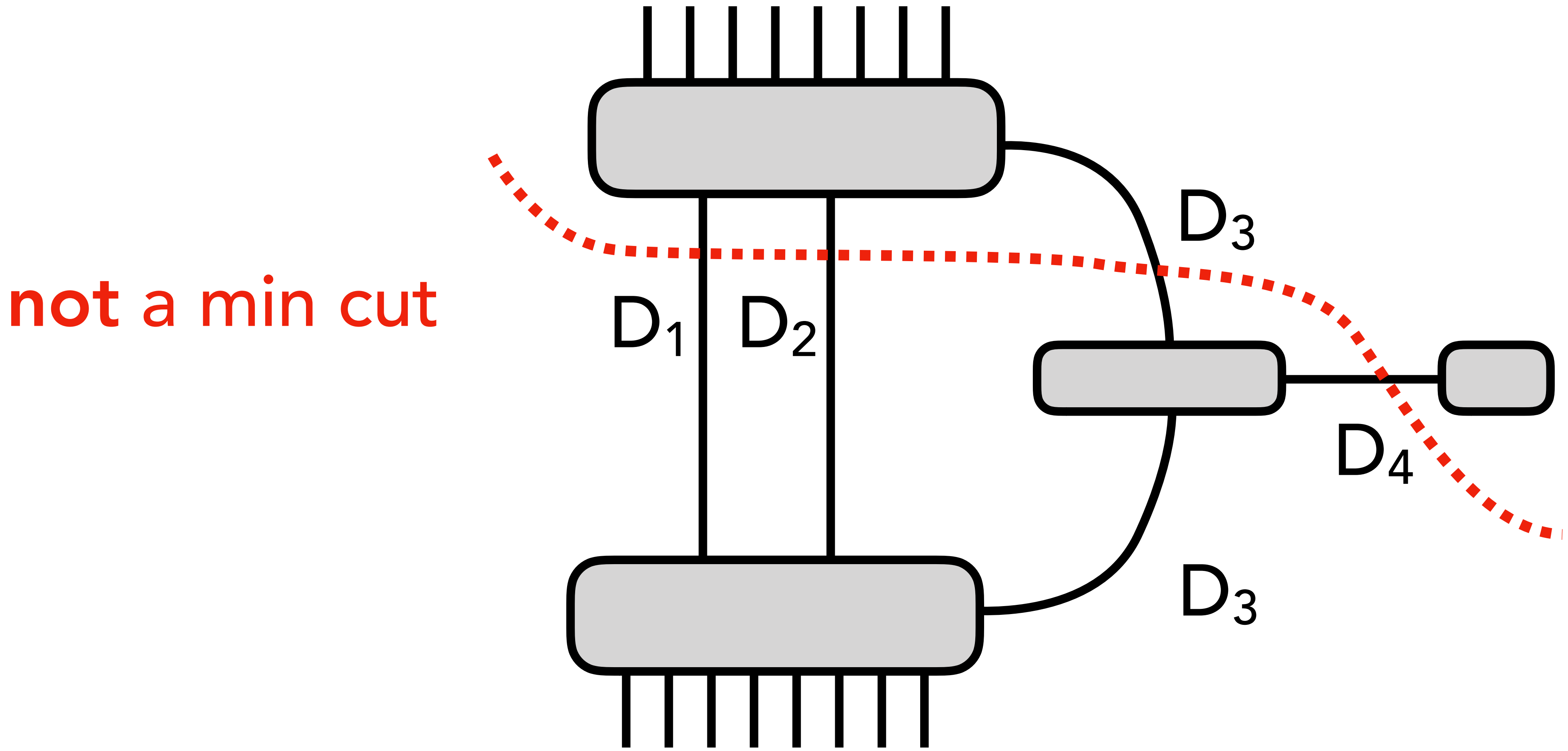
Can only have D non-zero eigenvalues



© Edwin Miles Stoudenmire, Flatiron Institute, 2018

Min-cut Argument

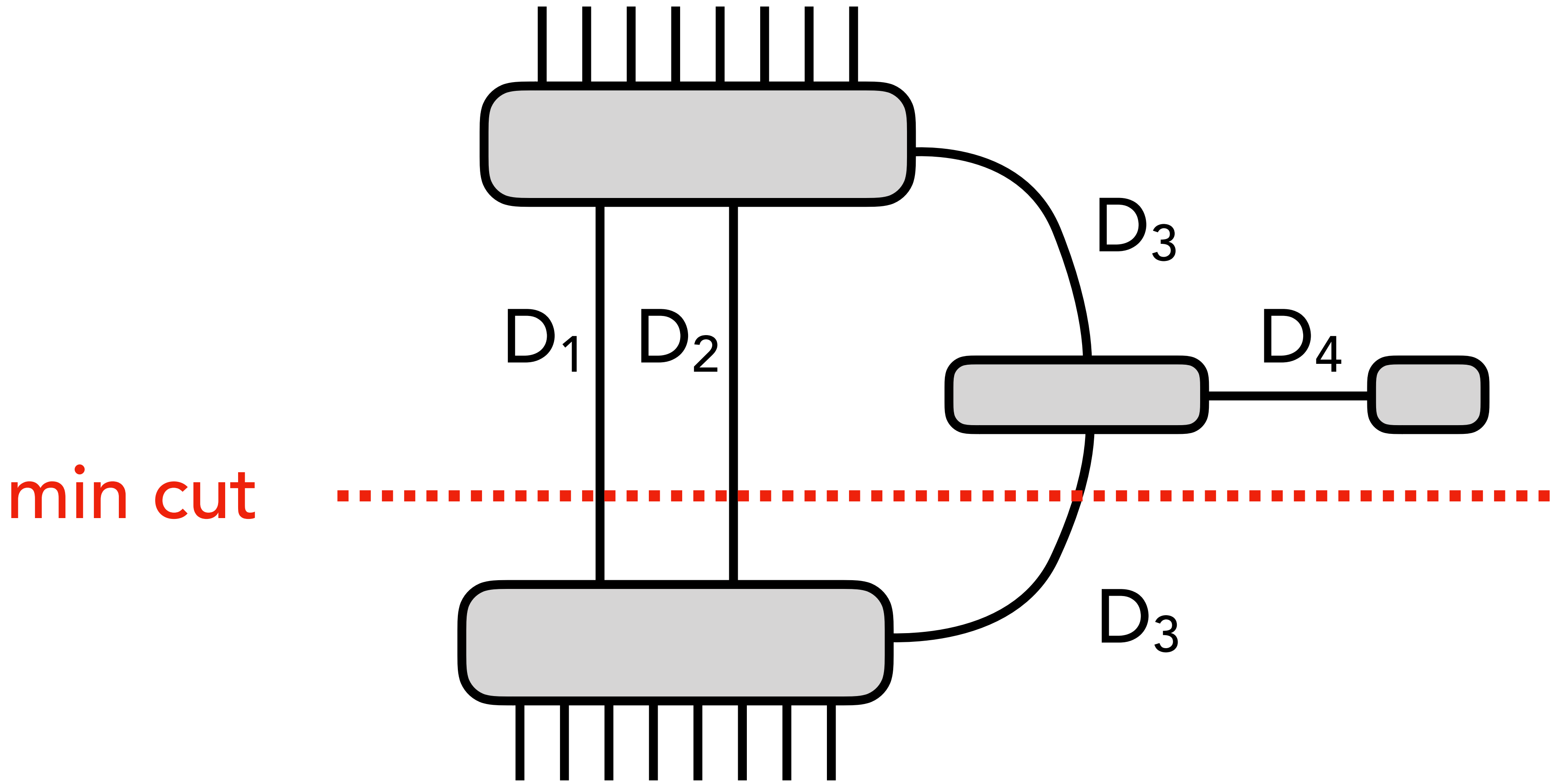
More generally, can show rank of a "structured" matrix is minimum of product of dimensions of lines needed to cut into two pieces



© Edwin Miles Stoudenmire, Flatiron Institute, 2018

Min-cut Argument

More generally, can show rank of a "structured" matrix is minimum of product of dimensions of lines needed to cut into two pieces

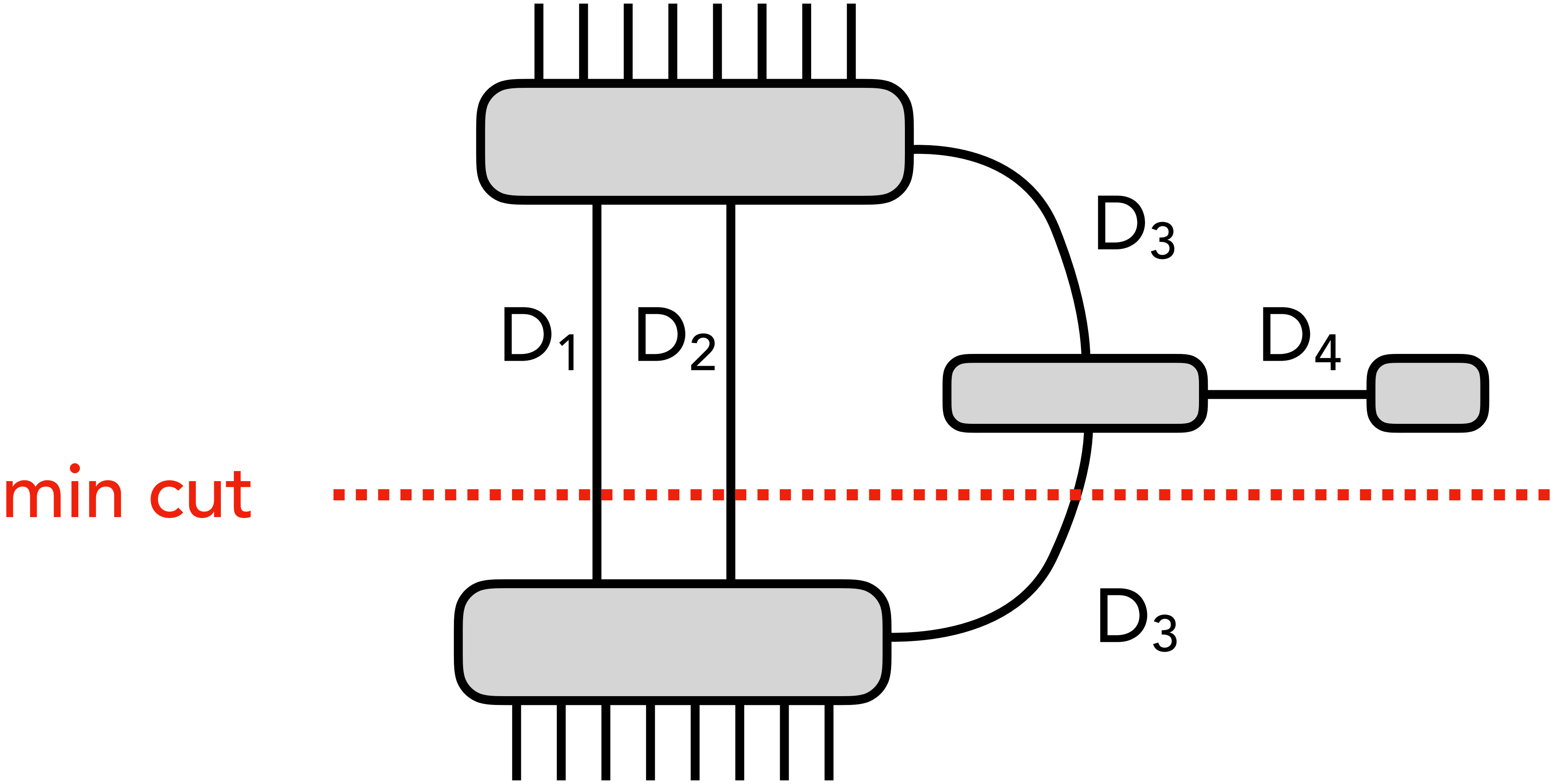


$$\boxed{\text{rank}(\rho) = D_1\, D_2\, D_3}$$

Min-cut Argument

If matrix is a density matrix, then maximum entropy occurs when all eigenvalues = 1/rank

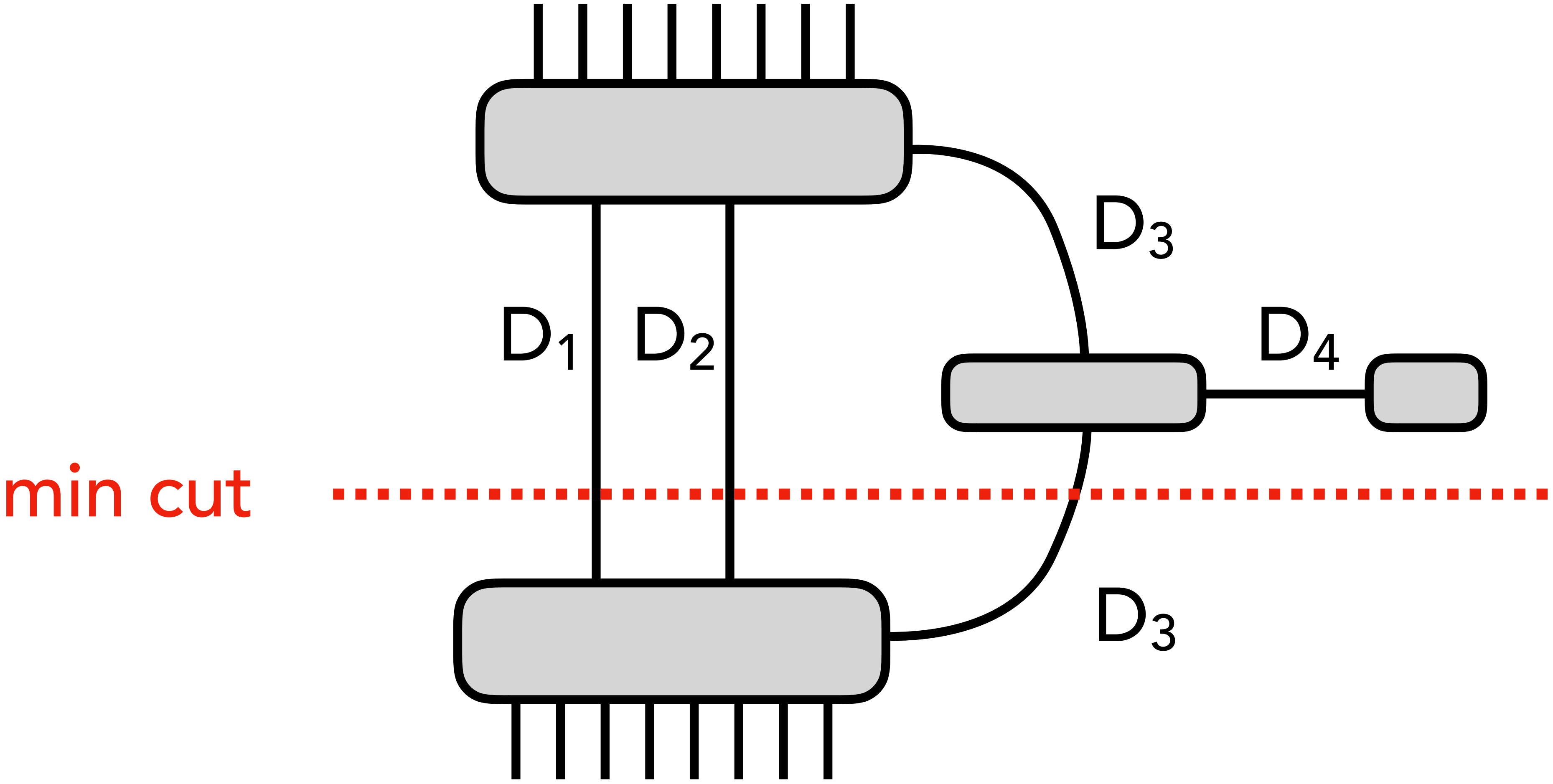
$$\Longrightarrow \quad S_{\max} = \log(\text{rank})$$



$$\boxed{S_{\max}(\rho) = \log(D_1 D_2 D_3)}$$

Min-cut Argument

Let's apply this argument to various tensor networks

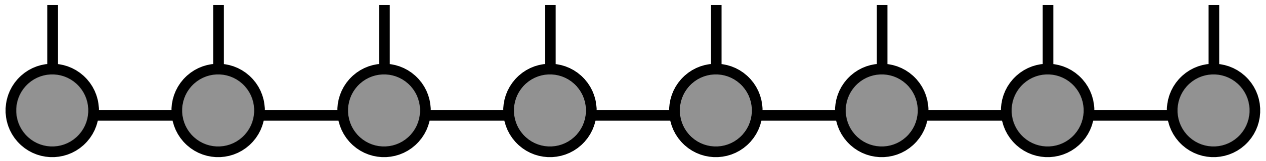


$$\boxed{S_{\text{max}}(\rho) = \log(D_1 D_2 D_3)}$$

Min-cut Argument

Matrix product state (MPS)

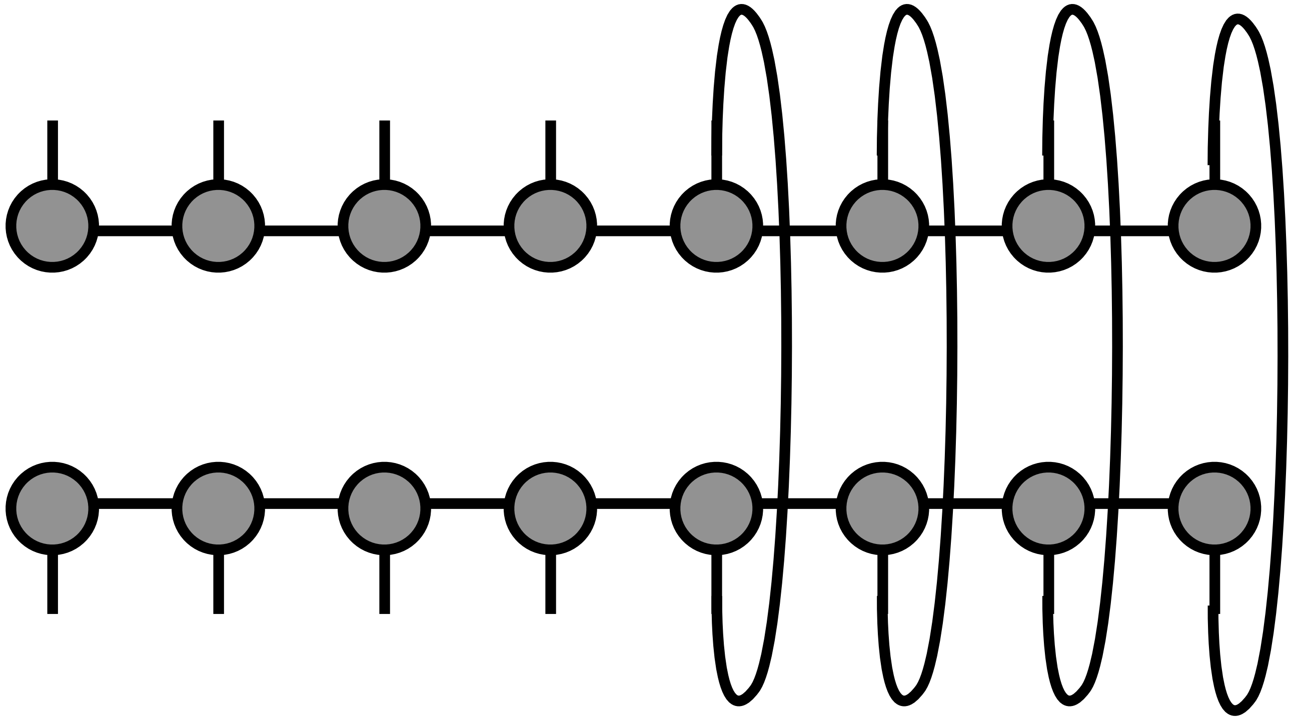
$|\Psi\rangle$



Min-cut Argument

Matrix product state (MPS)

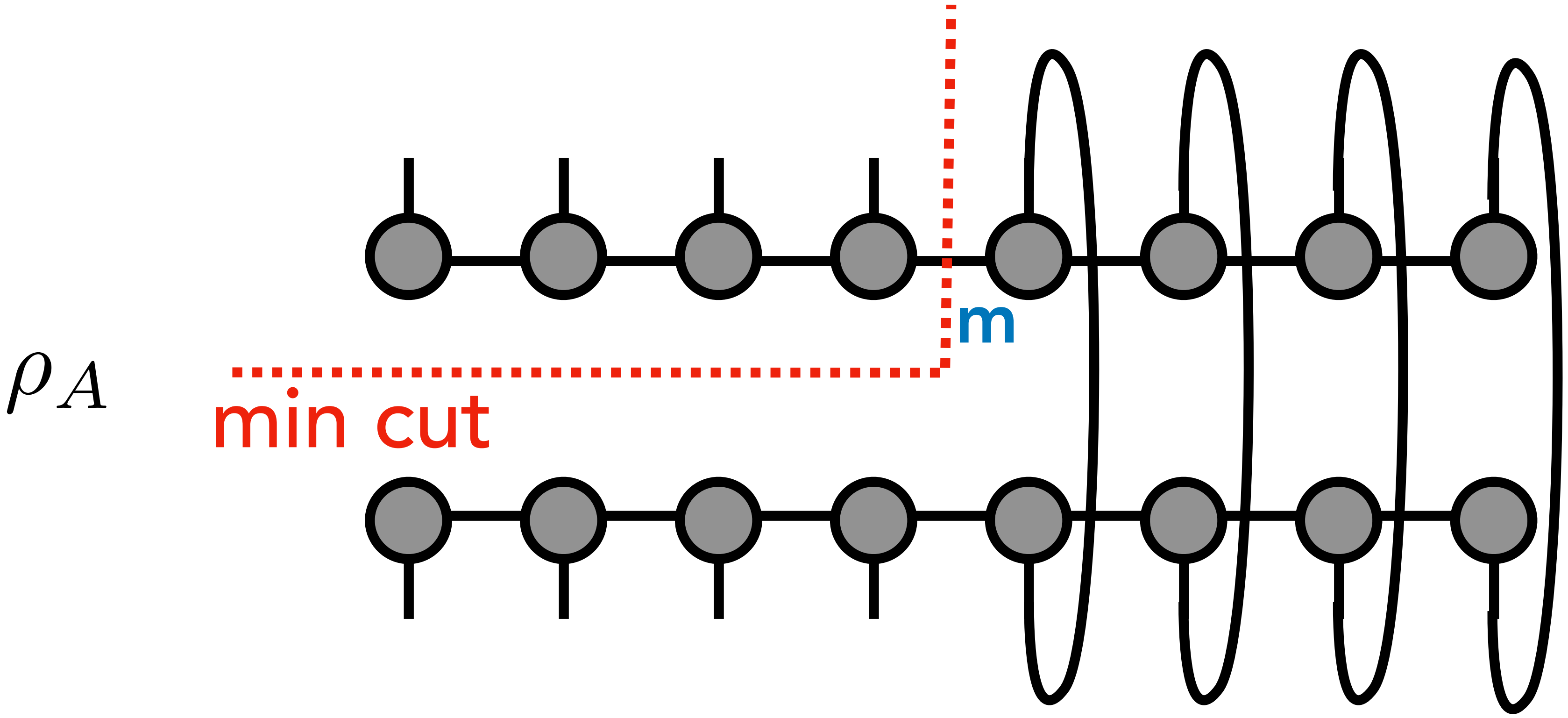
ρ_A



Independent of system size = *boundary law*

Min-cut Argument

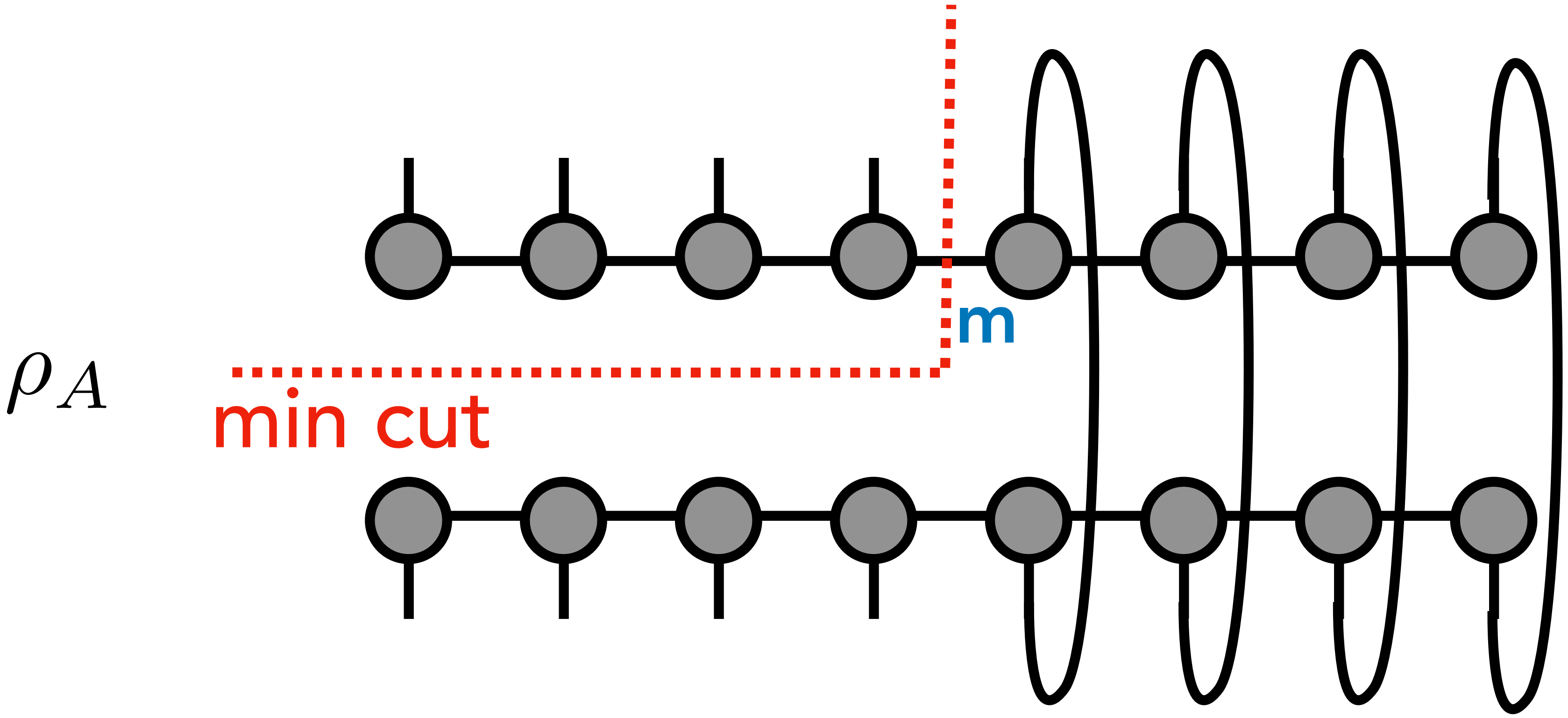
Matrix product state (MPS)



Independent of system size = *boundary law*

Min-cut Argument

Matrix product state (MPS)

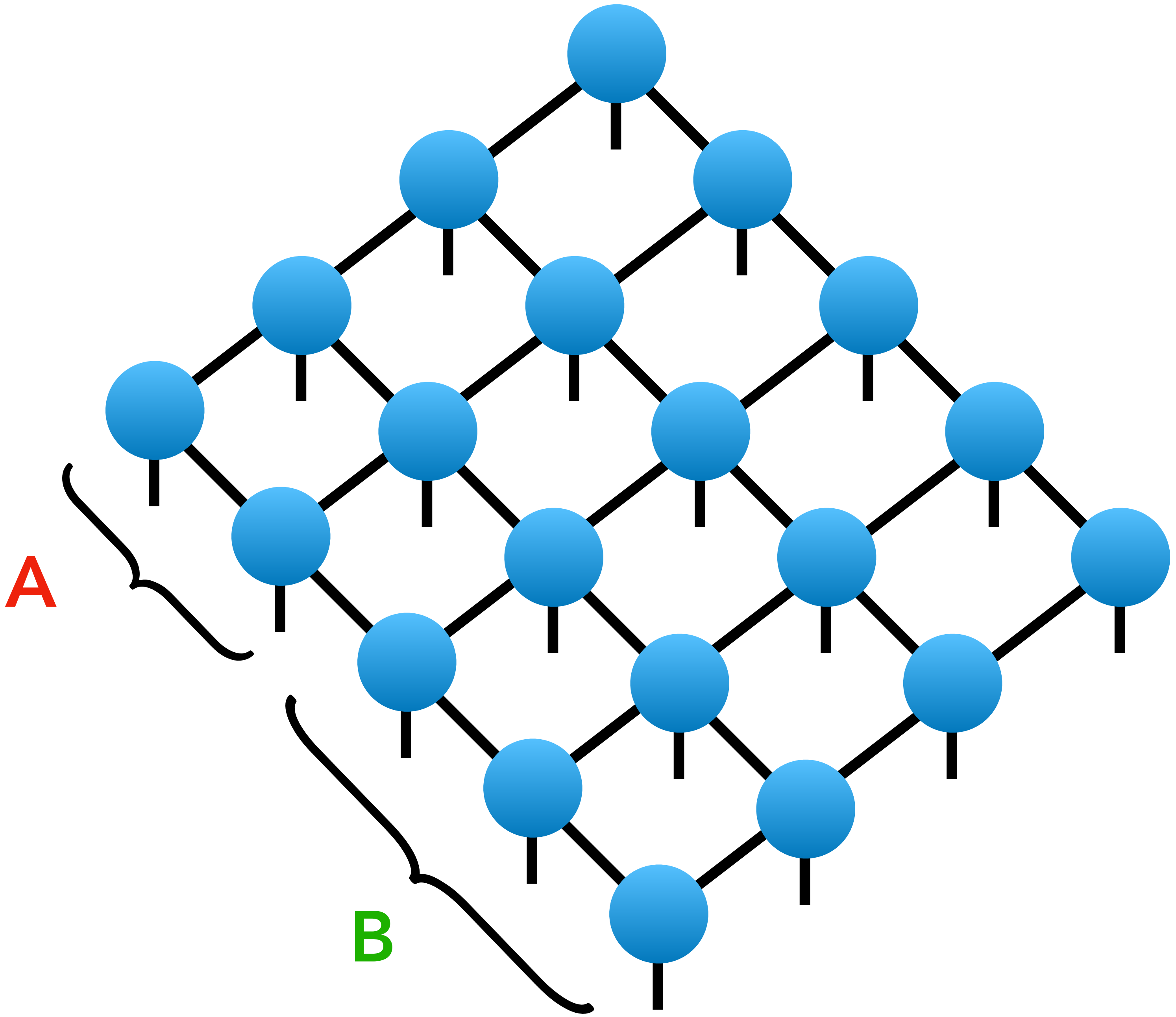


$S_{\max} = \log(\, \mathbf{m} \,)$

Independent of system size = *boundary law*

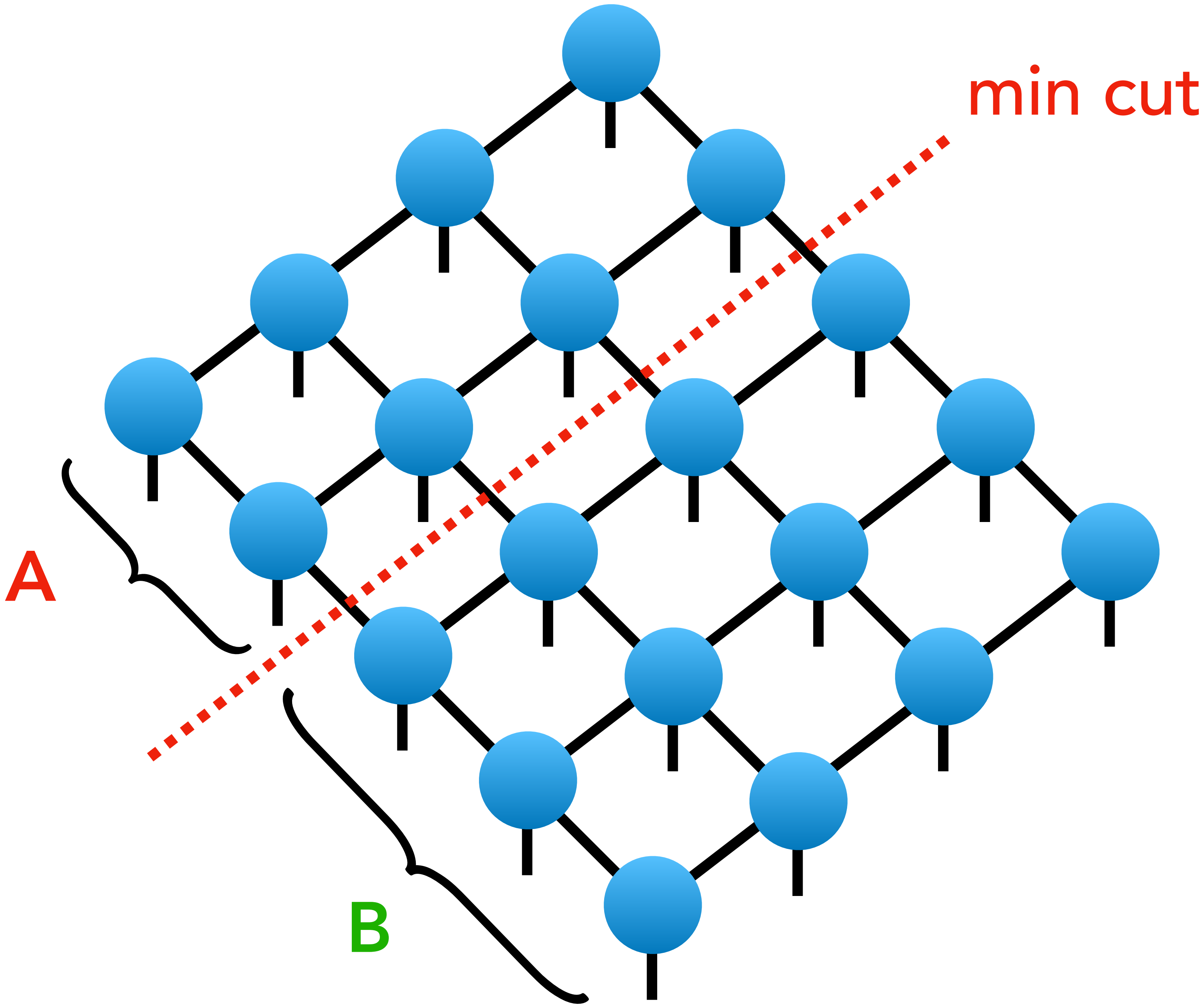
Min-cut Argument

PEPS wavefunction



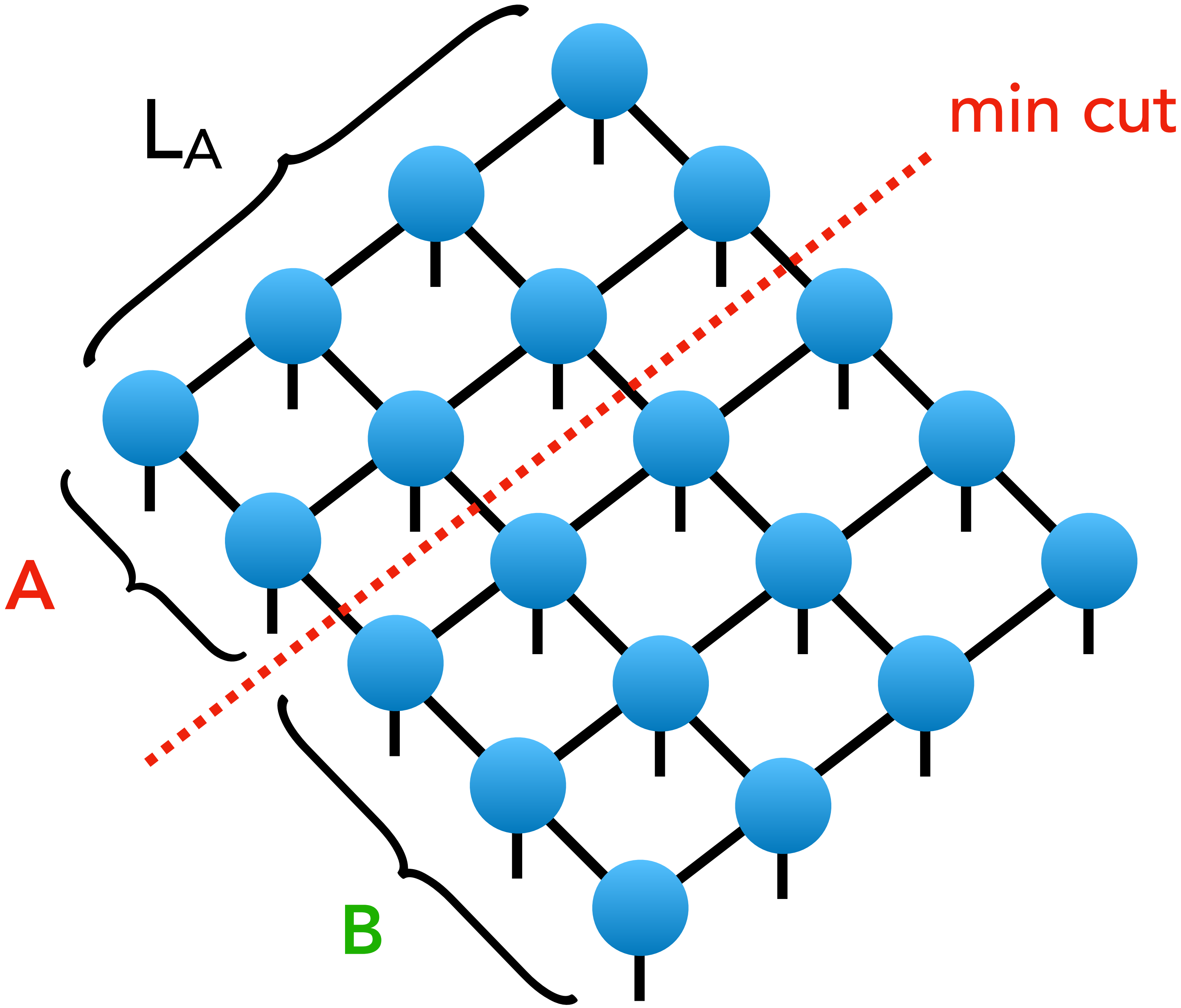
Min-cut Argument

PEPS wavefunction



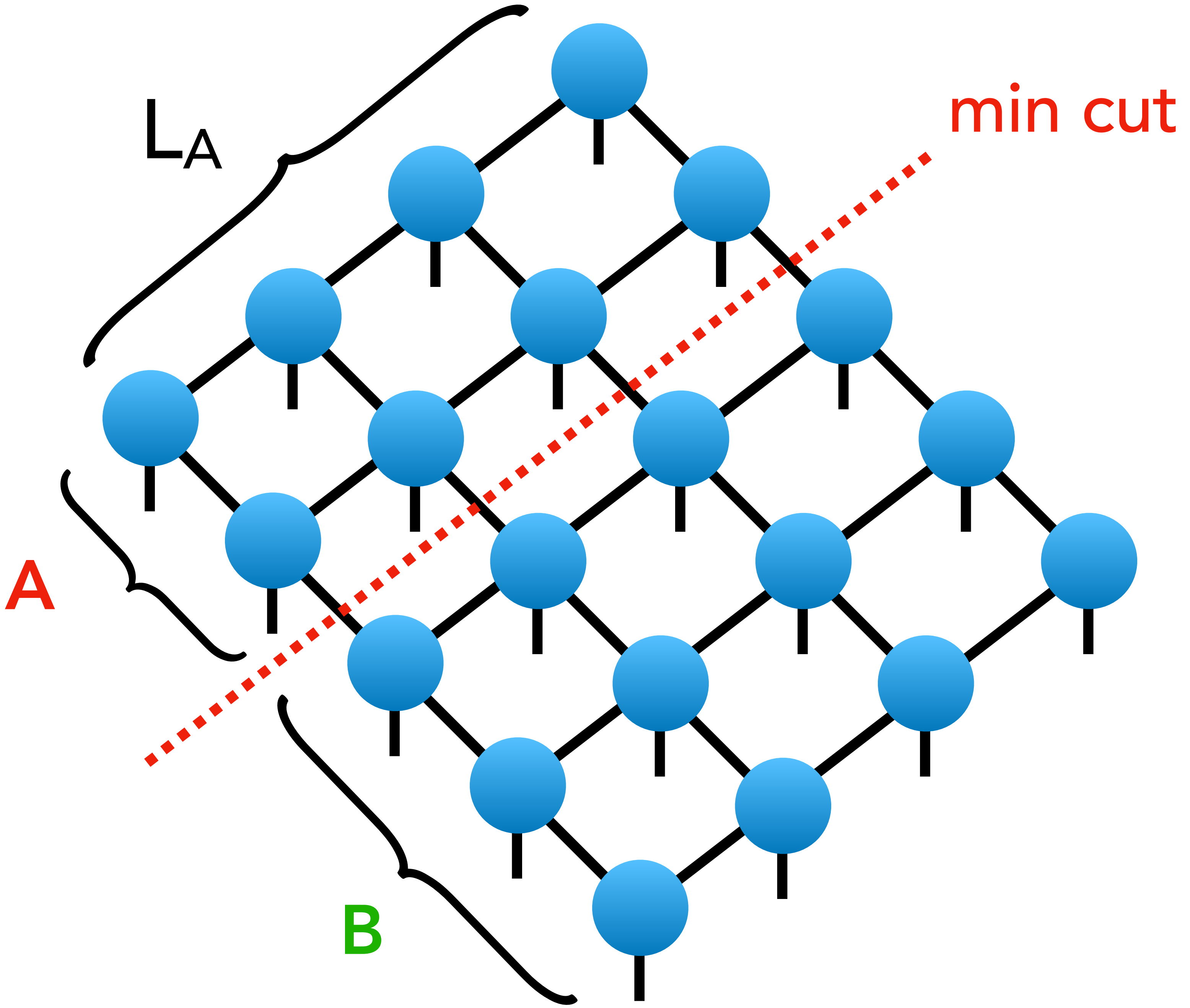
Min-cut Argument

PEPS wavefunction



Min-cut Argument

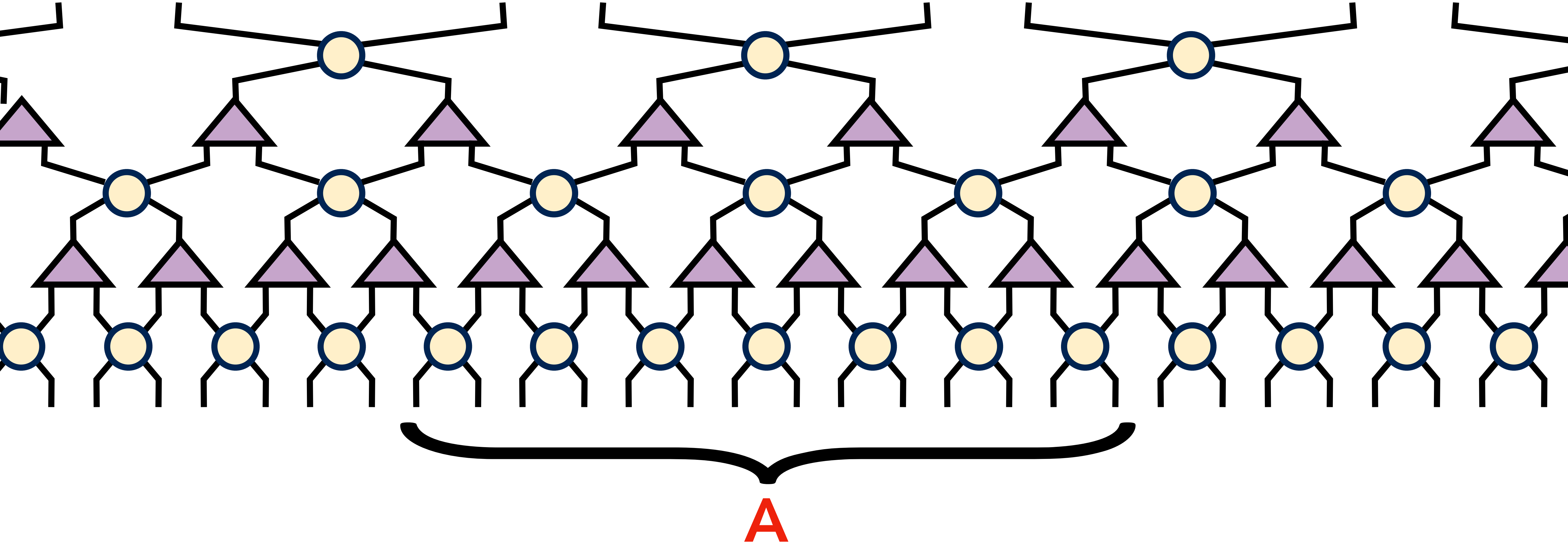
PEPS wavefunction



$$S_{\max} = \log(D^{L_A}) = L_A \log(D)$$

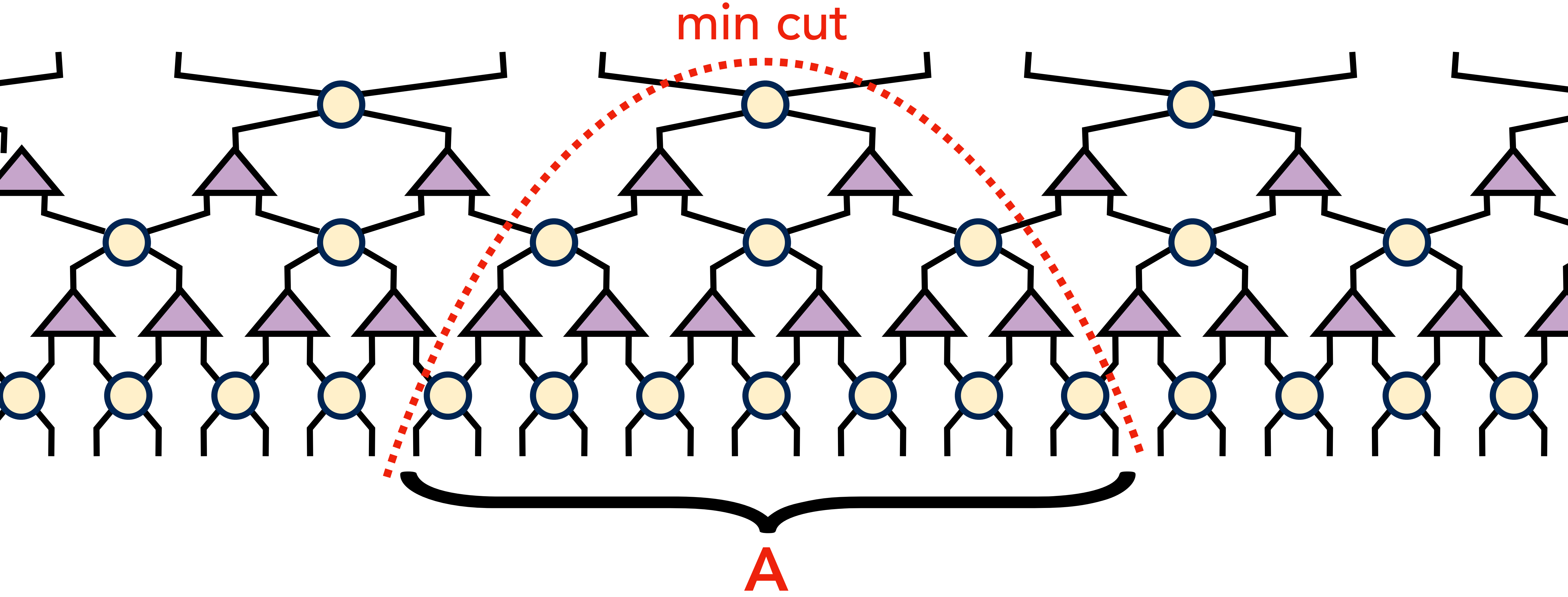
Min-cut Argument

MERA wavefunction



Min-cut Argument

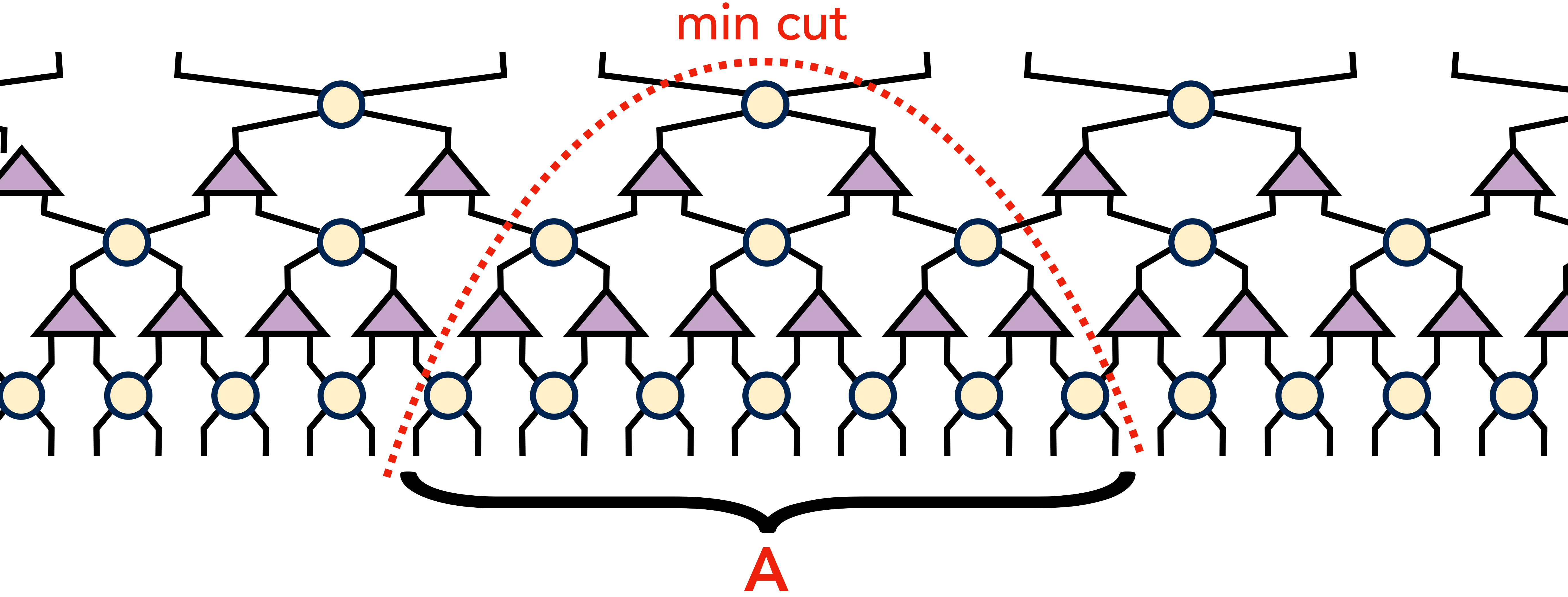
MERA wavefunction



© Edwin Miles Stoudenmire, Flatiron Institute, 2018

Min-cut Argument

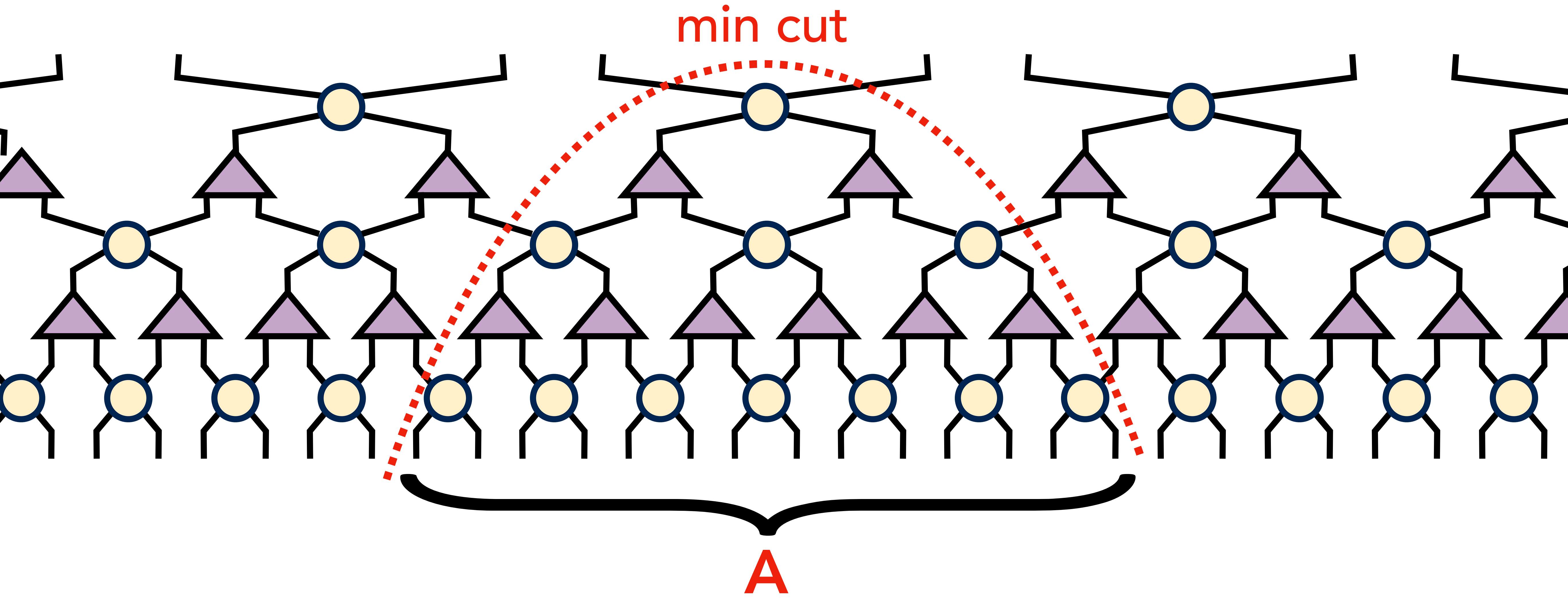
MERA wavefunction



Number of layers intersecting *min cut* of region A is log(L_A)

Min-cut Argument

MERA wavefunction



Number of layers intersecting *min cut* of region A is $\log(L_A)$

$$S_{\mathrm{max}} = \log(\, D^{\log(L_A)}\,) = \log(L_A)\log(\, D\,)$$

Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A
1	gapped	const.
1	critical	$\log(L_A)$
2	gapped	L_A
2	gapless, Fermi points	L_A
2	gapless, Fermi surface	$L_A \log(L_A)$

Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A
1	gapped	const.
1	critical	$\log(L_A)$
2	gapped	L_A
2	gapless, Fermi points	L_A
2	gapless, Fermi surface	$L_A \log(L_A)$

MPS

Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A	
1	gapped	const.	MPS
1	critical	$\log(L_A)$	MERA
2	gapped	L_A	
2	gapless, Fermi points	L_A	
2	gapless, Fermi surface	$L_A \log(L_A)$	

Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A	
1	gapped	const.	MPS
1	critical	$\log(L_A)$	MERA
2	gapped	L_A	PEPS
2	gapless, Fermi points	L_A	
2	gapless, Fermi surface	$L_A \log(L_A)$	

Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A	
1	gapped	$\mathrm{const.}$	MPS
1	critical	$\log(L_A)$	MERA
2	gapped	L_A	PEPS & MERA
2	gapless, Fermi points	L_A	
2	gapless, Fermi surface	$L_A \log(L_A)$	

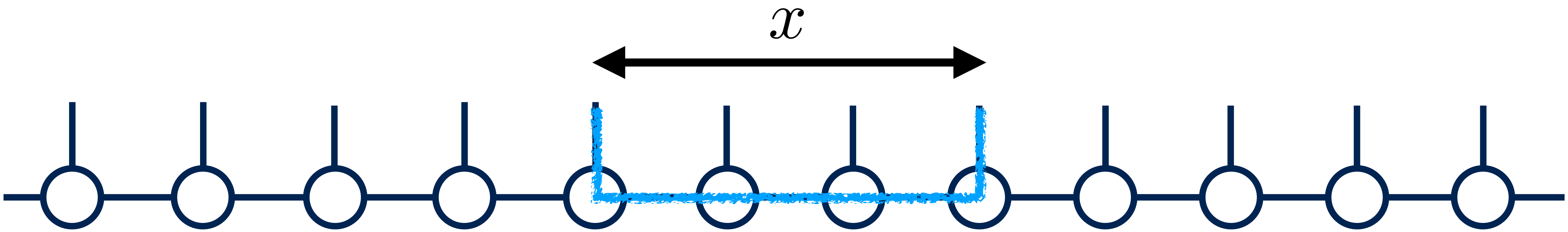
Entanglement Scaling

MPS, MERA, and PEPS realize most cases

Dimension	Gap ?	Scaling S with L_A	
1	gapped	const.	MPS
1	critical	$\log(L_A)$	MERA
2	gapped	L_A	PEPS & MERA
2	gapless, Fermi points	L_A	
2	gapless, Fermi surface	$L_A \log(L_A)$	"branching MERA"

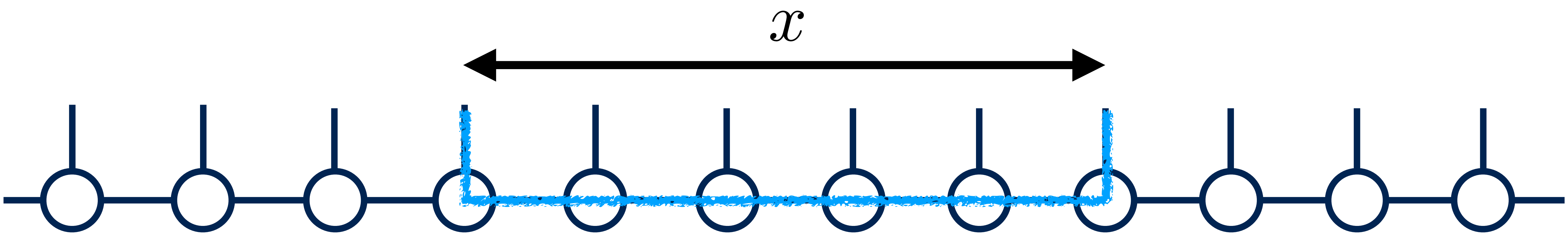
MERA Tensor Network

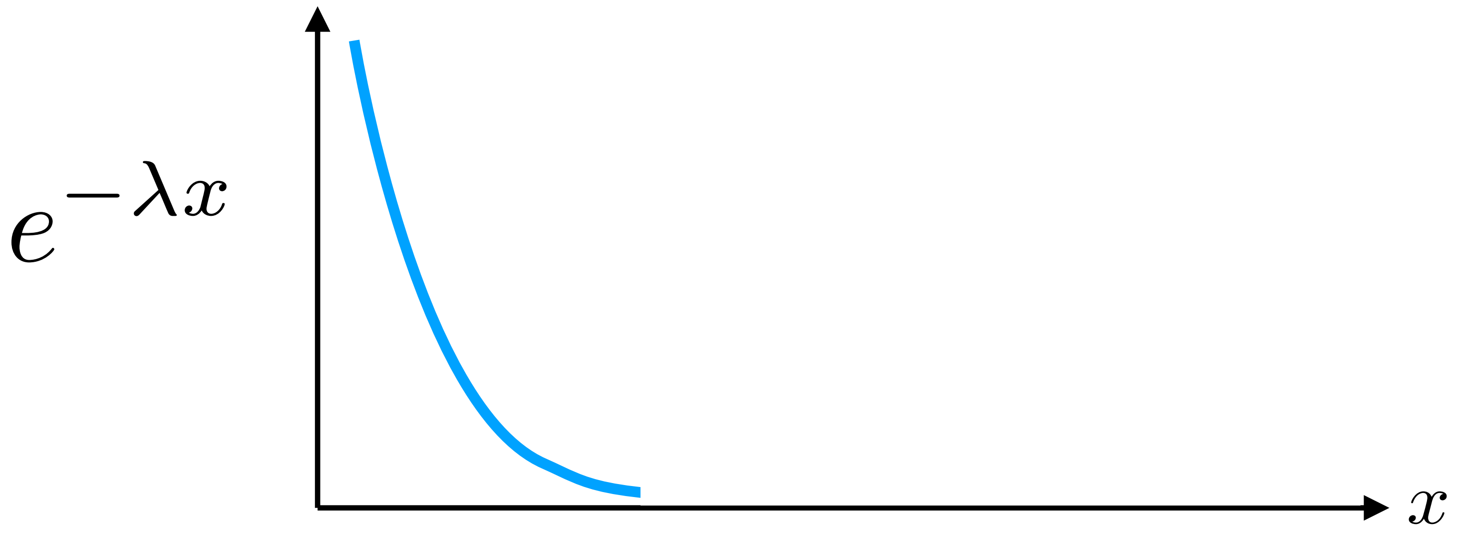
Matrix product state captures only exponential correlations



MERA Tensor Network

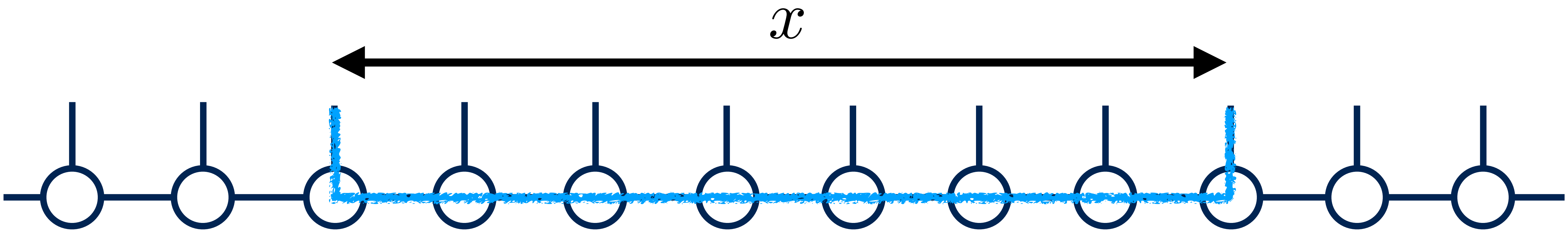
Matrix product state captures only exponential correlations

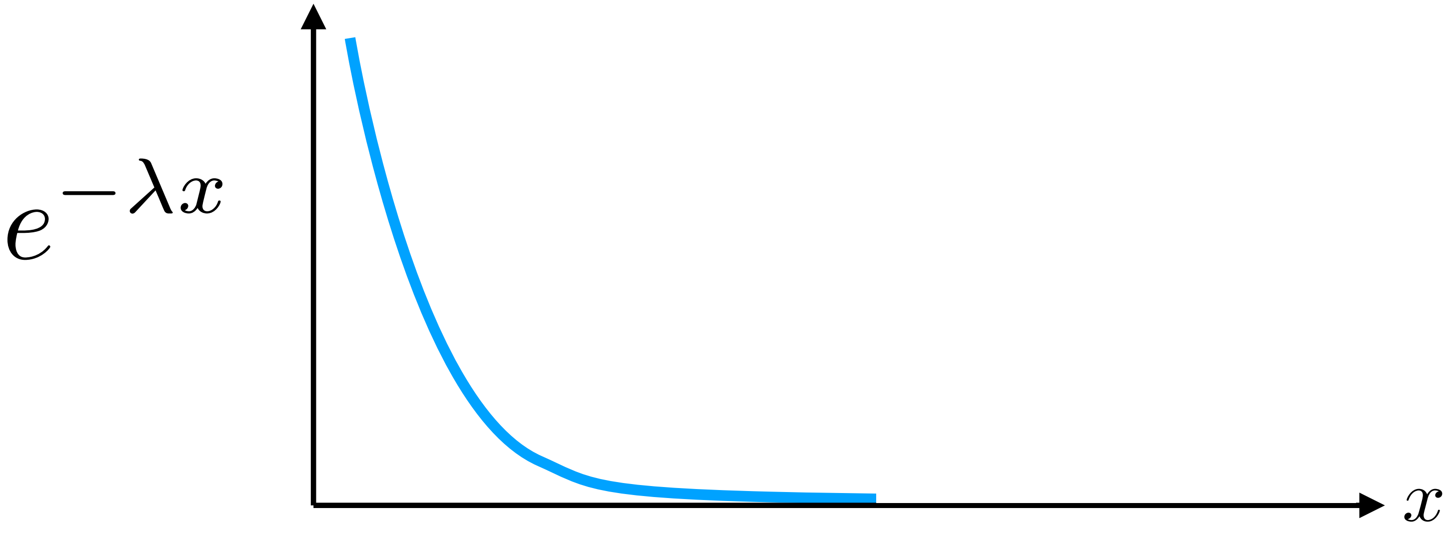




MERA Tensor Network

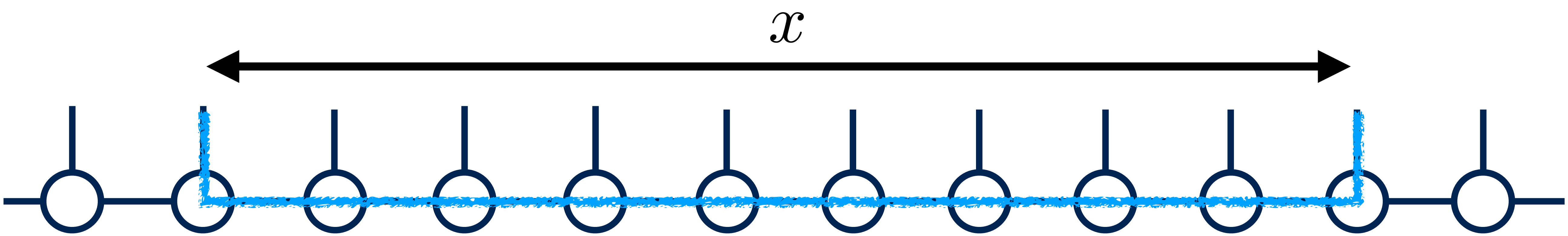
Matrix product state captures only exponential correlations

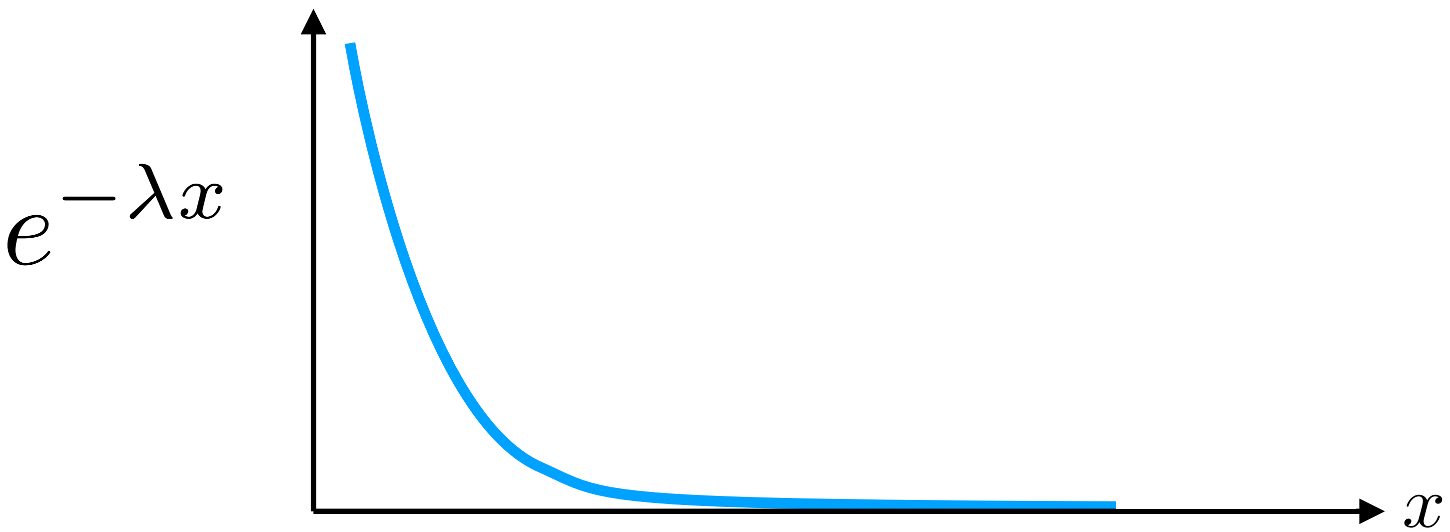




MERA Tensor Network

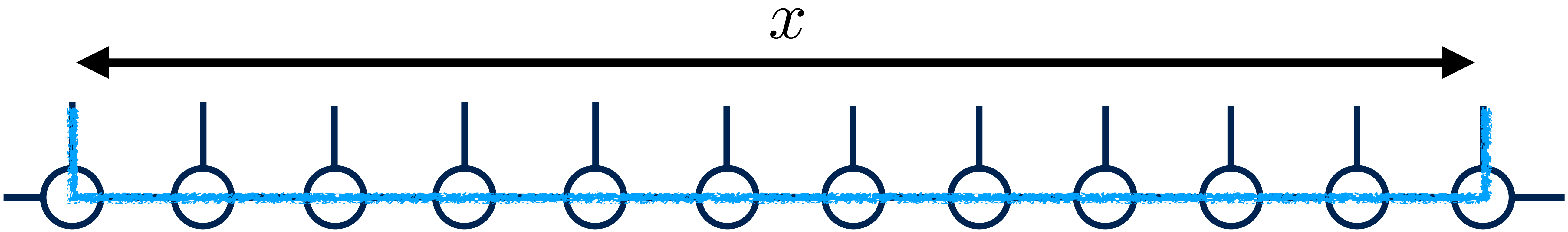
Matrix product state captures only exponential correlations

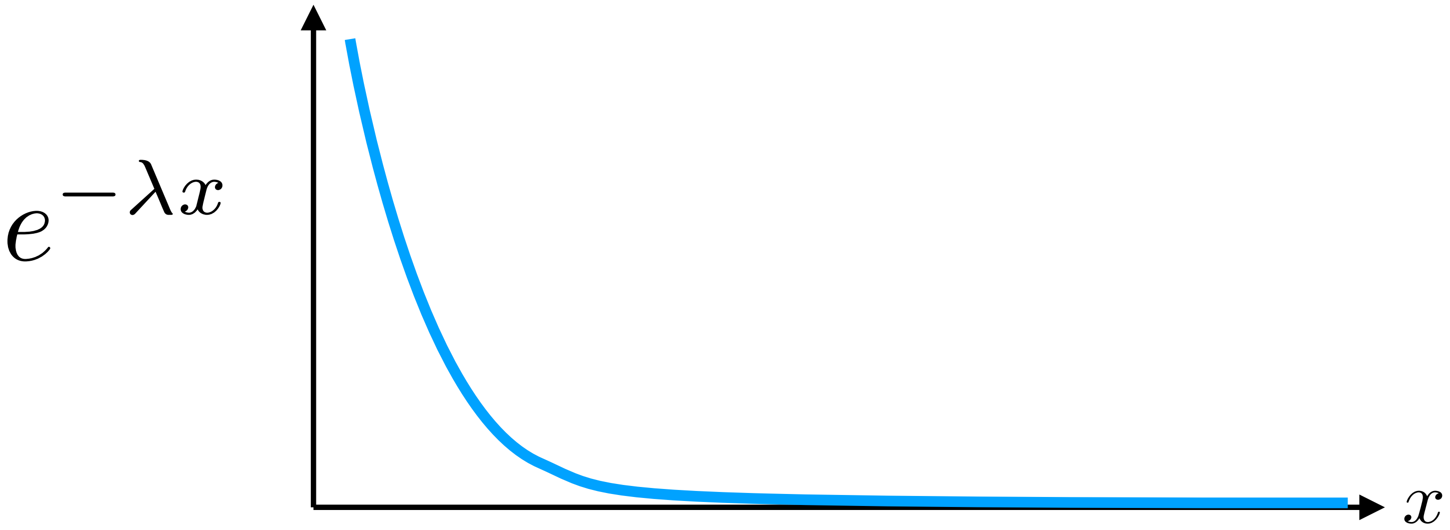




MERA Tensor Network

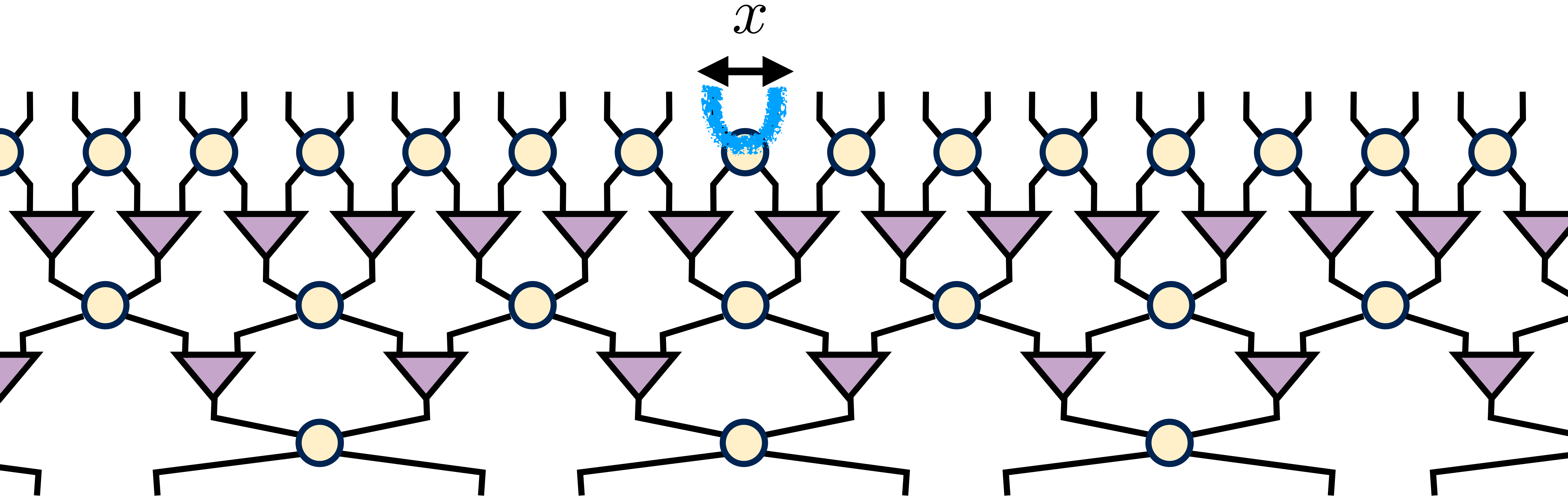
Matrix product state captures only exponential correlations





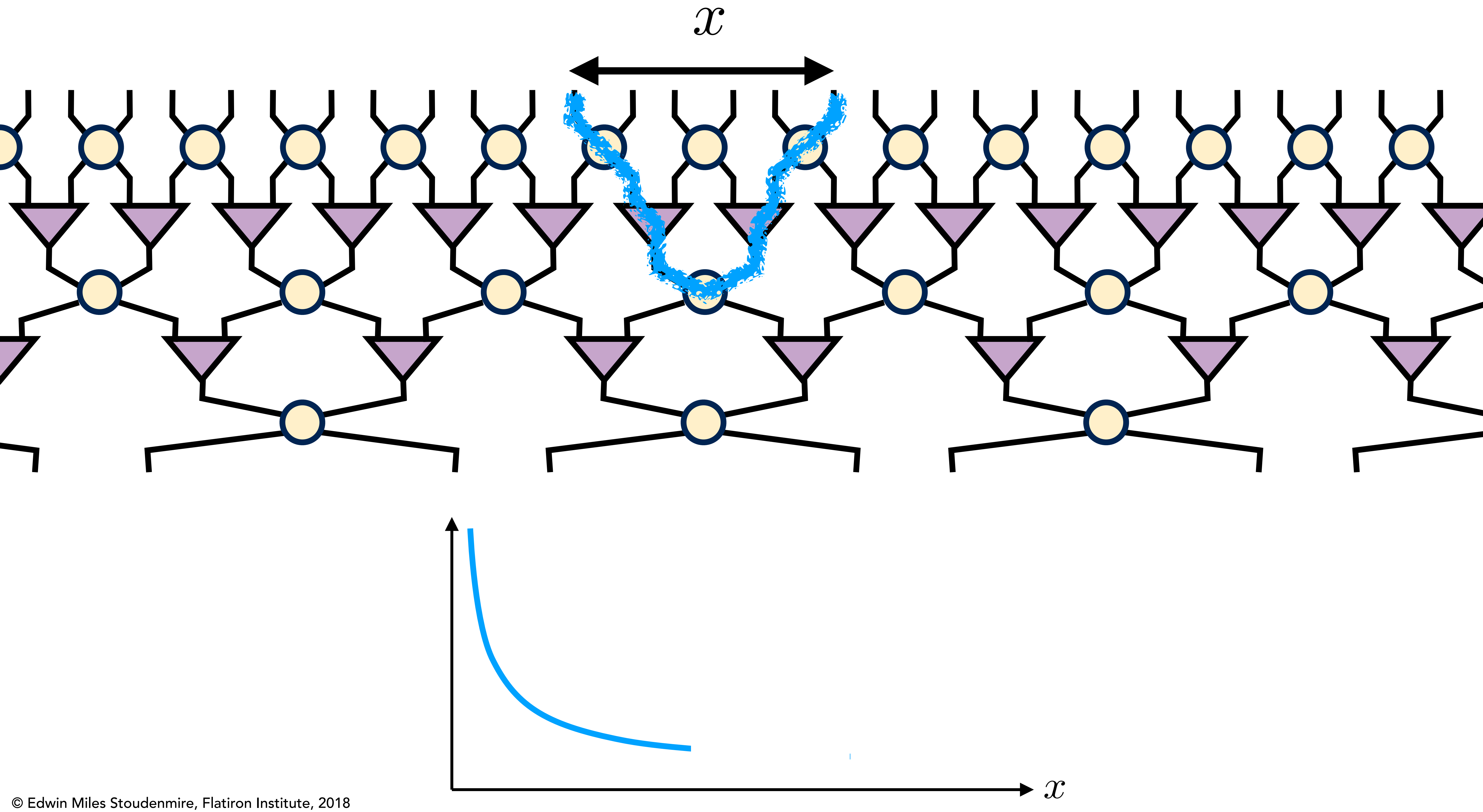
MERA Tensor Network

MERA layered architecture captures
power-law correlations



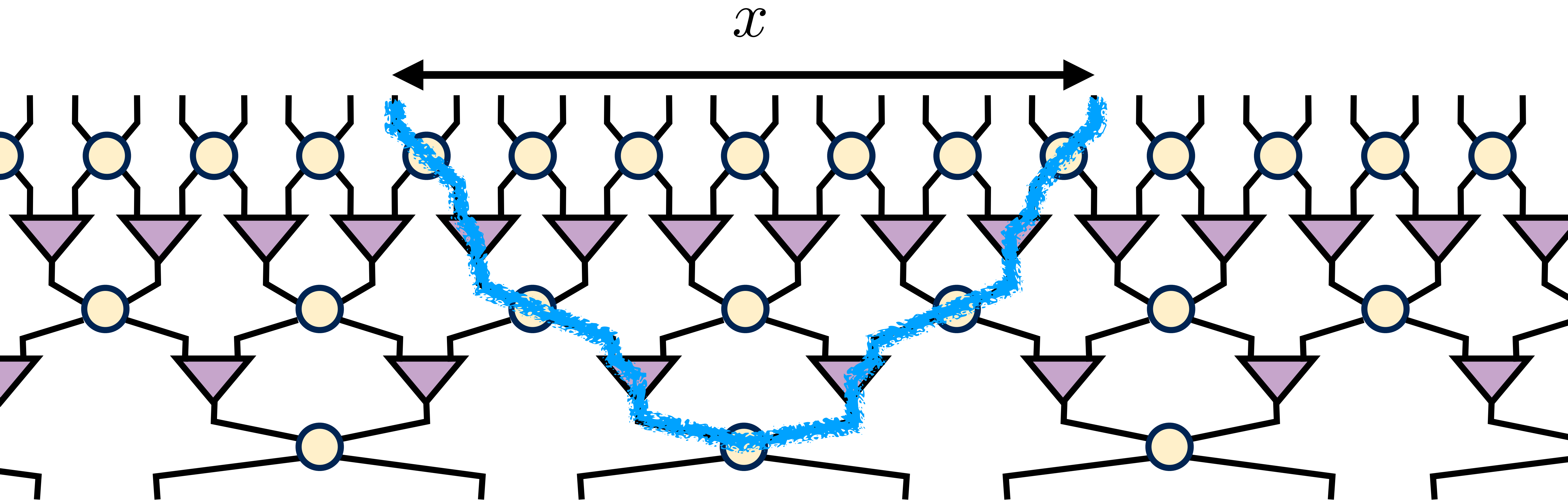
MERA Tensor Network

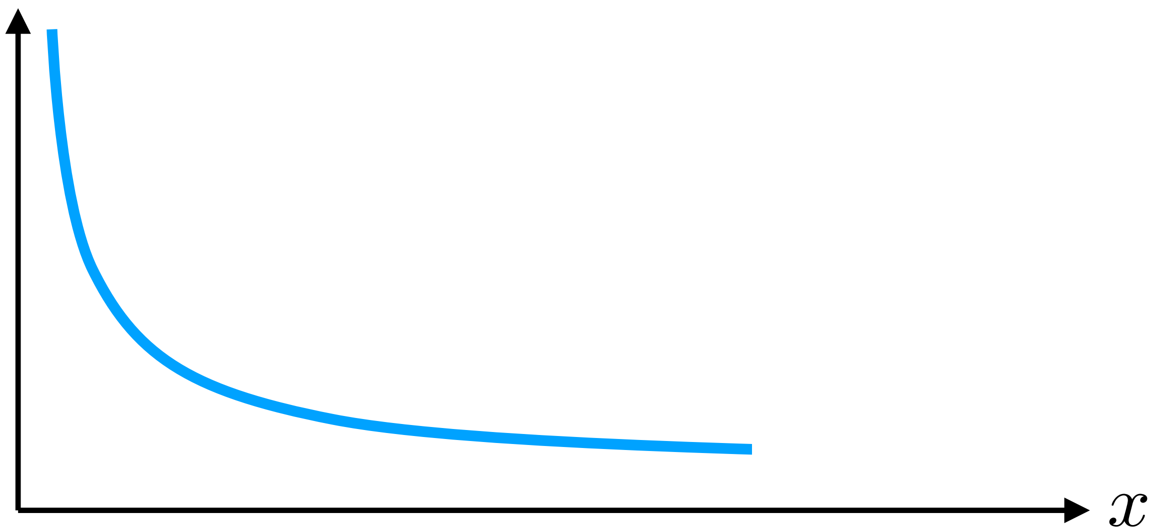
MERA layered architecture captures *power-law correlations*



MERA Tensor Network

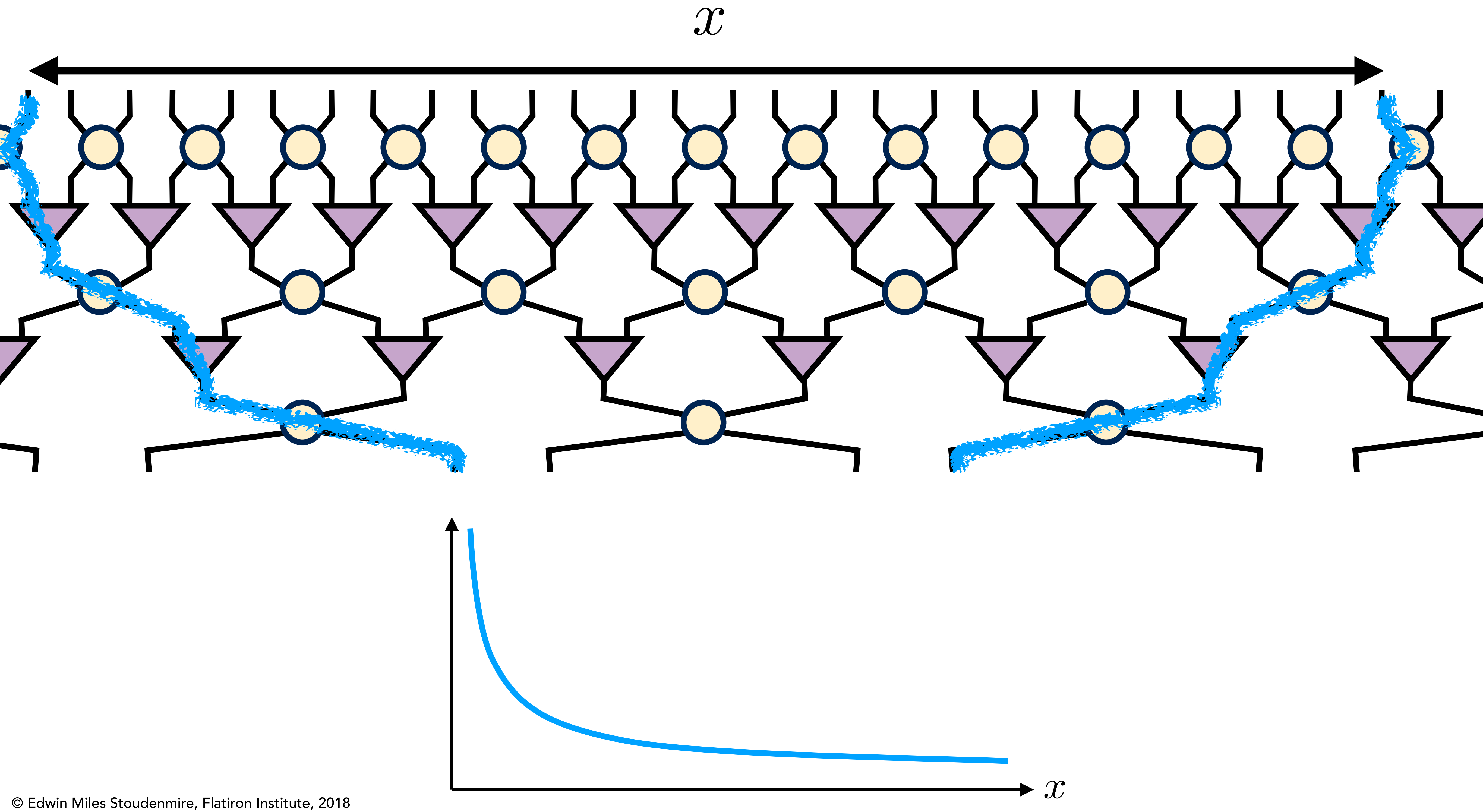
MERA layered architecture captures
power-law correlations





MERA Tensor Network

MERA layered architecture captures *power-law correlations*



Introduction to DMRG for Quantum Chemistry

A large part of quantum chemistry is calculating energies of molecules within the Born-Oppenheimer approximation

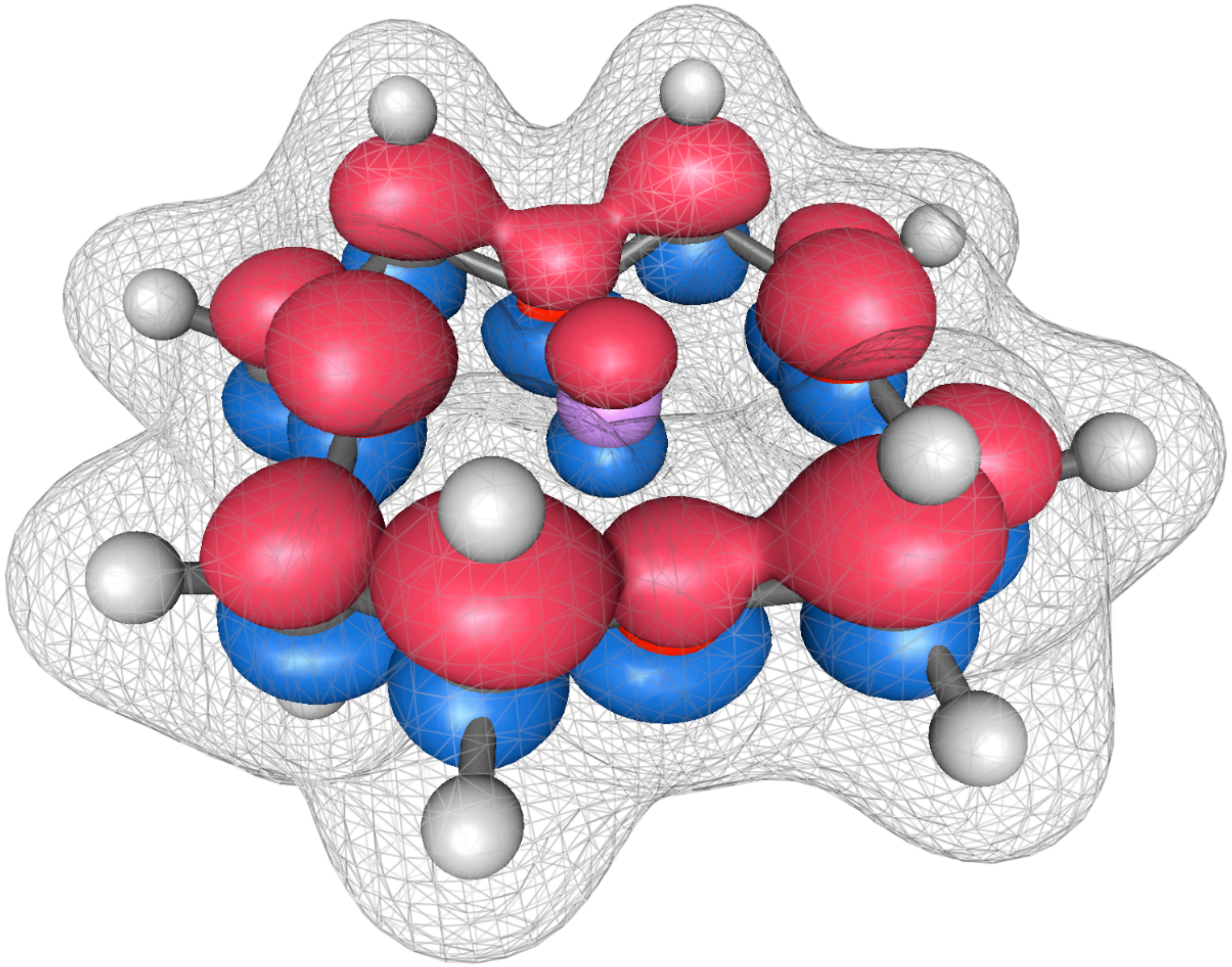


Figure credit: iqmol.org

A standard approach pioneered by John Pople is to use *Gaussian basis functions* to approximate the continuum

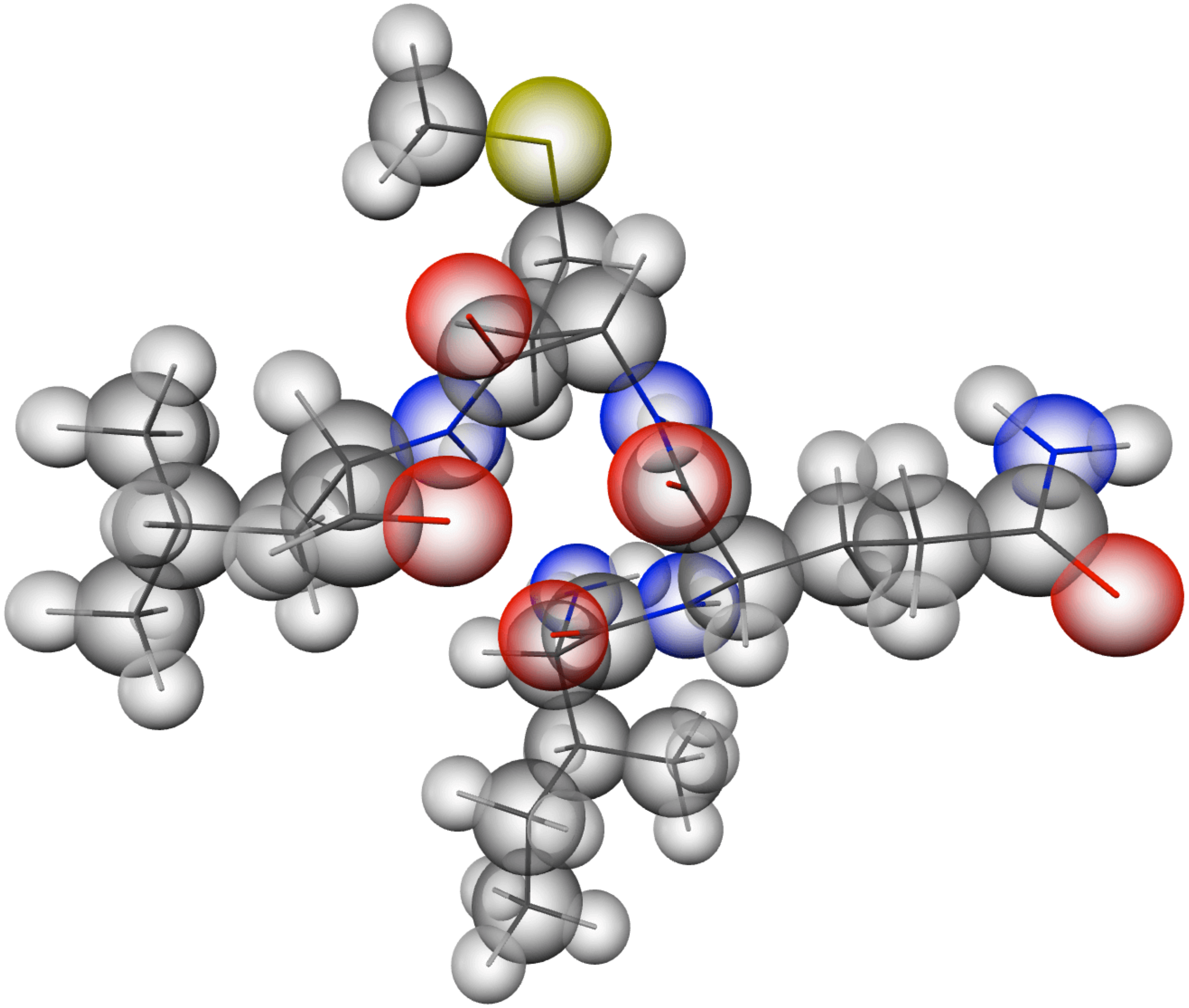
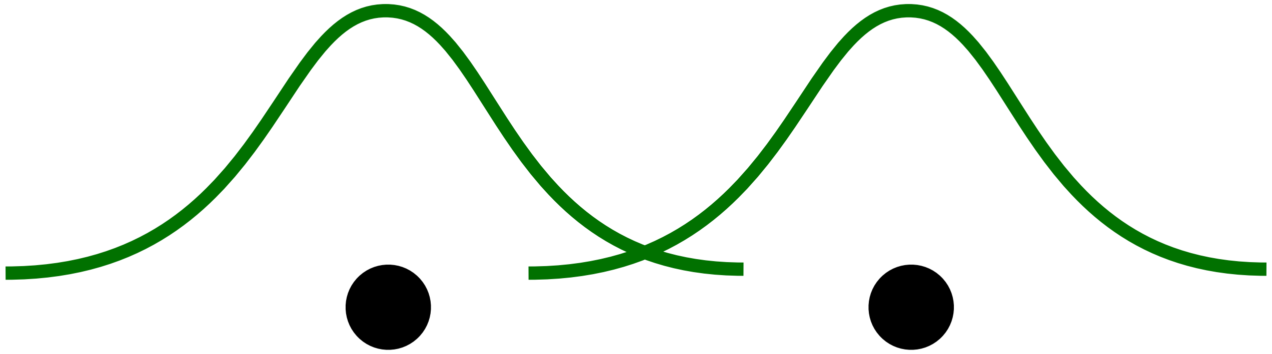


Figure credit: iqmol.org

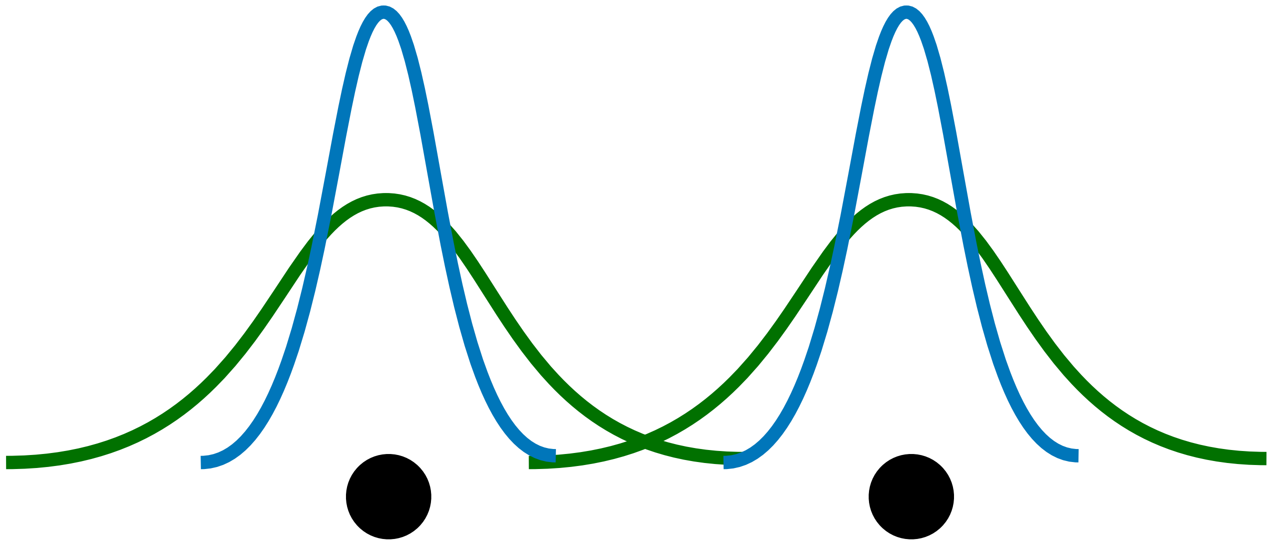
Consider an H_2 molecule

Cartoon of Gaussian basis sets:



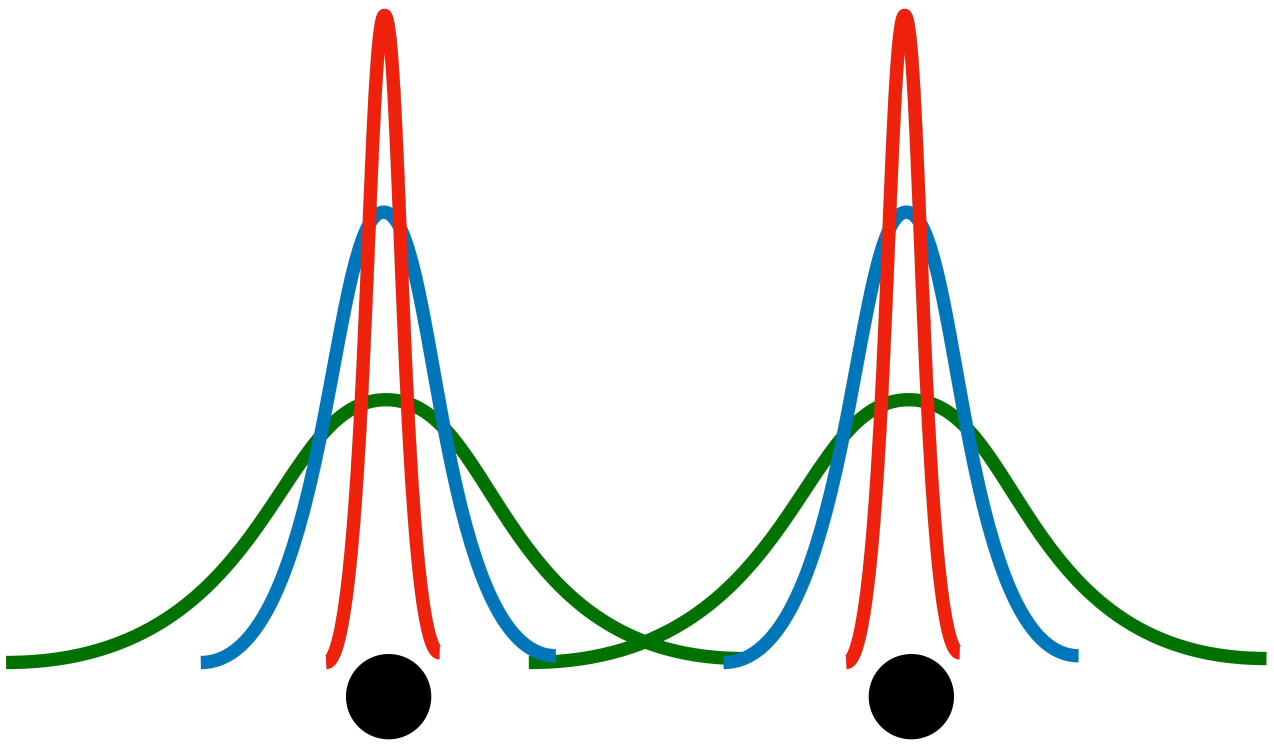
Consider an H_2 molecule

Cartoon of Gaussian basis sets:



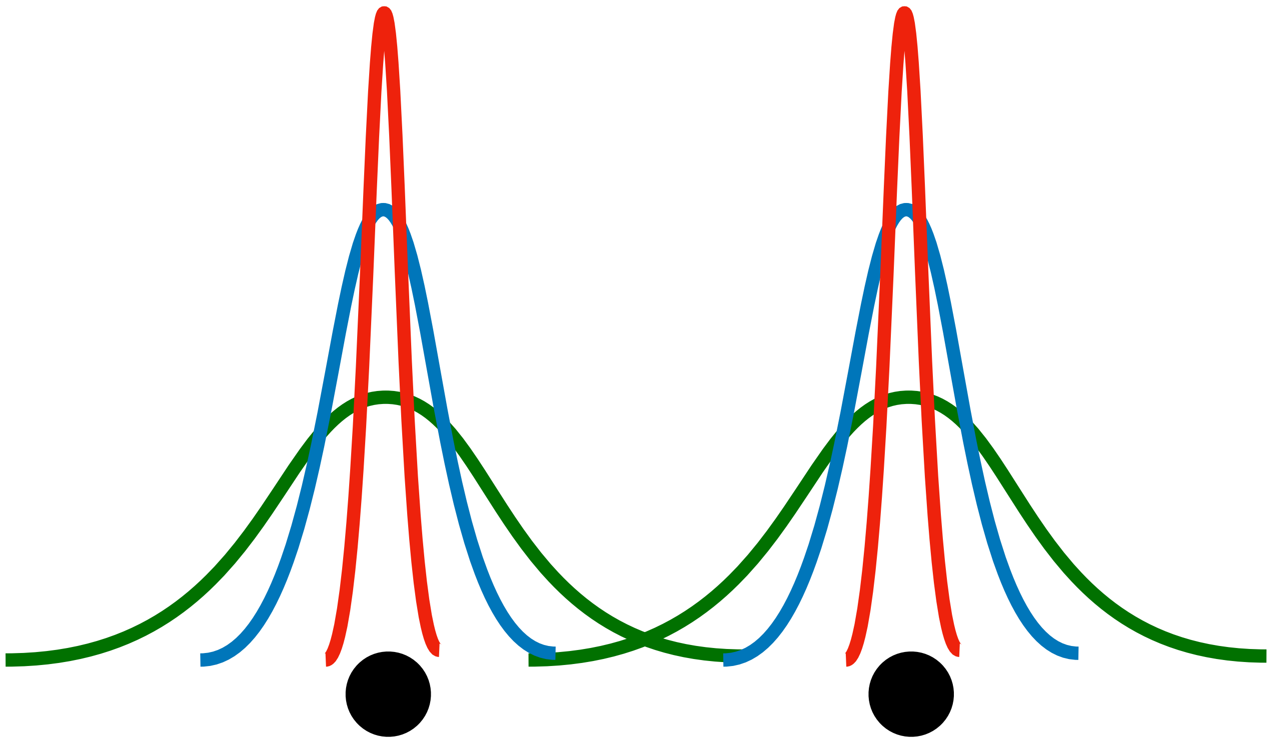
Consider an H_2 molecule

Cartoon of Gaussian basis sets:



Consider an H_2 molecule

Cartoon of Gaussian basis sets:



Basis sets also include linear combinations of Gaussians:

$$b_n(\mathbf{r}) = \sum_{i=1}^{N_n} c_{n,i} e^{-\zeta_{n,i}(\mathbf{r}-\mathbf{r}_A)^2}$$

And multiplicative factors: $x^p y^q z^s$

Fixed nuclei Hamiltonian – "electronic structure" problem:

$$H = \int_{\mathbf{r}} \hat{\psi}^\dagger_\sigma(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right] \hat{\psi}_\sigma(\mathbf{r})$$

$$+\frac{1}{2}\int_{\mathbf{r}\mathbf{r}'} \frac{1}{|\mathbf{r}-\mathbf{r}'|} \hat{\psi}^\dagger_\sigma(\mathbf{r}) \hat{\psi}^\dagger_{\sigma'}(\mathbf{r}') \hat{\psi}_{\sigma'}(\mathbf{r}') \hat{\psi}_\sigma(\mathbf{r})$$

Basis set approach: define $\hat{c}_n = \int_{\mathbf{r}} \phi_n(\mathbf{r}) \hat{\psi}(\mathbf{r})$

transform to discrete $\hat{c}_n$ basis and compute ground state

Orbital basis Hamiltonian (*i,j,k,l* label orbital 'sites'):

$$H = \sum_{ij} t_{ij} c^\dagger_{i\sigma} c_{j\sigma} + \sum_{ijkl} V_{ijkl} c^\dagger_{i\sigma} c^\dagger_{j\sigma'} c_{k\sigma'} c_{l\sigma}$$

$$t_{ij} = \int_{\mathbf{r}} \phi_i(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right] \phi_j(\mathbf{r})$$

$$V_{ijkl} = \int_{\mathbf{r}_1,\mathbf{r}_2} \frac{\phi_i(\mathbf{r_1})\phi_j(\mathbf{r}_2)\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

Orbital basis Hamiltonian (*i,j,k,l* label orbital 'sites'):

$$H = \sum_{ij} t_{ij} c^\dagger_{i\sigma} c_{j\sigma} + \sum_{ijkl} V_{ijkl} c^\dagger_{i\sigma} c^\dagger_{j\sigma'} c_{k\sigma'} c_{l\sigma}$$

$$t_{ij} = \int_{\mathbf{r}} \phi_i(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right] \phi_j(\mathbf{r})$$

$$V_{ijkl} = \int_{\mathbf{r}_1,\mathbf{r}_2} \frac{\phi_i(\mathbf{r_1})\phi_j(\mathbf{r}_2)\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

Point of using Gaussians is computing integrals quickly!

Wavefunction in orbital basis:

$$|\Psi\rangle = \sum_{\{n\}} \Psi^{n_1 n_2 \cdots n_k} |n_1 n_2 \cdots n_k\rangle$$

$n_j = \{0,\ \uparrow,\ \downarrow,\ \uparrow\downarrow\}$ orbital occupancy basis

$\Psi^{n_1 n_2 \cdots n_k}$ tensor with 4^k components

MPS Approximation

Simple approximation: single Slater determinant

$$\Psi^{n_1 n_2 \cdots n_k} \approx \sum_{\{i\}} \psi^{n_1} \psi^{n_2} \psi^{n_3} \ldots \psi^{n_k}$$

Occupancy of orbitals independent of each other

How to improve?

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

MPS Approximation

Add internal indices ⟶ include correlations

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

MPS Approximation

Add internal indices ⟶ include correlations

$$\Psi^{n_1 n_2 \cdots n_k} \approx \sum_{\{i\}} \psi^{n_1} \psi^{n_2} \psi^{n_3} \ldots \psi^{n_k}$$

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

MPS Approximation

Add internal indices ⟶ include correlations

$$\Psi^{n_1 n_2 \cdots n_k} \approx \sum_{\{i\}} \psi^{n_1}_{i_1} \psi^{n_2}_{i_1} \psi^{n_3} \ldots \psi^{n_k}$$

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

MPS Approximation

Add internal indices ⟶ include correlations

$$\Psi^{n_1 n_2 \cdots n_k} \approx \sum_{\{i\}} \psi^{n_1}_{i_1} \psi^{n_2}_{i_1 i_2} \psi^{n_3}_{i_2} \ldots \psi^{n_k}$$

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

MPS Approximation

Add internal indices ⟶ include correlations

$$\Psi^{n_1 n_2 \cdots n_k} \approx \sum_{\{i\}} \psi^{n_1}_{i_1} \psi^{n_2}_{i_1 i_2} \psi^{n_3}_{i_2 i_3} \cdots \psi^{n_k}_{i_{k-1}}$$

Results in matrix product state form of wavefunction

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

DMRG for Quantum Chemistry

DMRG has many advantages for chemistry, especially when wavefunction *strongly correlated*

Modest-size MPS wavefunction can represent sum of <u>exponentially many</u> Slater determinants

Yet DMRG can struggle for:
- strong "dynamic" correlations
- molecules extended in 2D or 3D

Chan and Sharma, Annu. Rev. Phys. Chem. 2011.62:465-48

DMRG for Quantum Chemistry

But current basis-set DMRG approach has additional problems not fundamental to DMRG itself

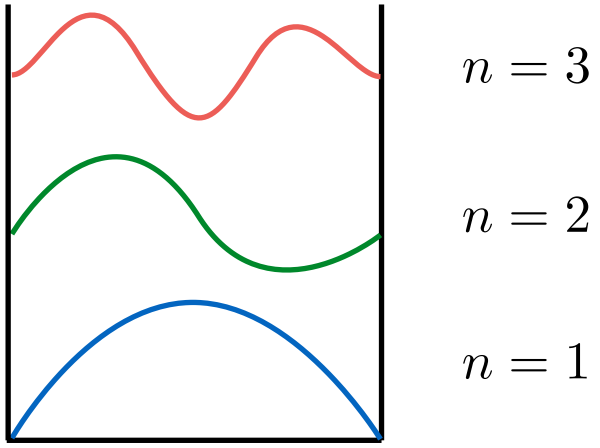
Orthogonalizing basis set produces long tails, despite efforts to localize functions

Resulting Hamiltonian has N^4 terms, and also DMRG wavefunction hard to represent (highly entangled)

Alternatives to pure basis-set approach?

Two Continuum Approaches

Consider 1D particles in a box



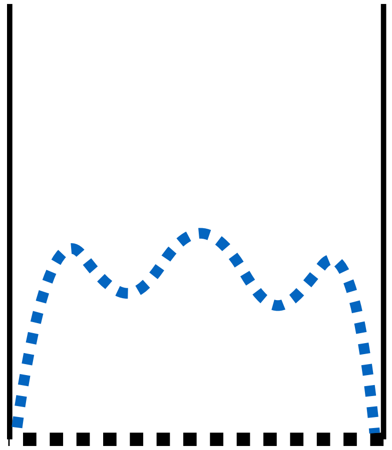
Approach 1: basis set $c_n = \int_x \phi_n(x)\hat{\psi}(x)$

$$H = -\frac{1}{2}\int_x \hat{\psi}^\dagger(x)\frac{\partial^2}{\partial x^2}\hat{\psi}(x) \quad \to \quad H = \sum_{nm} t_{nm} c^\dagger_n c_m$$

- Loss of locality
- Must compute integrals

\+ Variational

Two Continuum Approaches



Consider 1D particles in a box

Approach 2: grid approximation $c_j = \sqrt{a}\ \hat{\psi}(x_j)$

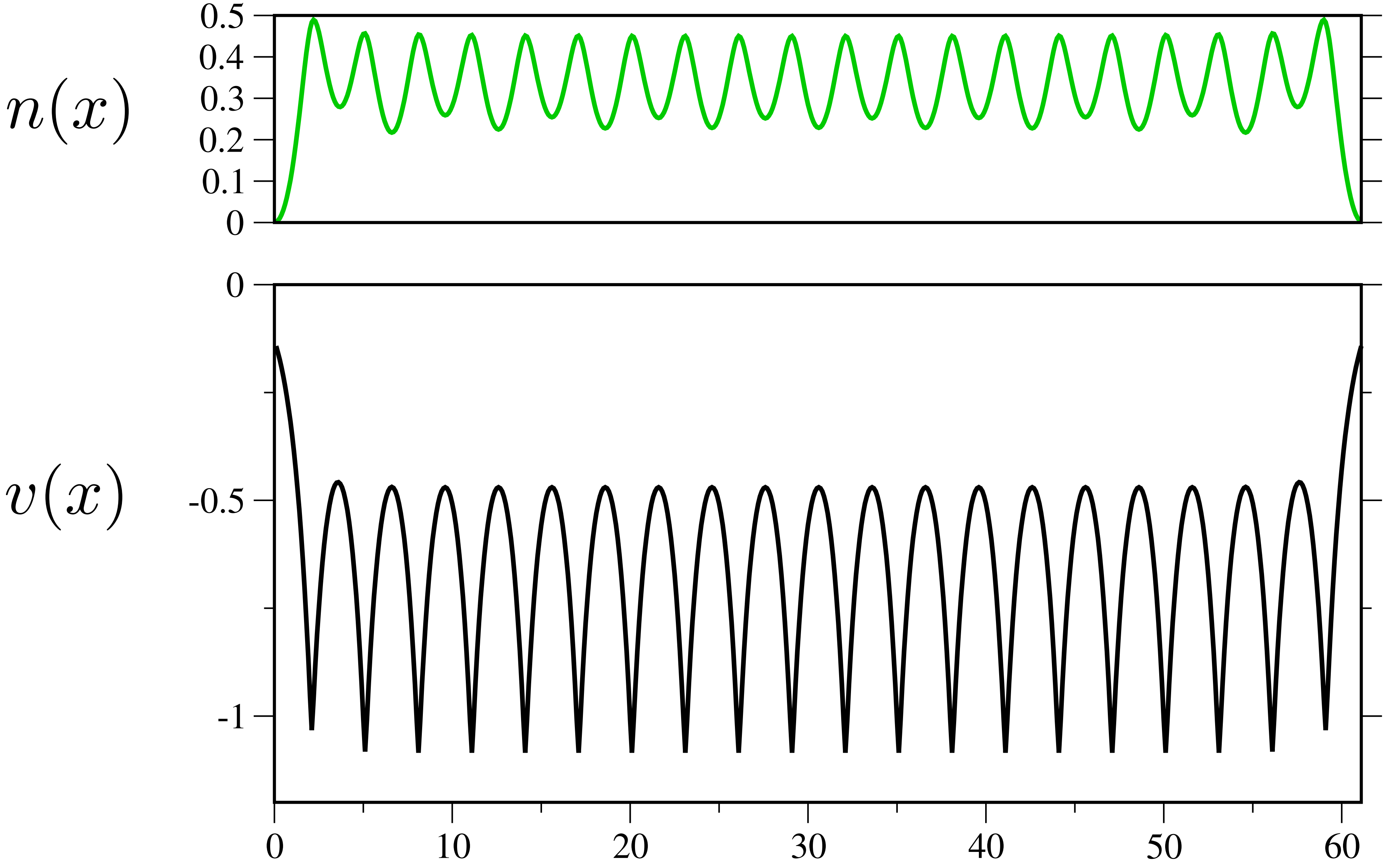
$$H = -\frac{1}{2}\int_x \hat{\psi}^\dagger(x)\frac{\partial^2}{\partial x^2}\hat{\psi}(x)$$

$$\searrow$$

$$H \approx -\frac{1}{2a^2}\sum_j (c_j^\dagger c_{j+1} - 2n_j + c_{j+1}^\dagger c_j) + \mathcal{O}(a^2)$$

- \+ Local / short range
- \+ No integrals to compute
- \- Not variational

Can use grid + DMRG for interacting many-body physics in 1D continuum:



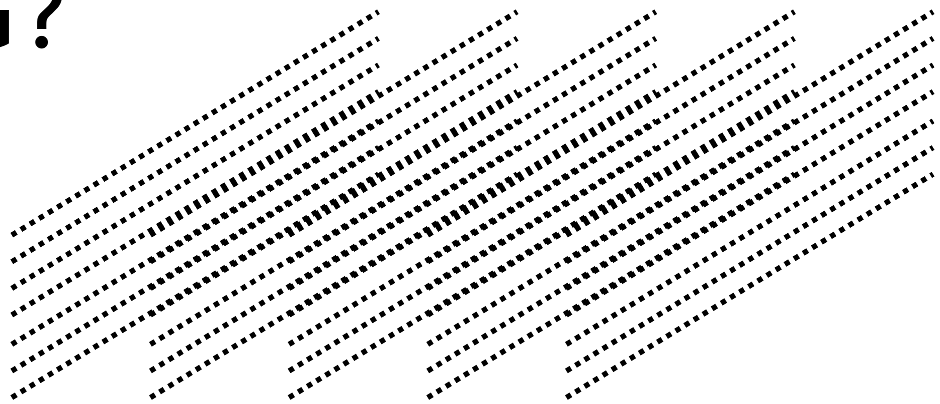
Wagner et al., PRB 90, 045109 (2014)

Stoudenmire, Wagner, White, Burke, PRL 109, 056402 (2012)

What about 3D continuum? (molecules, cold atoms)

1. Basis set DMRG: works well, but could scale better

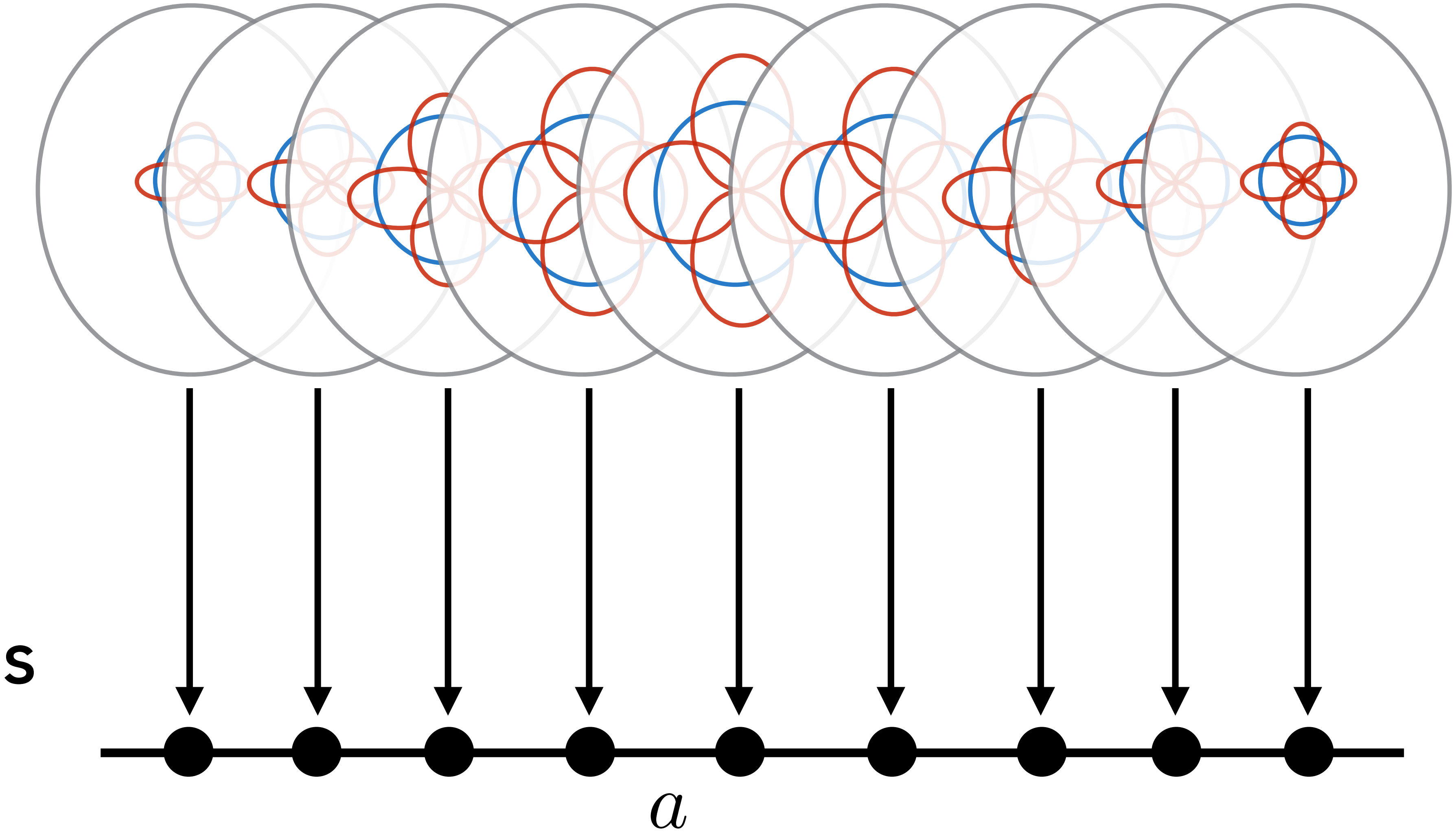
2. 3D Grid DMRG?



Too many sites!

3. Hybrid grid + basis set DMRG…

Slicing Approach



Slice 3D chemical basis sets along x-direction:

Map to 1D 'chain' with 1000's of sites (small $\Delta x = a$):

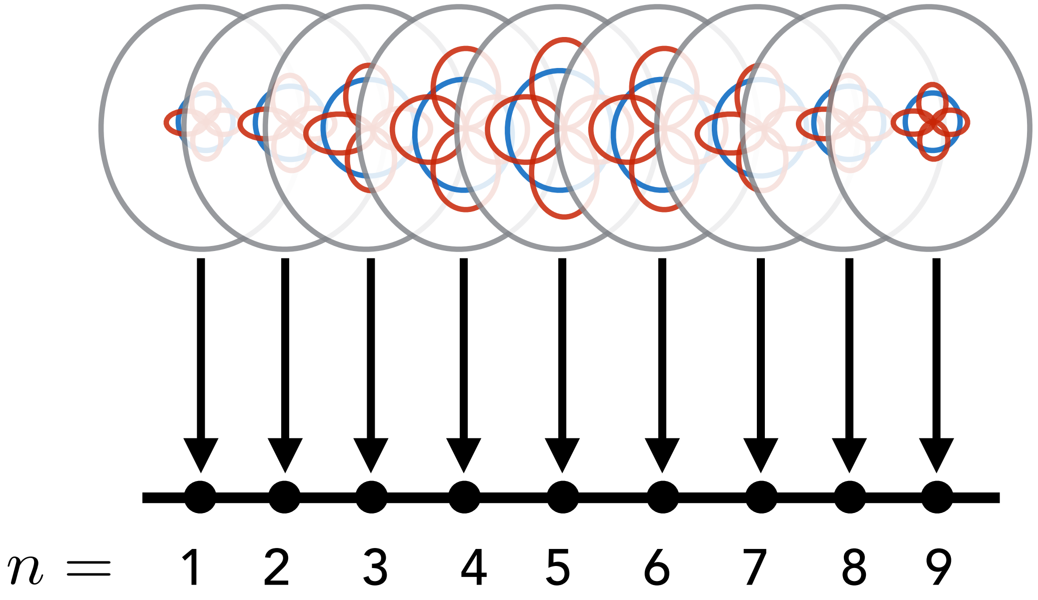
- Can reach chemical accuracy for $a \lesssim 0.1$
- Scalable to 1000's of atoms

Sliced Basis Set

Slices roughly equivalent to using basis set of "functions":

$$\phi_{nj}(\mathbf{r}) = \delta^{\frac{1}{2}}(x - x_n)\varphi_{nj}(y, z)$$

$$x_n = n\,a$$



Functions on different slices $n' \neq n$ automatically ortho.

In transverse direction, orthogonalize the basis set [$\varphi_{nj}(y, z)$ orthogonal for same n and $j' \neq j$]

Constructing a Sliced Basis

1. On each slice, collect standard basis functions from each atom, projected onto the slice

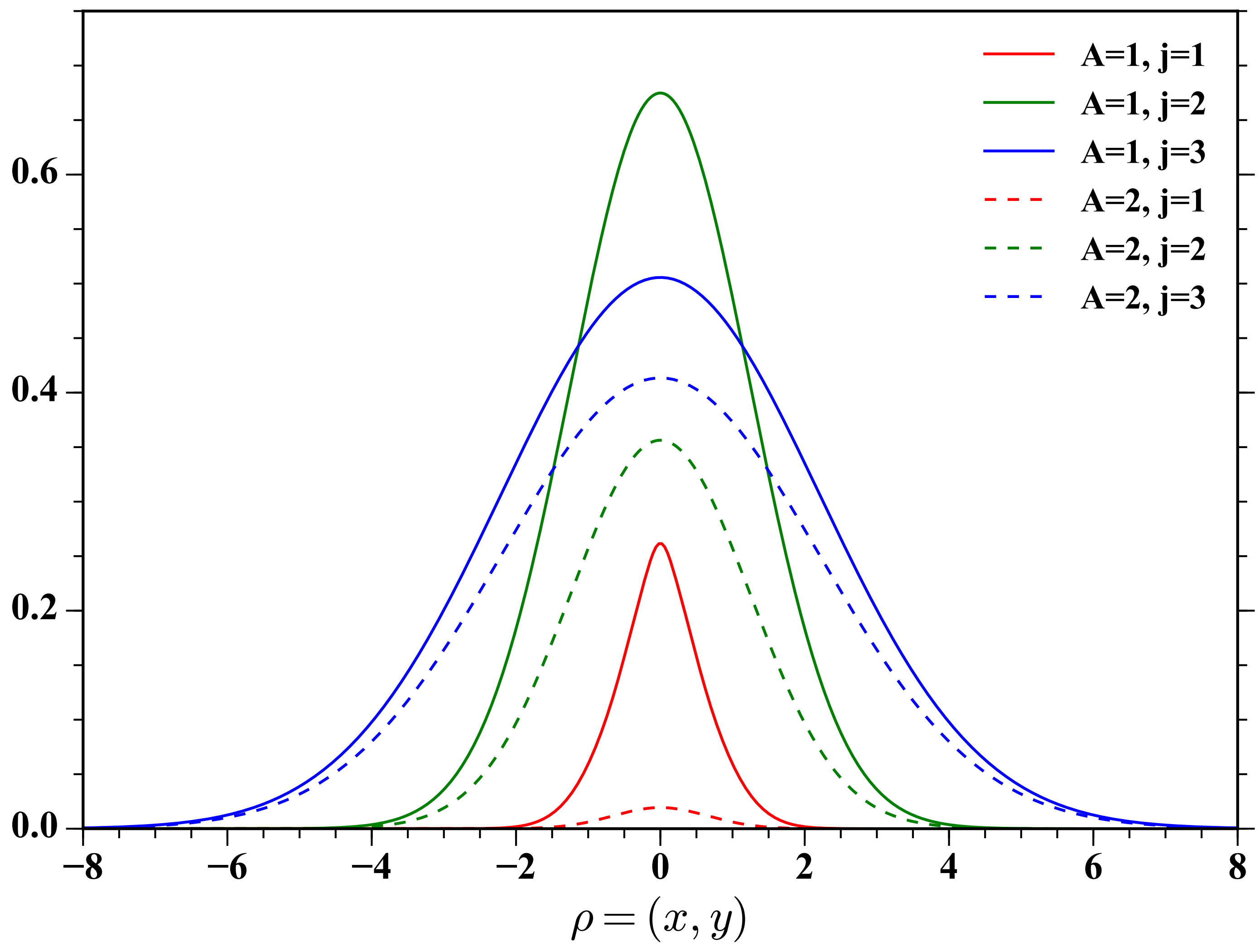
   $$\phi_{niA}(y,z) = \phi_{iA}(x_n,y,z)$$

   Projected functions are not normalized

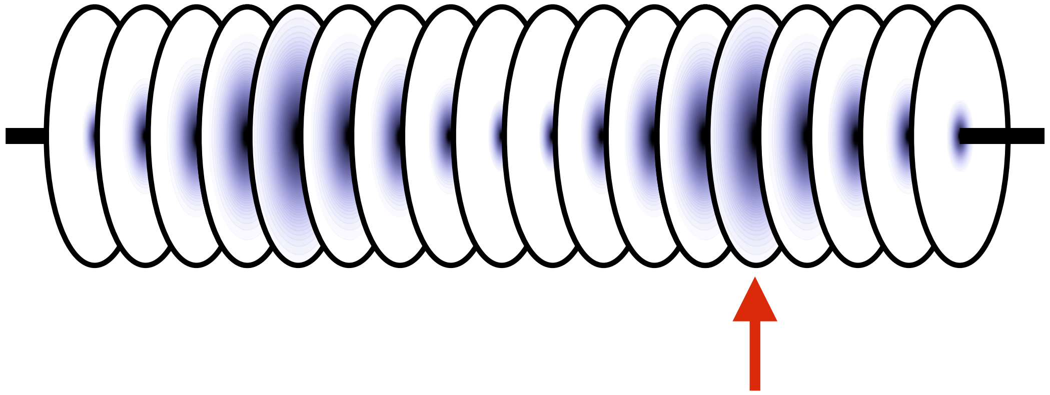
2. Orthogonalize projected functions symmetrically, truncating any with negligible contribution

Constructing a Sliced Basis

Example #1: H_2 molecule, slice through nucleus



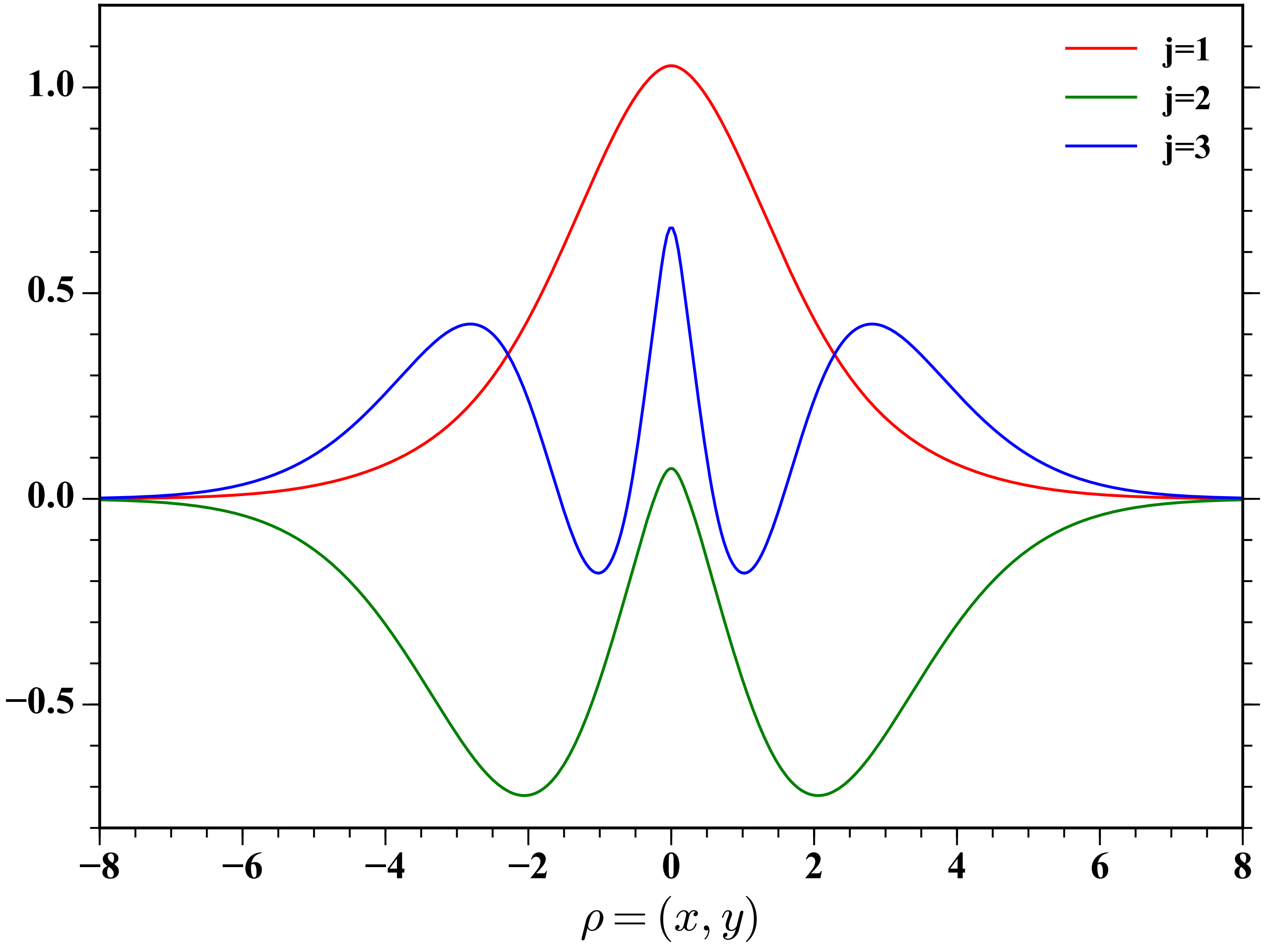
Original basis

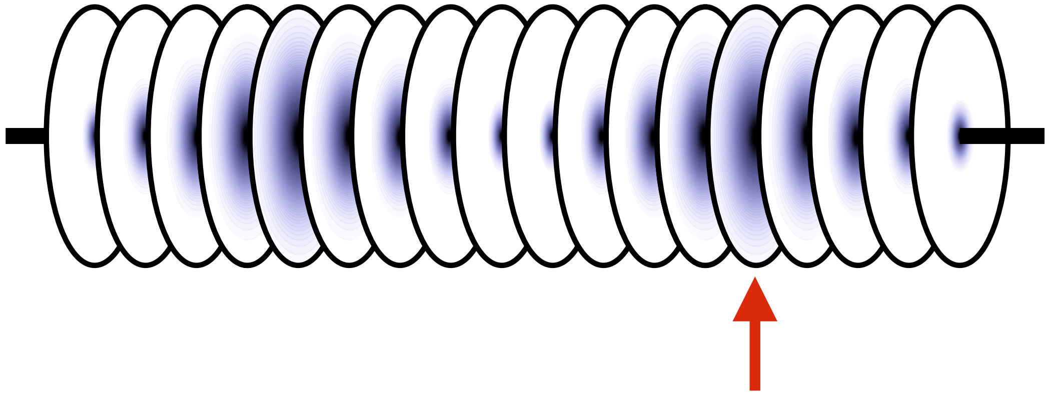


Orbitals taken from cc-pvTZ,
S orbitals only

Constructing a Sliced Basis

Example #1: H_2 molecule, slice through nucleus



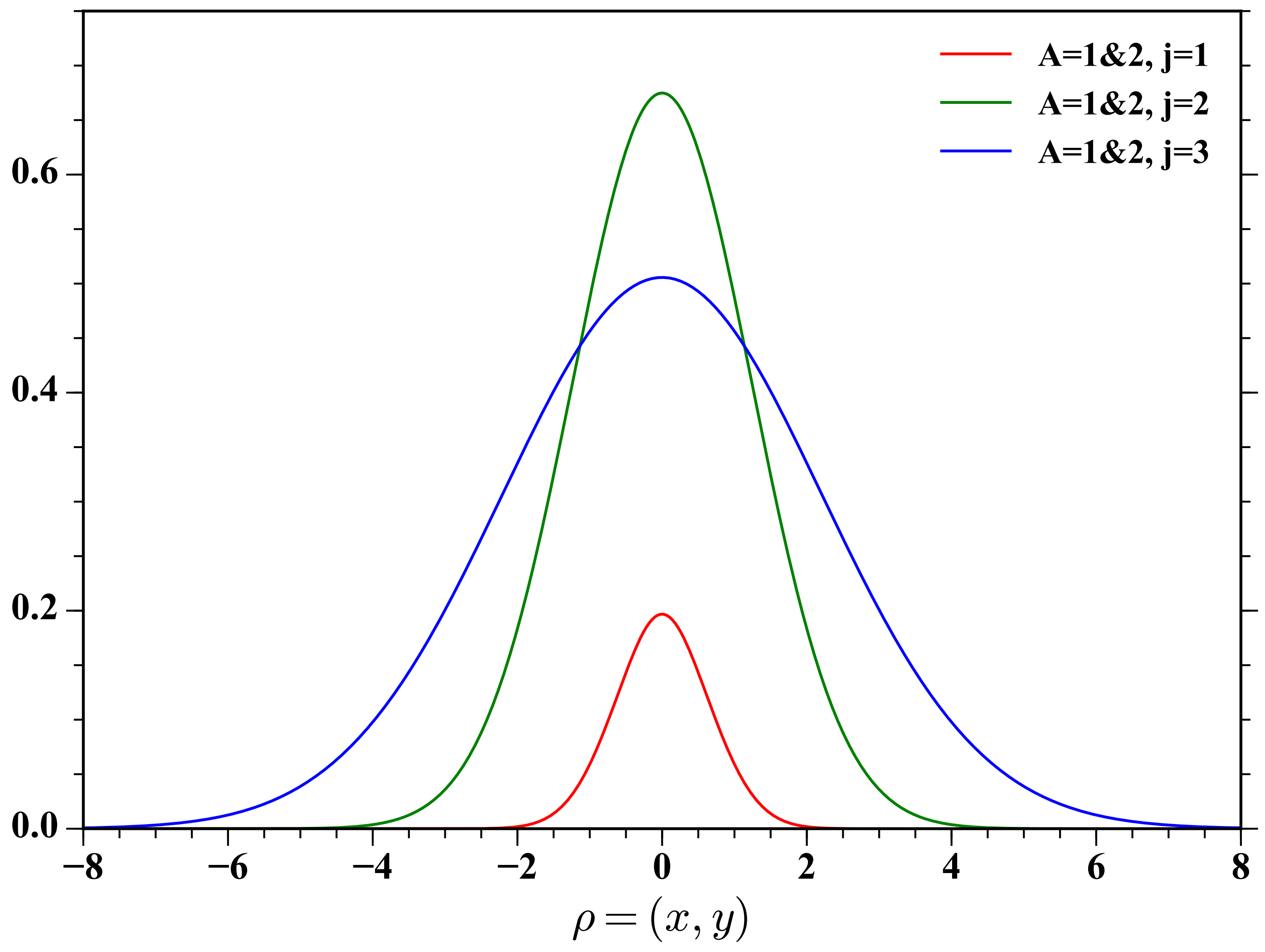


Orbitals taken from cc-pvTZ,
S orbitals only

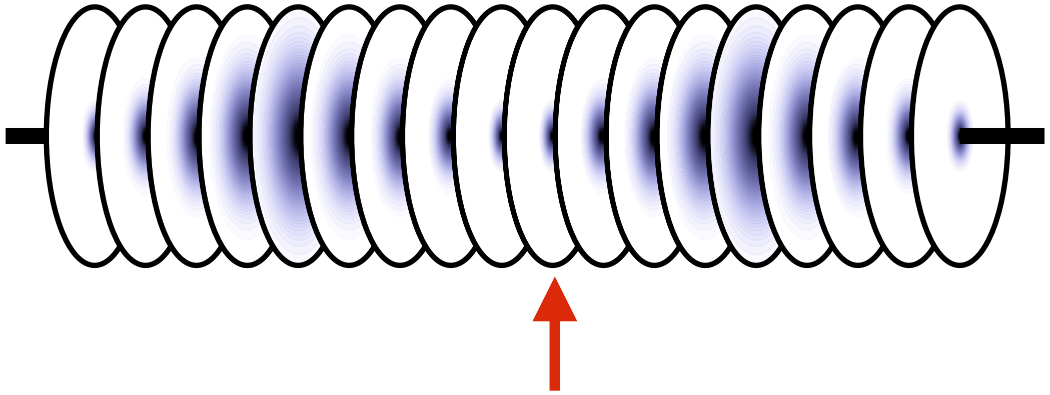
Orthogonal basis

Constructing a Sliced Basis

Example #2: H_2 molecule, slice through bond



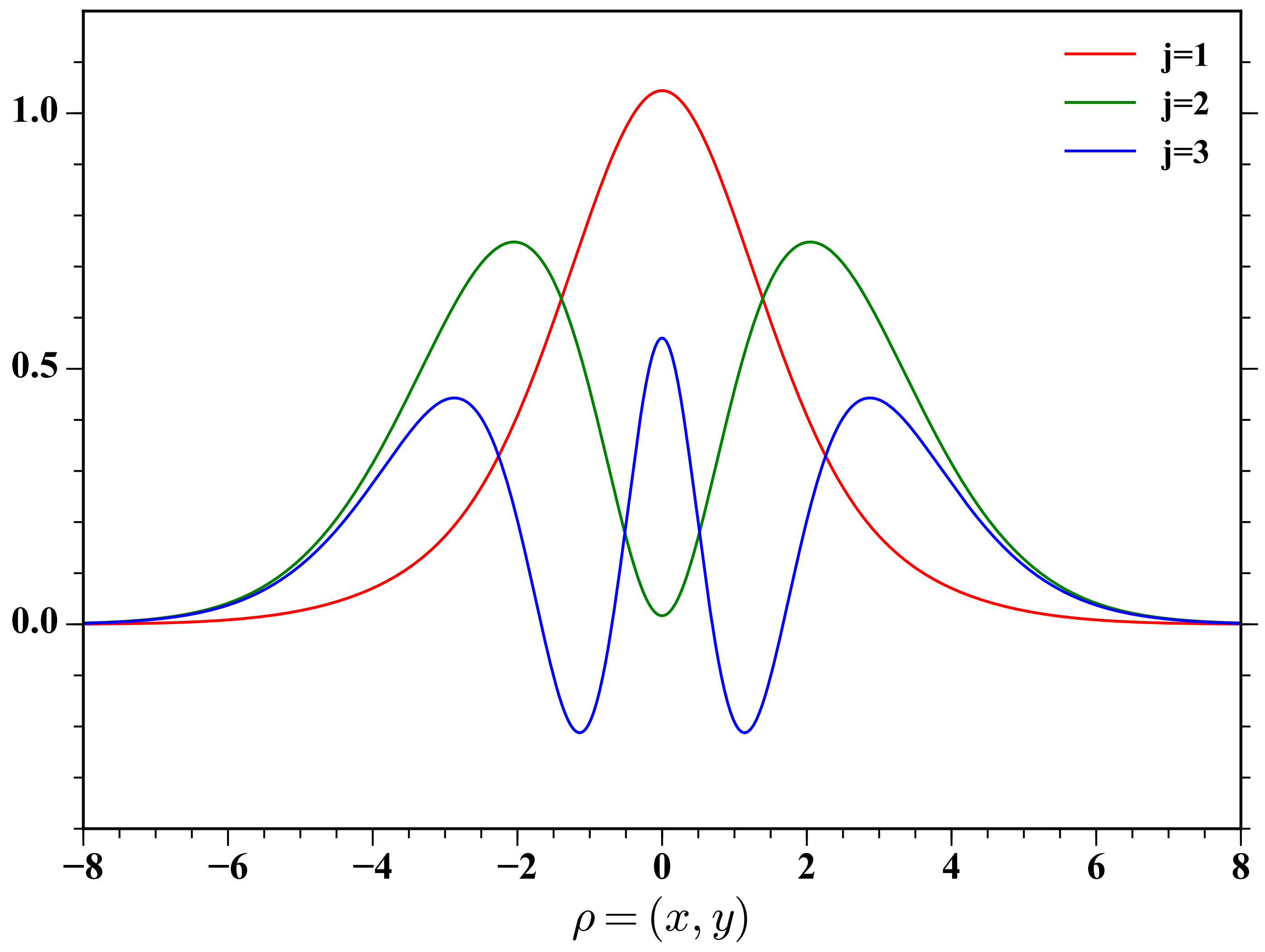
Original basis

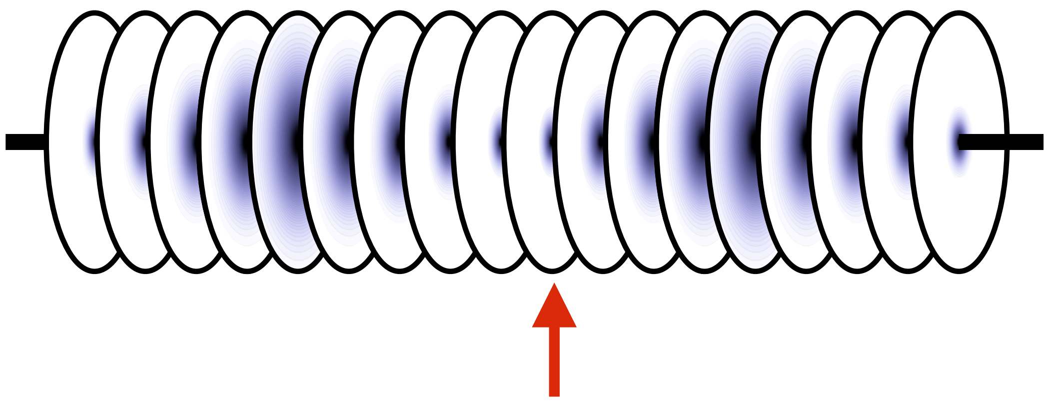


Orbitals taken from cc-pvTZ,
S orbitals only

Constructing a Sliced Basis

Example #2: H_2 molecule, slice through bond





Orbitals taken from cc-pvTZ,
S orbitals only

Orthogonal basis

Sliced Basis Set

$$\phi_{nj}(\mathbf{r}) = \delta^{\frac{1}{2}}(x - x_n)\varphi_{nj}(y, z)$$

Kinetic Energy:

Delta function in x really means "grid approximation" to Hamiltonian

Replace $\frac{\partial^2}{\partial x^2}$ with higher-order discretized derivative

Not variational, but can keep grid error below basis error (controlled: goes to zero as grid spacing "a" goes to zero)

Sliced Basis Set

$$\phi_{nj}(\mathbf{r}) = \delta^{\frac{1}{2}}(x - x_n)\varphi_{nj}(y, z)$$

Interaction Energy:

Normally must deal with N^4 interaction terms

$$V_{ijkl} = \int_{\mathbf{r}_1,\mathbf{r}_2} \frac{\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

Sliced Basis Set

$$\phi_{nj}(\mathbf{r}) = \delta^{\frac{1}{2}}(x - x_n)\varphi_{nj}(y,z)$$

Interaction Energy:

Normally must deal with N^4 interaction terms

$$V_{ijkl} = \int_{\mathbf{r}_1,\mathbf{r}_2} \frac{\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

But treat slices as orthogonal. Then V_{ijkl} non-zero only if i,l on same slice and j,k on same slice

Sliced Basis Set

$$\phi_{nj}(\mathbf{r}) = \delta^{\frac{1}{2}}(x - x_n)\varphi_{nj}(y,z)$$

Interaction Energy:

Normally must deal with N^4 interaction terms

$$V_{ijkl} = \int_{\mathbf{r}_1,\mathbf{r}_2} \frac{\phi_i(\mathbf{r_1})\phi_j(\mathbf{r}_2)\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

But treat slices as orthogonal. Then V_{ijkl} non-zero only if i,l on same slice and j,k on same slice

Number of terms: $N^4 \rightarrow N_x^2 \; N_{\text{orb}}^4$

Form of Hamiltonian

Start from fixed nuclei Hamiltonian:

$$H = \int_{\mathbf{r}} \hat{\psi}^{\dagger}_{\sigma}(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right] \hat{\psi}_{\sigma}(\mathbf{r})$$

$$+ \frac{1}{2}\int_{\mathbf{r}\mathbf{r}'} v_{\mathrm{ee}}(\mathbf{r}-\mathbf{r}')\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\hat{\psi}^{\dagger}_{\sigma'}(\mathbf{r}')\hat{\psi}_{\sigma'}(\mathbf{r}')\hat{\psi}_{\sigma}(\mathbf{r})$$

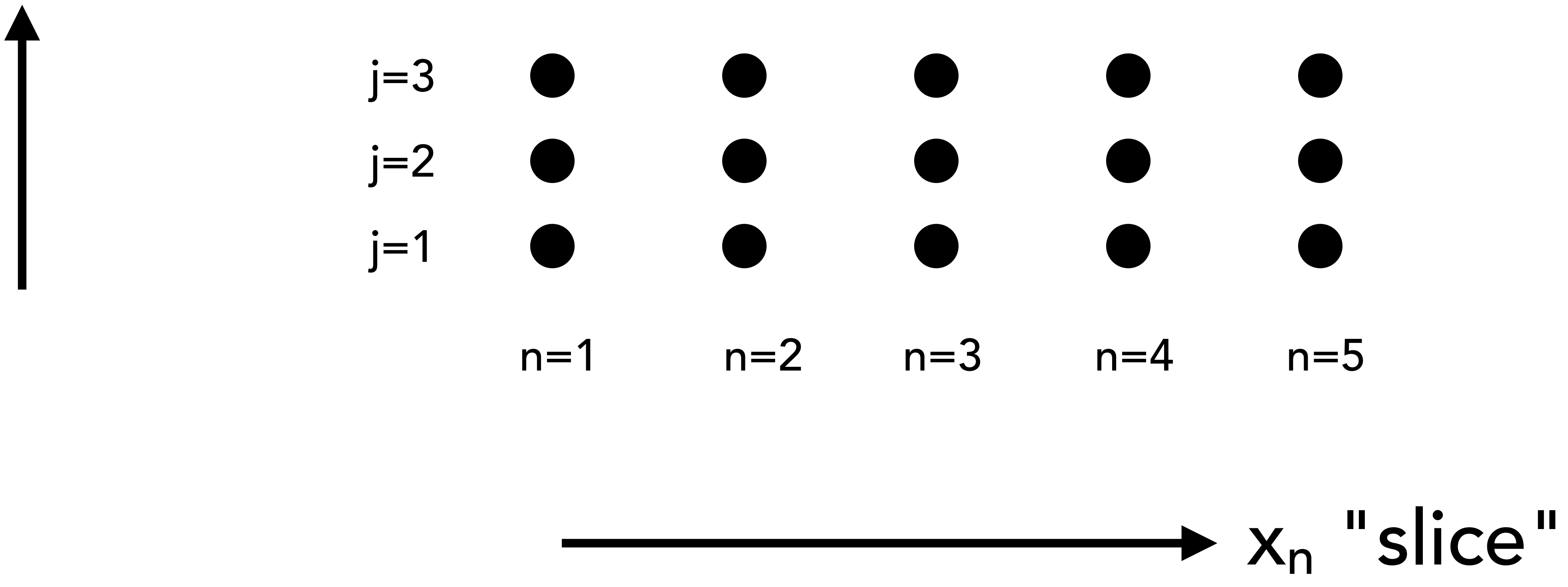
Define $\hat{c}_{nj\sigma} = \sqrt{a}\int_{y,z} \varphi_{nj}(x,y)\ \hat{\psi}_{\sigma}(x_n,y,z)$

and transform to $\hat{c}_{nj\sigma}$ basis

Form of Hamiltonian

Discrete (n,j) labels can be viewed as a "ladder" lattice

orbital # j



j=3

j=2

j=1

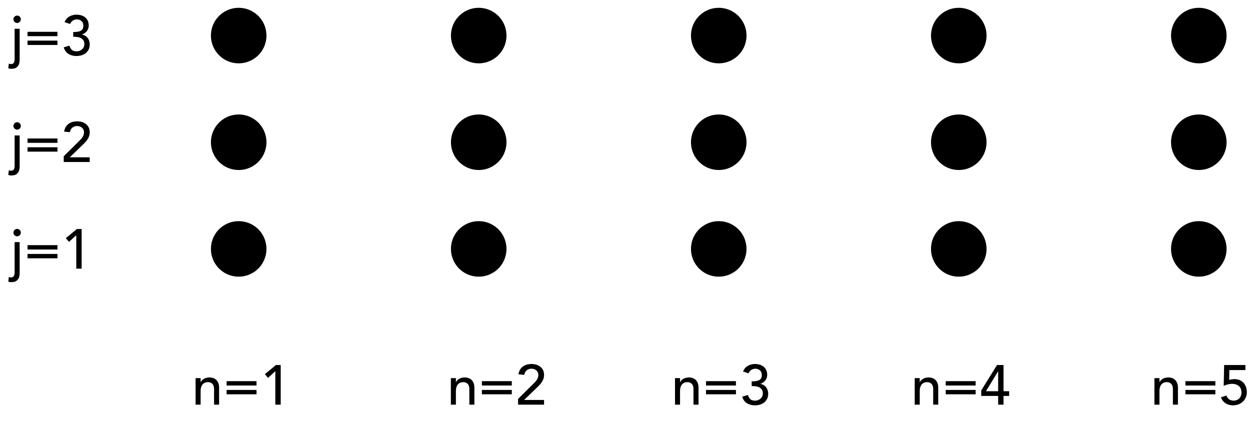
n=1 n=2 n=3 n=4 n=5

x_n "slice"

Form of Hamiltonian

Hamiltonian turns into three pieces

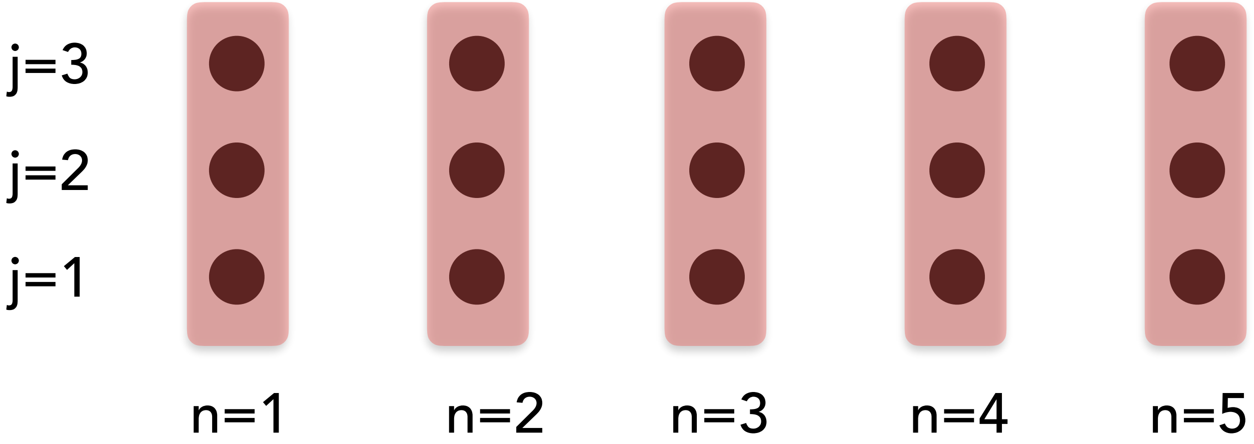
$$H =$$



Form of Hamiltonian

Hamiltonian turns into three pieces

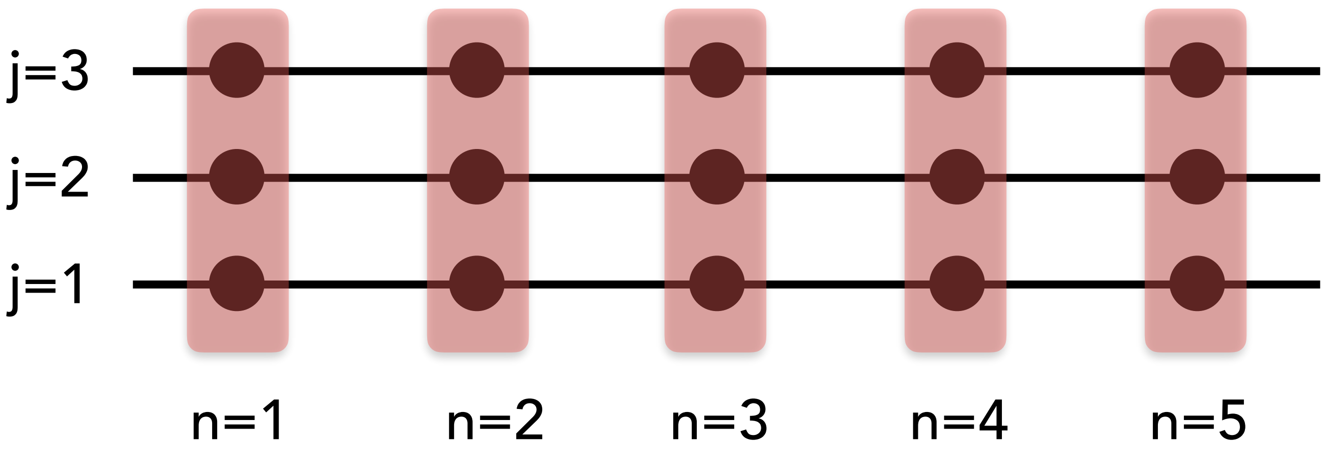
$$H = \sum_n H_n^{(\text{slice})}$$



Form of Hamiltonian

Hamiltonian turns into three pieces

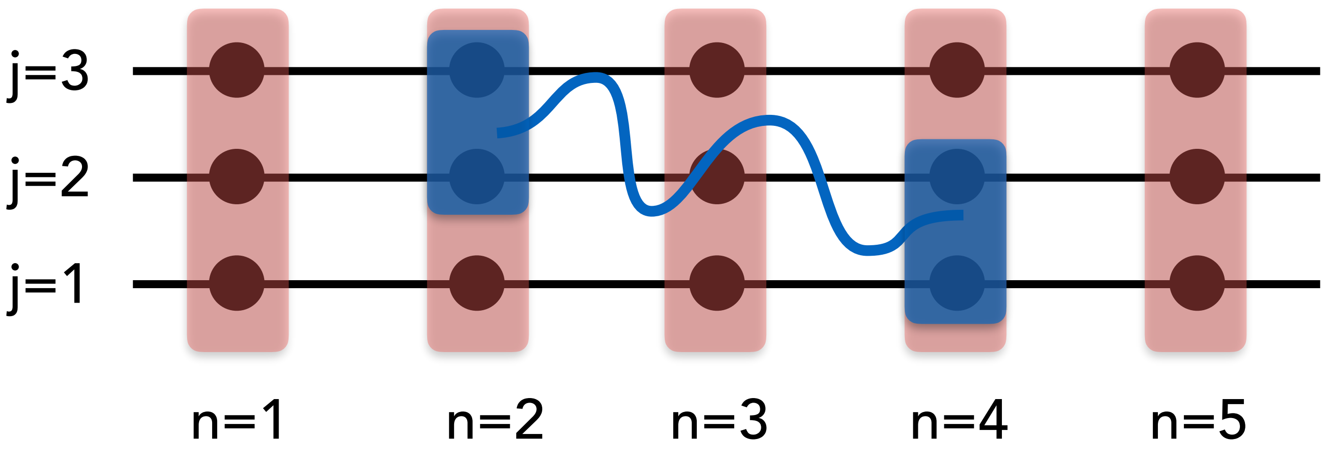
$$H = \sum_n H_n^{(\text{slice})} + T^{(\text{grid})}$$



Form of Hamiltonian

Hamiltonian turns into three pieces

$$H = \sum_n H_n^{(\mathrm{slice})} + T^{(\mathrm{grid})} + V$$



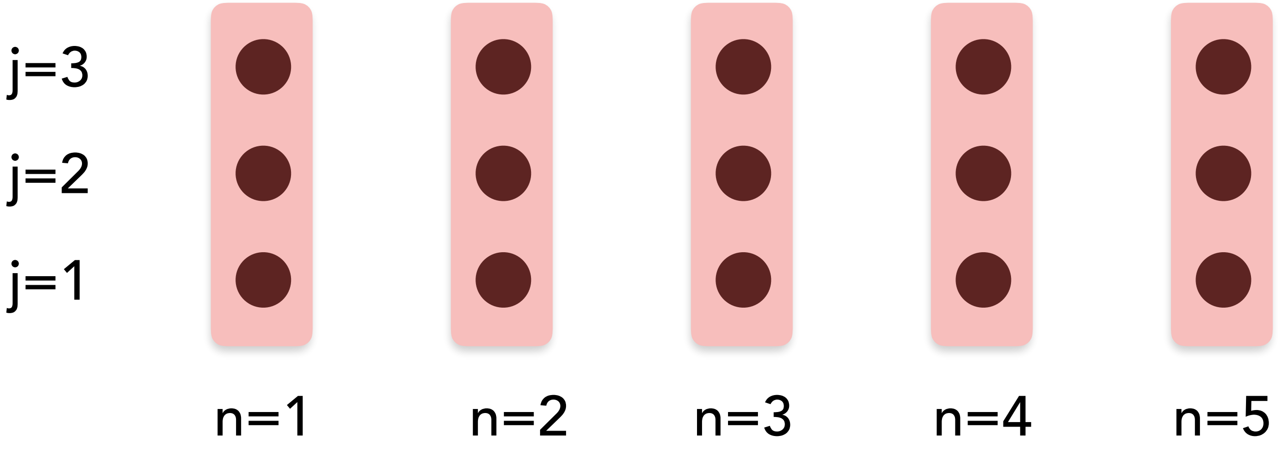
Form of Hamiltonian

$$H = \sum_n H_n^{(\text{slice})} + T^{(\text{grid})} + V$$

$$H_n^{(\text{slice})} = \sum_{ij} t_n^{ij} c^\dagger_{ni} c_{nj} + \sum_{ijkl} V_n^{ijkl} c^\dagger_{ni} c^\dagger_{nj} c_{nk} c_{nl}$$

Full complexity of orbital basis, but confined to each slice

$(N_{\text{orb}})^4$ terms but N_{orb} is small



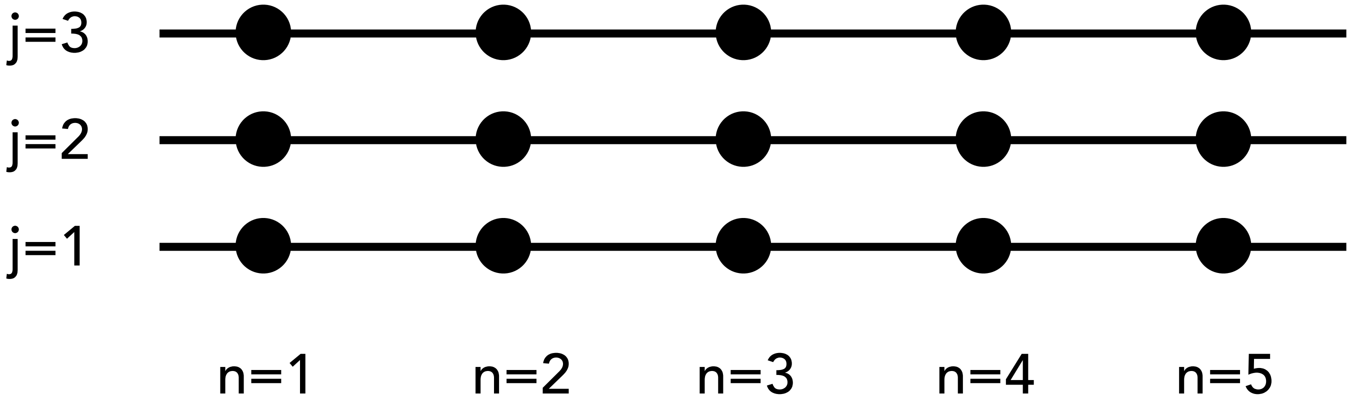
Form of Hamiltonian

$$H = \sum_n H_n^{(\mathrm{slice})} + T^{(\mathrm{grid})} + V$$

$$T^{(\mathrm{grid})} = -\frac{1}{2a^2} \sum_{nj} (c^\dagger_{nj} c_{n+1,j} - 2 n_{nj} + c^\dagger_{n+1,j} c_{nj})$$

Grid kinetic energy: hop without changing orbital j

Can use higher-order discrete deriv. to reduce error in a



Form of Hamiltonian

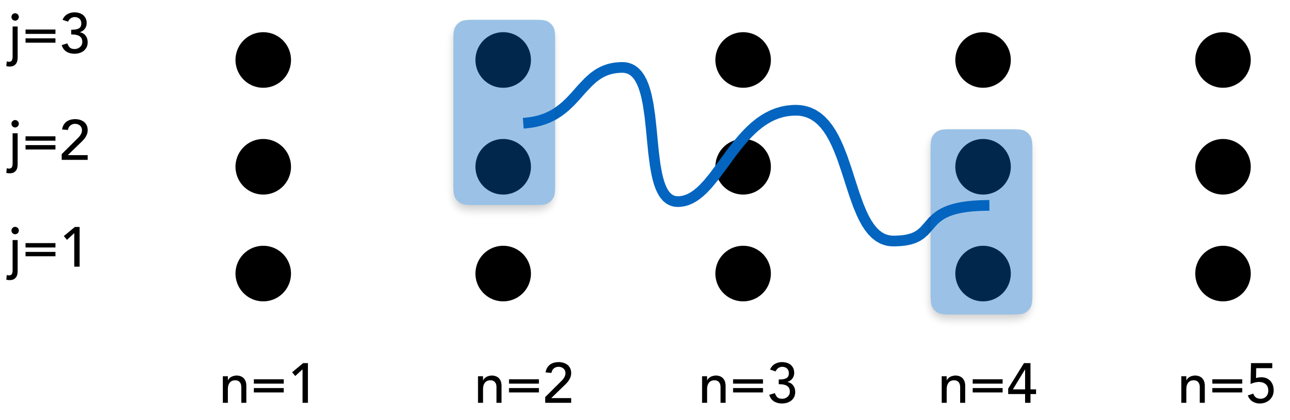
$$H = \sum_n H_n^{(\mathrm{slice})} + T^{(\mathrm{grid})} + V$$

$$V = \sum_{nn'} \sum_{ijkl} V_{mjk}^{nil} (c^\dagger_{ni} c_{nl})(c^\dagger_{mj} c_{mk})$$

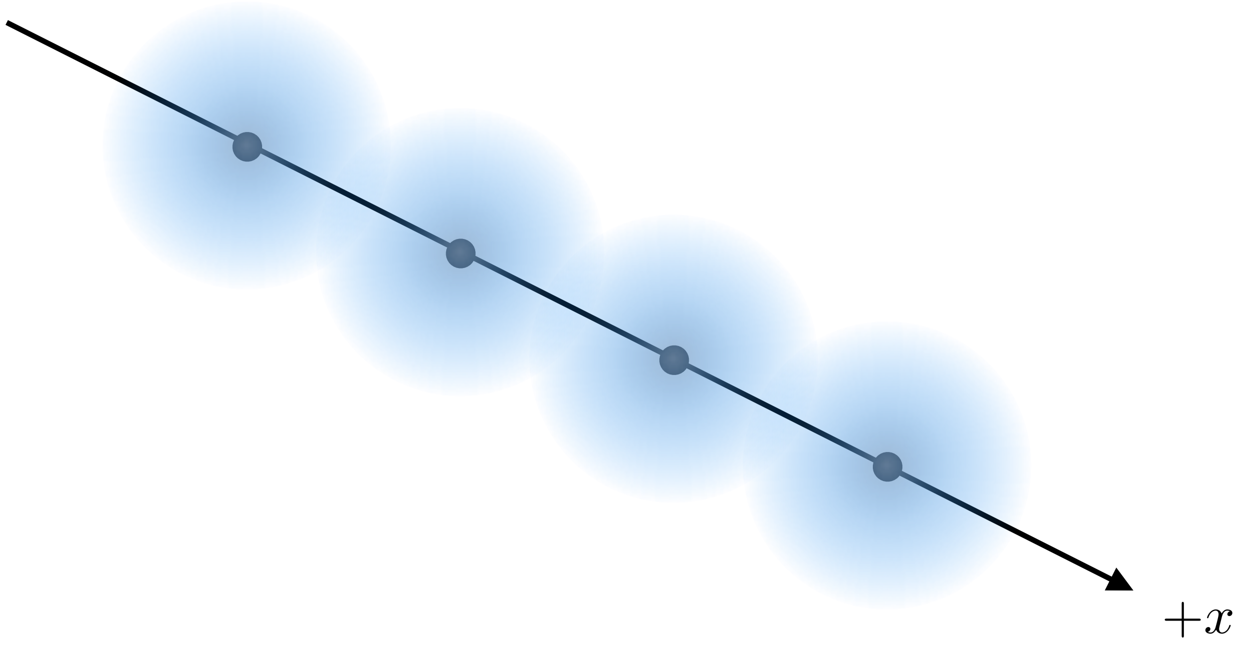
Operators $c^\dagger$ and c paired within rungs

Consists of $N^2 \, N^4_{\mathrm{orb}}$ terms ($<< N^4$)

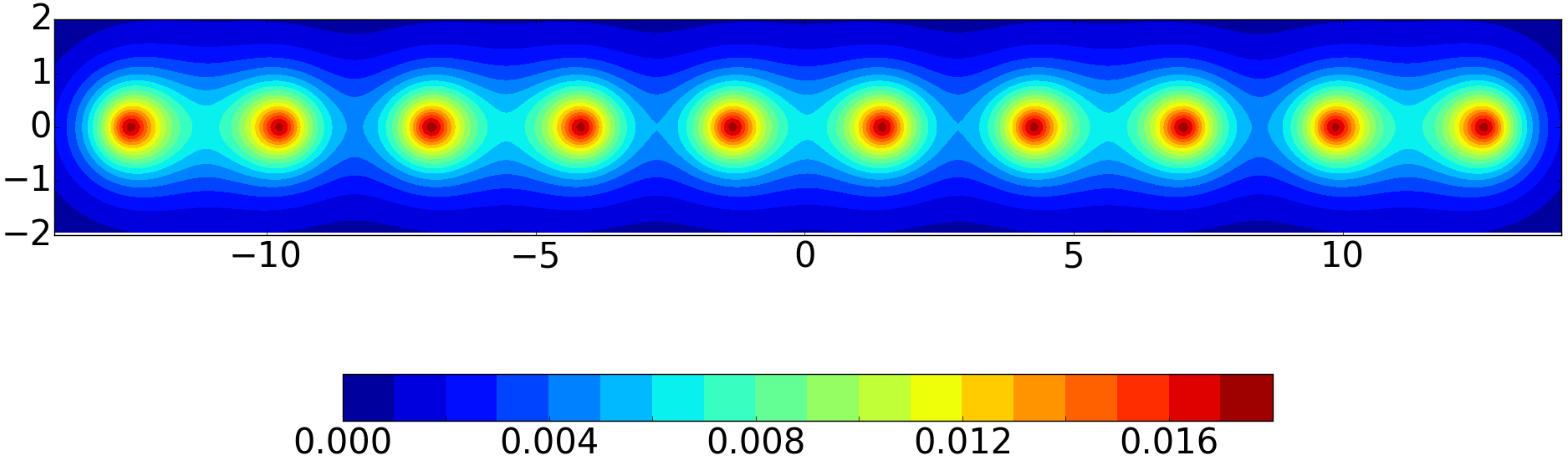
Can get even better scaling with compression (matrix product operator) tricks



Results for Hydrogen Chains

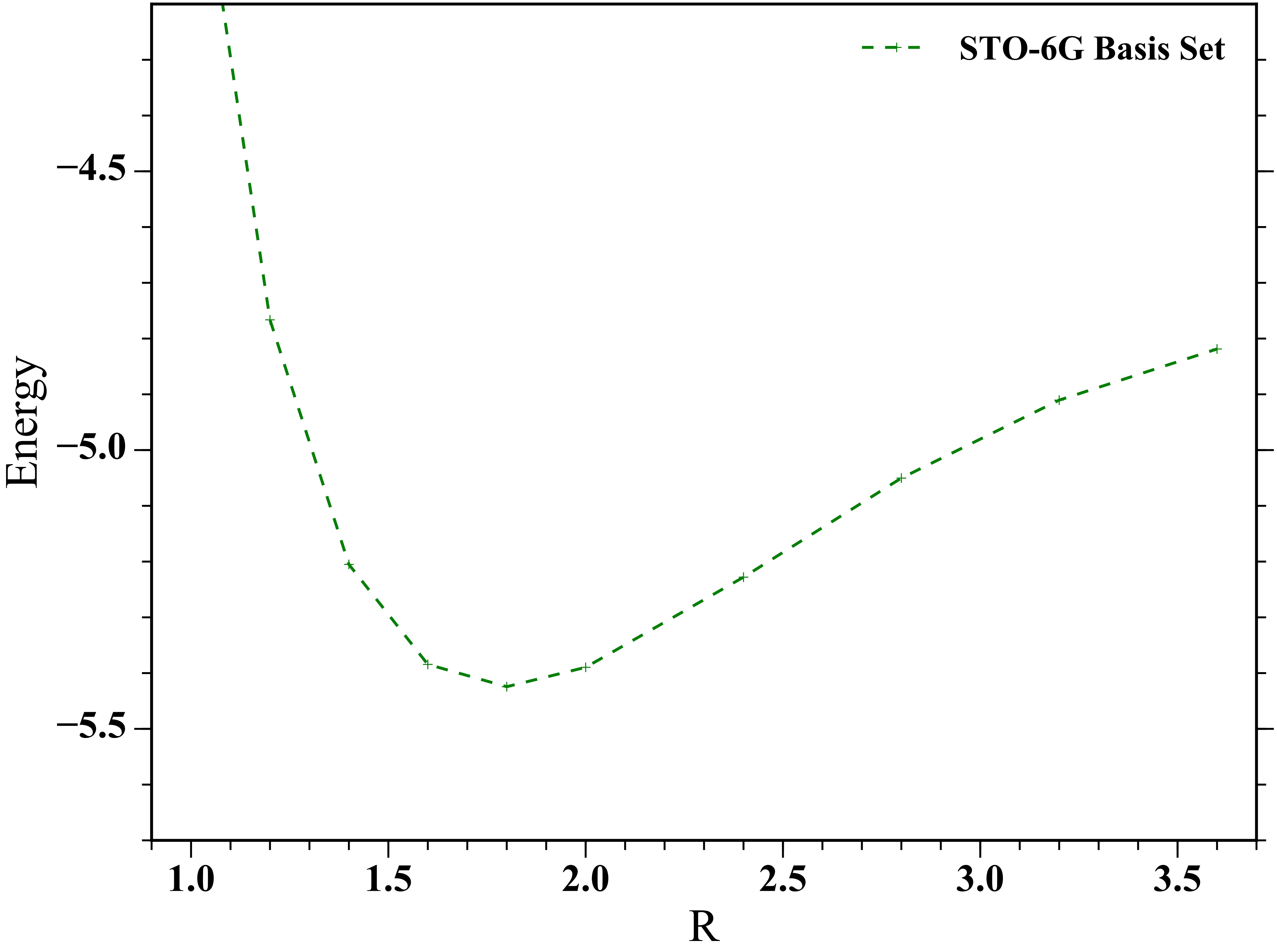


Density plot – chain of 10 hydrogen atoms

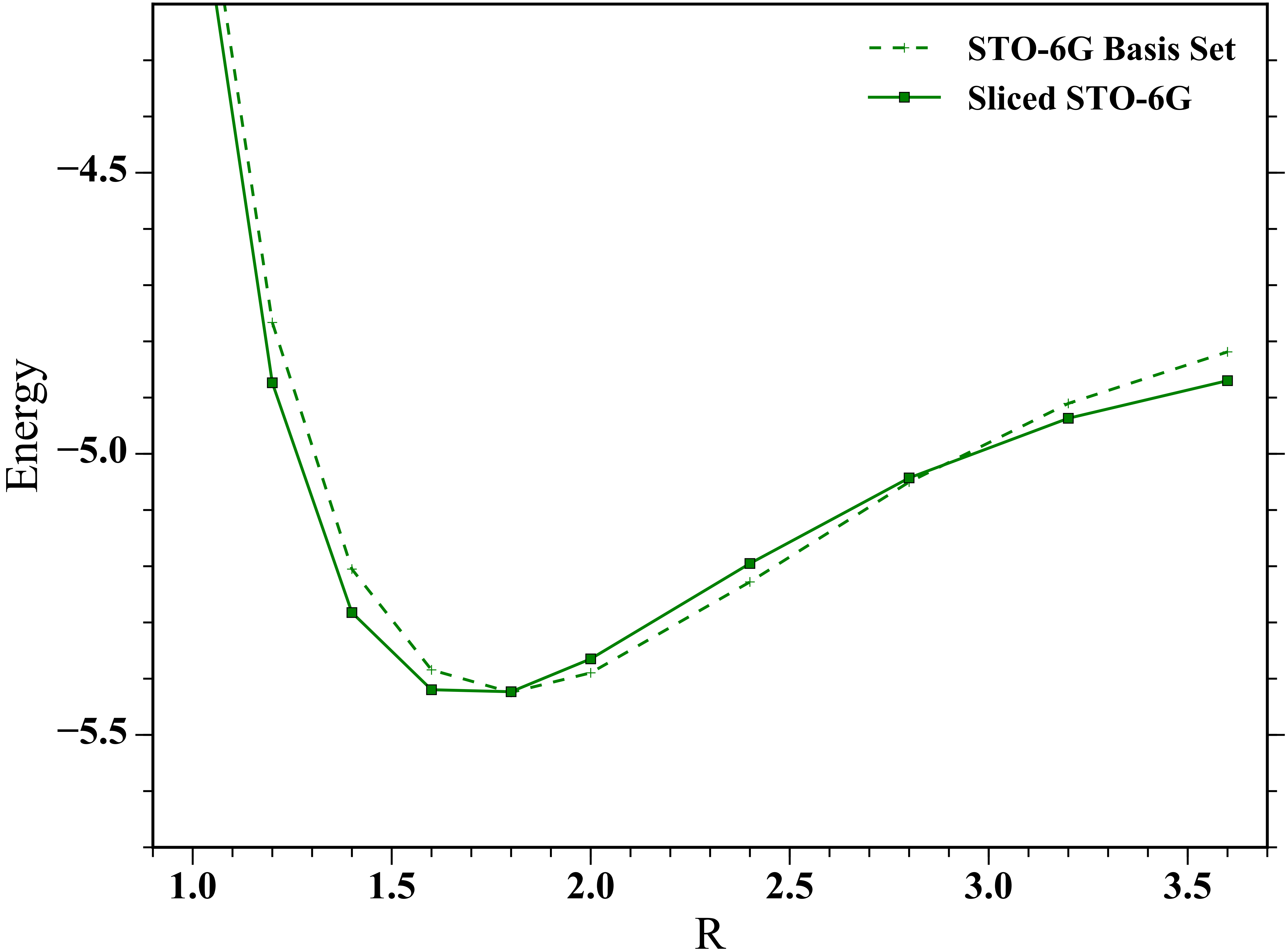


Spacing R=2.8; cc-pvDZ derived basis set

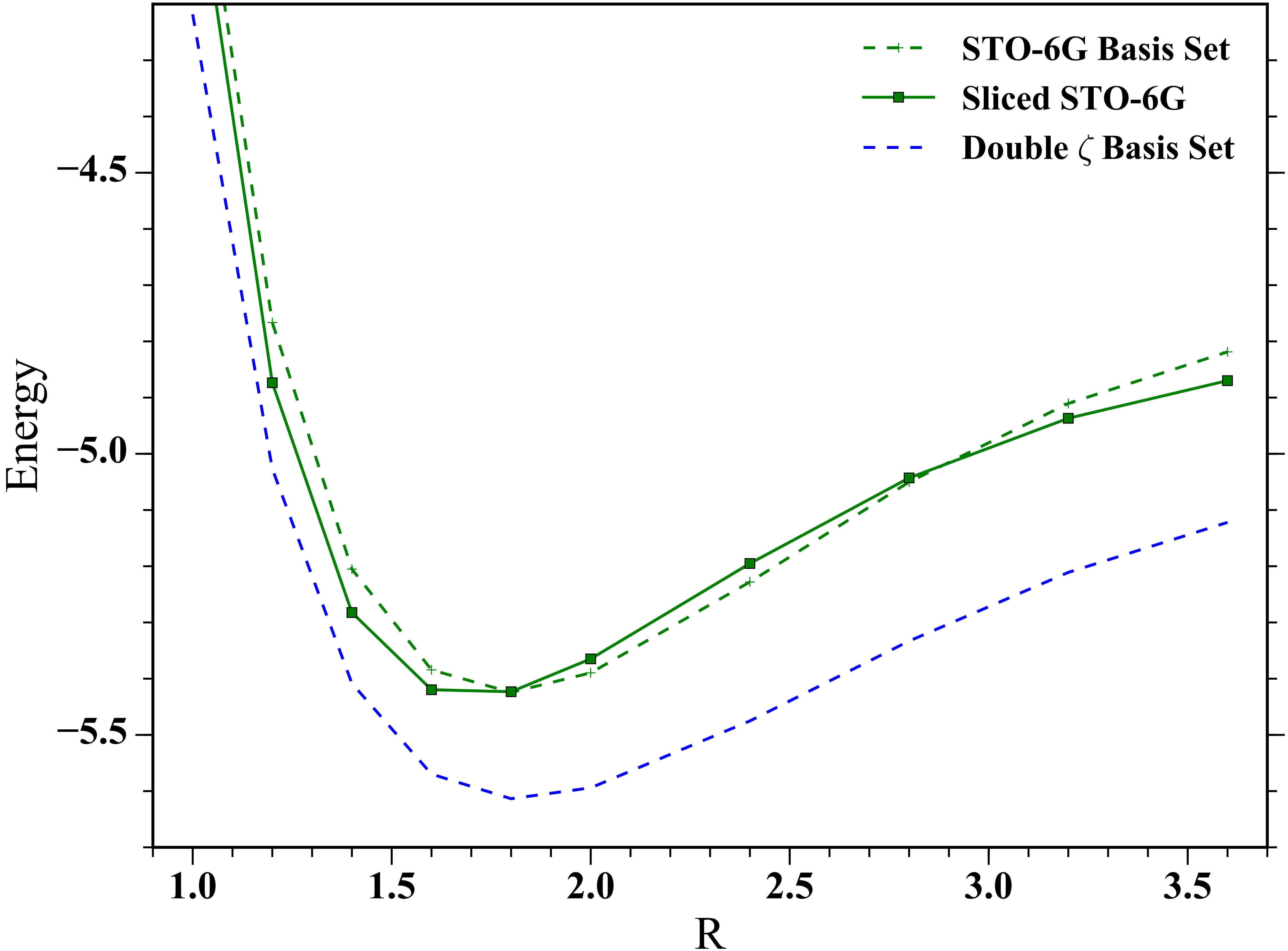
Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



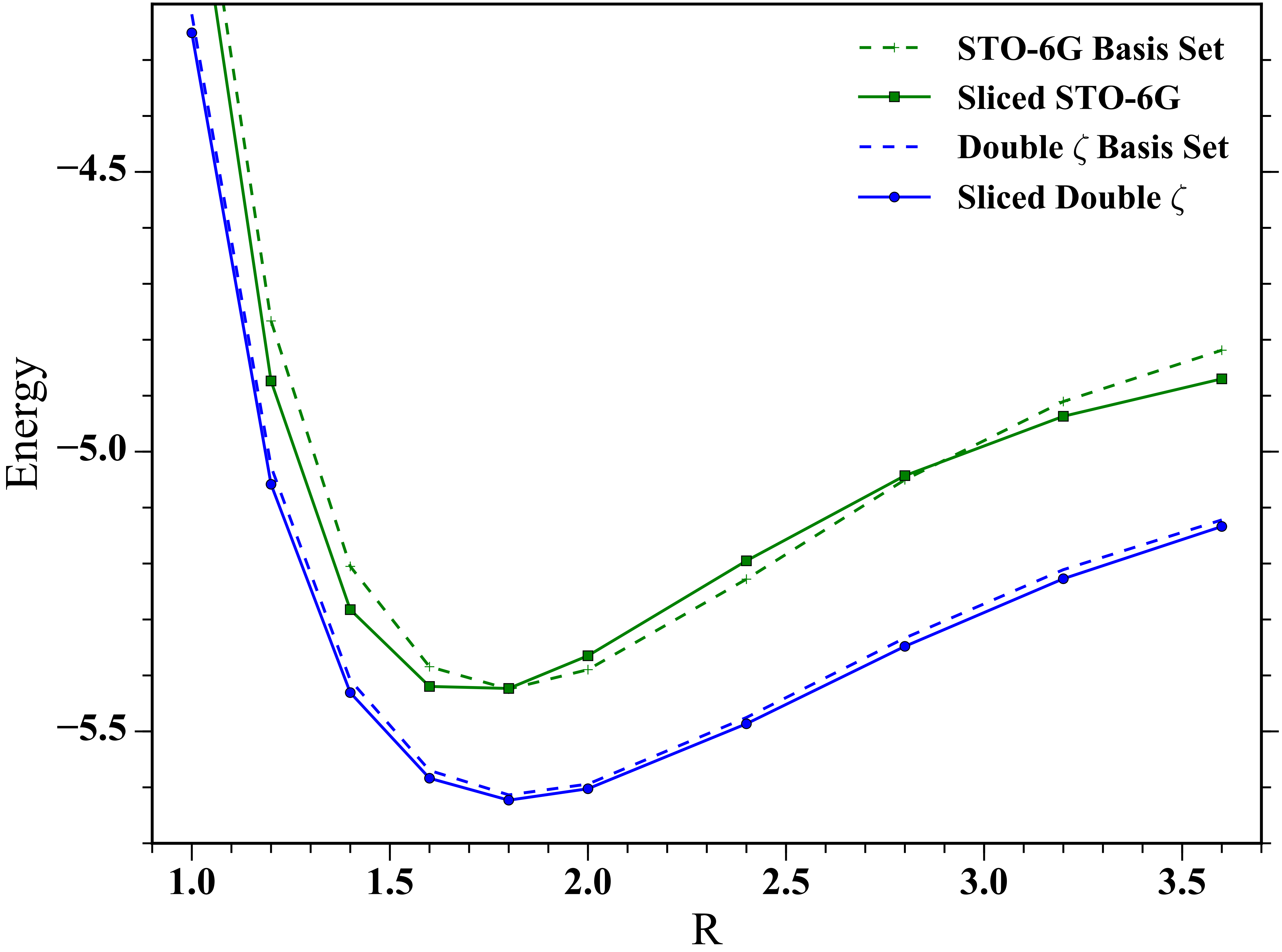
Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



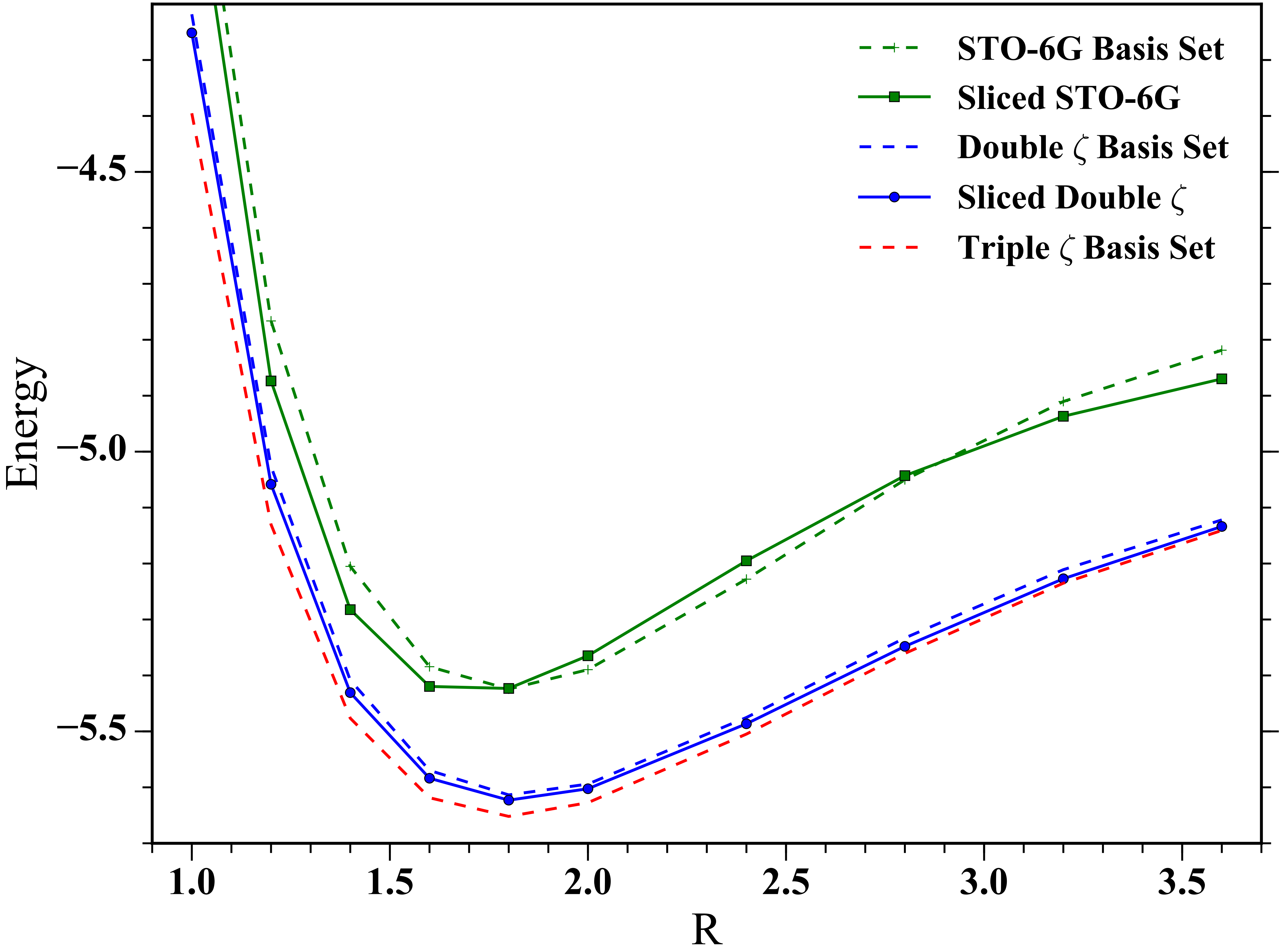
Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



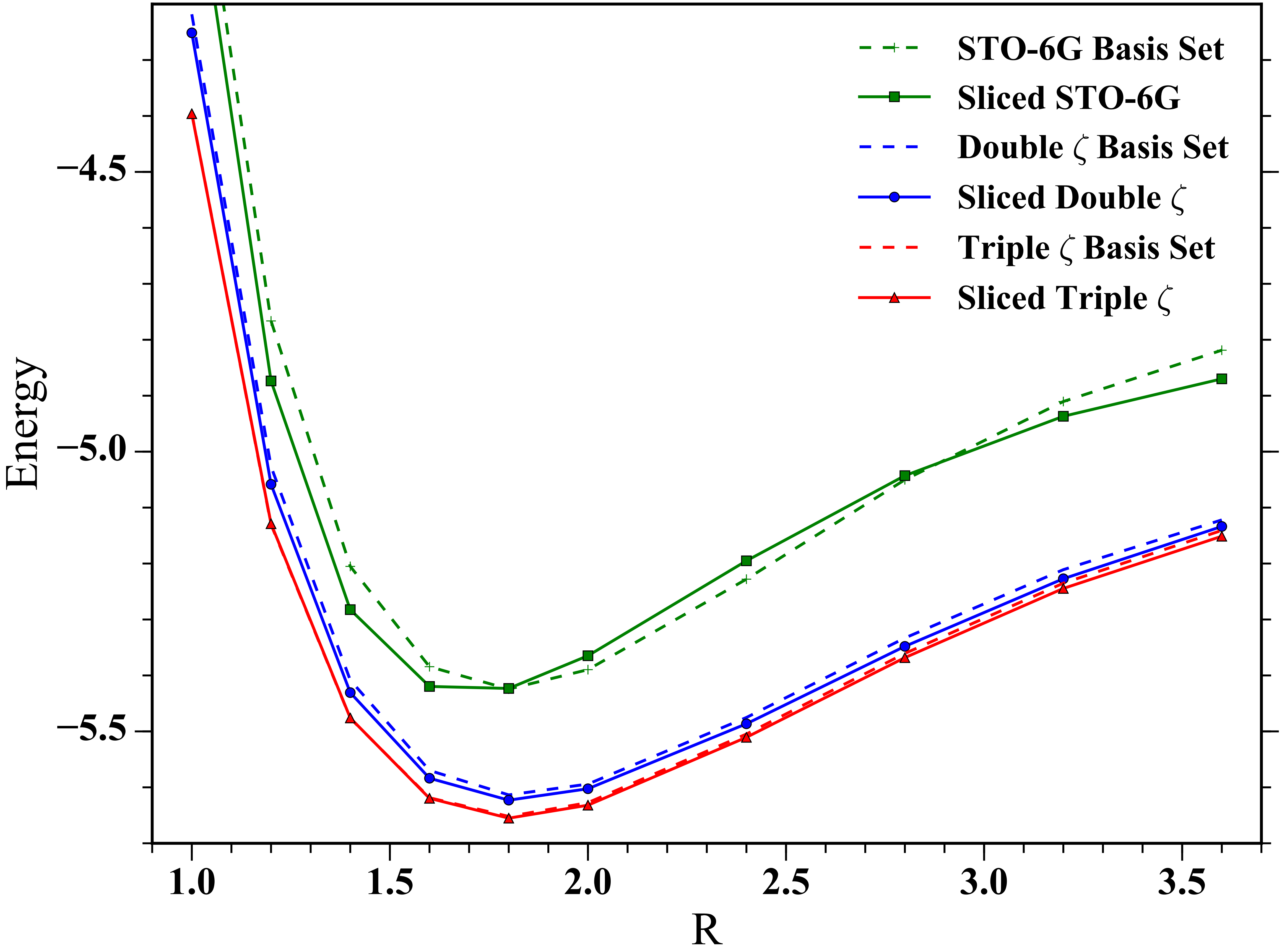
Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



Chains of N_a = 10 hydrogen atoms, spacing R (fixed nuclei)



Advantages of Slicing Approach

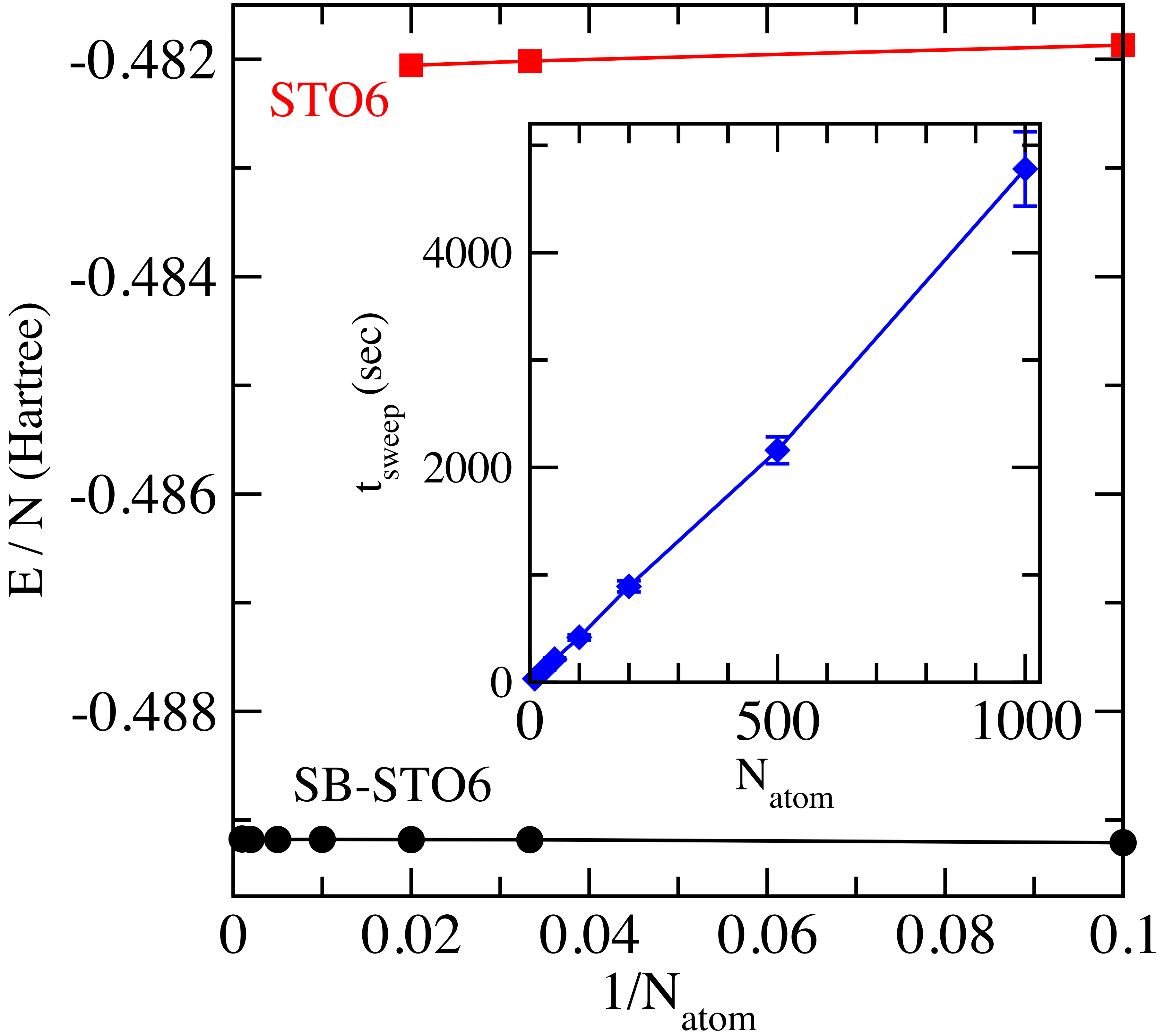
Other main advantage of slicing approach is simpler Hamiltonian.

$$H = \sum_j H_j^{(\text{slice})} + T^{(\text{grid})} \quad + V$$

- $\sum_j H_j^{(\text{slice})} + T^{(\text{grid})}$ parts are very short range, easy for DMRG
- V contains $N^2(N_{\text{orb}})^4$ terms, but can compress into a matrix product operator (MPO) of bond dimension only about $M_{\text{MPO}} \sim 100$, weakly dependent on system size

$$N^2(N_{\text{orb}})^4 \;\rightarrow\; (M_{\text{MPO}})^2$$

Linear scaling to 1000 hydrogen atoms



Advantages of Slicing Approach

Other advantanges to explore

- can optimize transverse functions within DMRG
- can use tensor networks (tree network; MERA) to initialize and optimize
- easier for applying time-dependent DMRG in chemistry context?

Conclusions & Future directions

- Hybrid grid + basis set ("sliced" basis) approach can extend usefulness of DMRG for chemistry
- Better scaling and treatment of correlations
- Approach is ideal for 1D trapped cold atoms too
- New MPO technology can be used for real-time evolution / dynamics studies

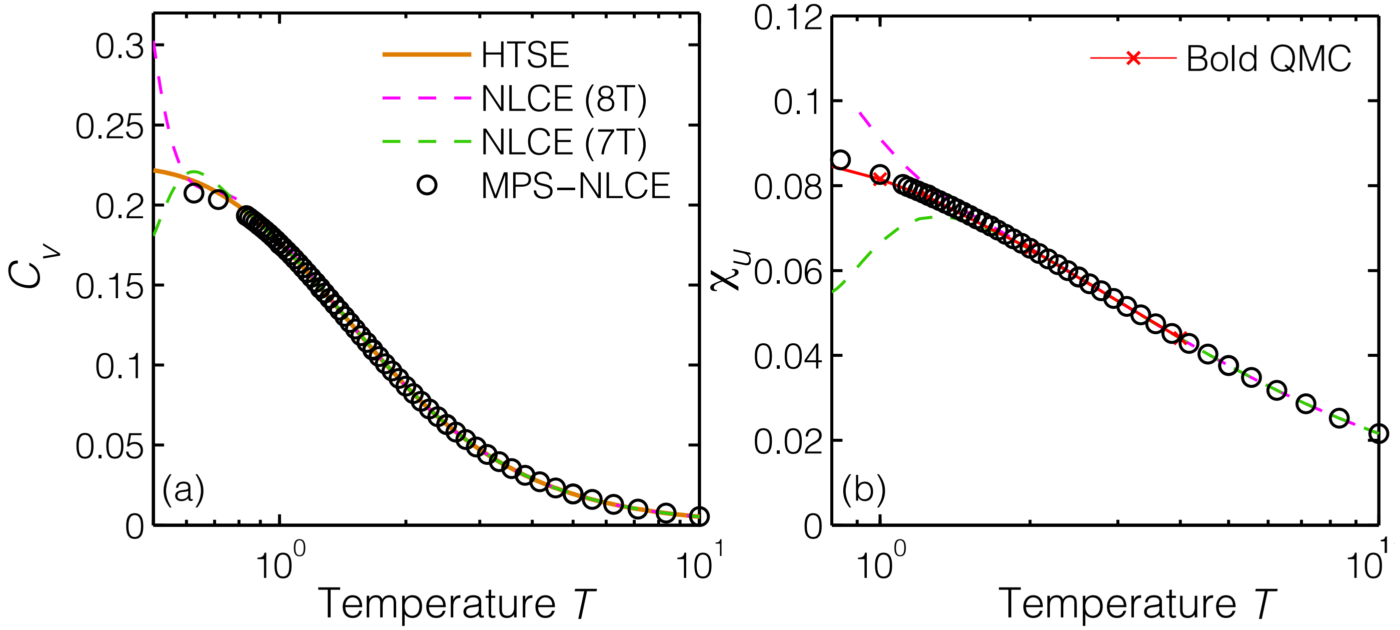
Finite Temperature Systems with MPS

Finite Temperature Systems with Matrix Product States

Two different, complementary techniques:

- purification / ancilla
- minimally entangled typical thermal states (METTS)

Highlight of applications



Finite Temperature Quantum Mech

Given Hamiltonian, for example $\hat{H} = \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j$

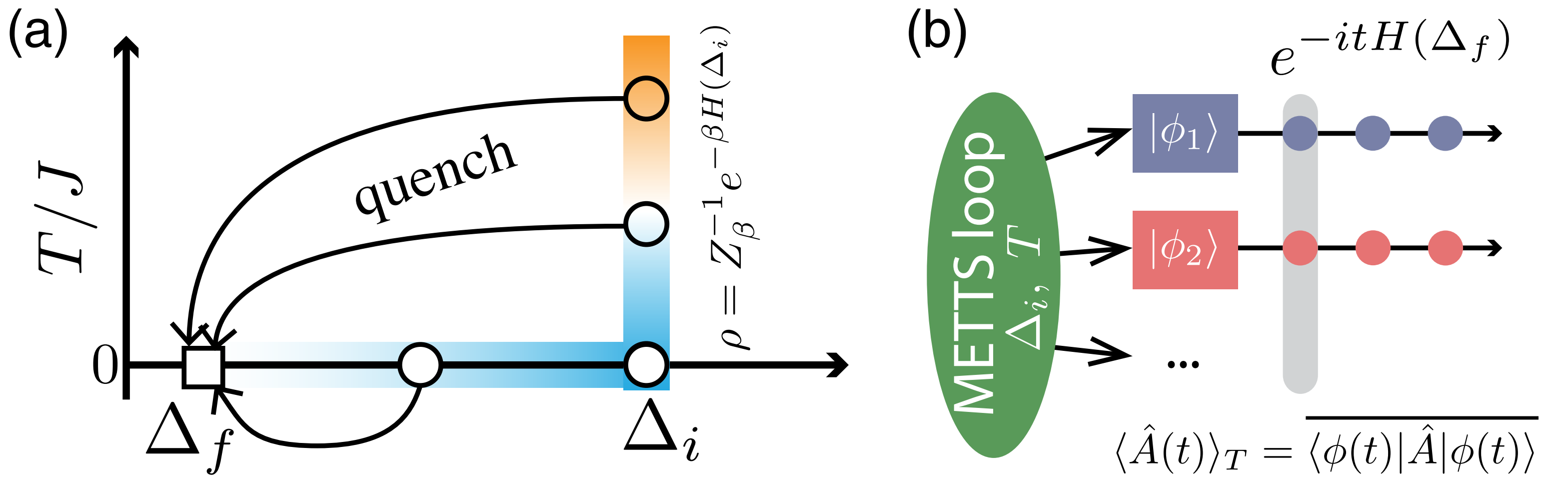
Compute finite T observables

$$\hat{\rho} = \frac{1}{Z} e^{-\hat{H}/T} \qquad \longrightarrow \qquad \langle \hat{A} \rangle = \mathrm{Tr}[\hat{\rho} \hat{A}]$$

Finite Temperature Quantum

Lots of interesting physics at finite T (especially in 2 or 3 dimensions)

- Critical behavior
- Magnetization plateuax
- Pseudogap crossover in high Tc materials
- Quenches & dynamics from finite T systems (Bonnes, Essler, Lauchli)



Finite Temperature Quantum

Usual prescription: "just" obtain all eigenstates

$$\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle$$

Then finite T density matrix is

$$\hat{\rho} = \frac{1}{Z}\sum_n e^{-\epsilon_n/T}|\epsilon_n\rangle\langle\epsilon_n|$$

Finite Temperature Quantum

Usual prescription: "just" obtain all eigenstates

$$\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle$$

Thermal averages given by

$$\langle\hat{A}\rangle = \frac{1}{Z}\sum_n e^{-\epsilon_n/T}\langle\epsilon_n|\hat{A}|\epsilon_n\rangle$$

Finite Temperature Quantum

But eigenstates terrible numerically!

$$\xcancel{\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle}$$

- Most are very highly entangled (volume law)

Finite Temperature Quantum

But eigenstates terrible numerically!

$$\xcancel{\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle}$$

- Most are very highly entangled (volume law)
- Exponentially small energy spacing

Finite Temperature Quantum

But eigenstates terrible numerically!

$$\xcancel{\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle}$$

- Most are very highly entangled (volume law)
- Exponentially small energy spacing
- Poor (worst) basis for sampling specific heat

Finite Temperature Quantum

But eigenstates terrible numerically!

$$\xcancel{\hat{H}|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle}$$

- Most are very highly entangled (volume law)
- Exponentially small energy spacing
- Poor (worst) basis for sampling specific heat
- Retain "quantum" behavior at high T

Finite T With Tensor Networks

Why tensor networks for finite T numerics?

Finite T With Tensor Networks

Why tensor networks for finite T numerics?

- No sign problem (can treat fermions, frustrated magnets)

Finite T With Tensor Networks

Why tensor networks for finite T numerics?

- No sign problem (can treat fermions, frustrated magnets)
- No sampling error in some regimes

Finite T With Tensor Networks

Why tensor networks for finite T numerics?

- No sign problem (can treat fermions, frustrated magnets)
- No sampling error in some regimes
- Starting point for dynamics

Finite T With Tensor Networks

Why tensor networks for finite T numerics?

- No sign problem (can treat fermions, frustrated magnets)
- No sampling error in some regimes
- Starting point for dynamics
- Interesting perspective on finite T physics

Finite T With Tensor Networks

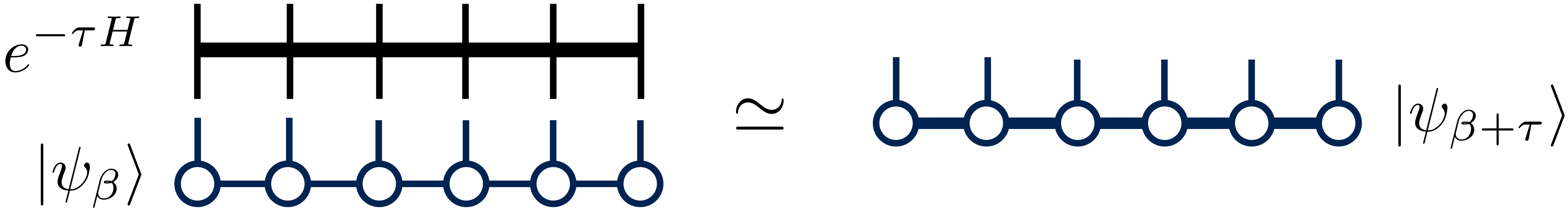
Two major approaches for finite T:

1. Ancilla or purification
2. METTS sampling (minimally entangled typical thermal states)

Ancilla / Purification Method

Surprisingly simple idea: directly make density matrix

Leverage algorithms for time evolving MPS

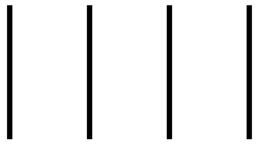


Ancilla / Purification Method

Surprisingly simple idea: directly make density matrix

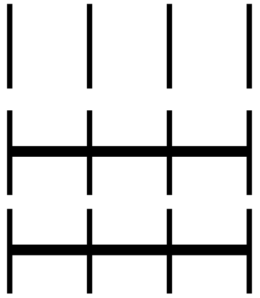
Overview:

1. start with identity operator



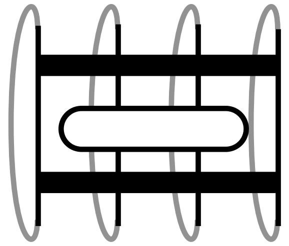
$$\hat{I}_1\hat{I}_2\hat{I}_3\hat{I}_4$$

2. evolve in imaginary time for time $\beta/2$



$$\left.\begin{matrix}e^{-\tau H}\\ e^{-\tau H}\end{matrix}\right\} \left(\frac{\beta}{2\tau}\right) \text{ times}$$

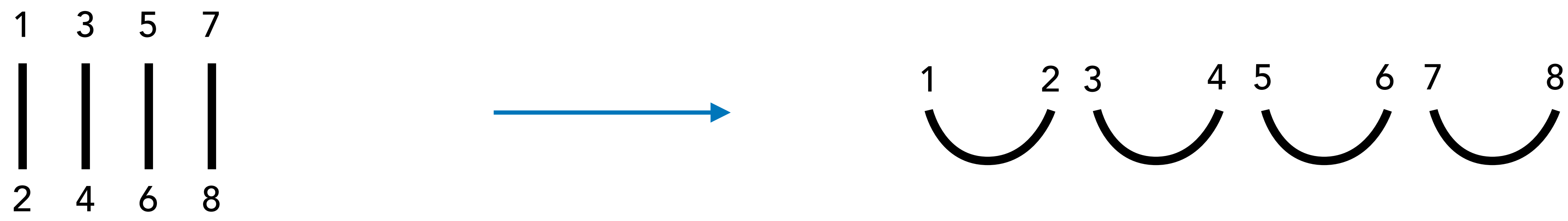
3. take trace expectation value of operator



$$\mathrm{Tr}[e^{-\beta/2H}\hat{A}\,e^{-\beta/2H}]/Z$$

Ancilla / Purification Method

View starting identity operator as a "wavefunction"



For case of spin 1/2 system, each pair mapped to wavefunction as:

$$|\uparrow\rangle_1\langle\uparrow|_2 + |\downarrow\rangle_1\langle\downarrow|_2 \longrightarrow |\uparrow\rangle_1|\uparrow\rangle_2 + |\downarrow\rangle_1|\downarrow\rangle_2$$

Easy to represent as a matrix product state (MPS)

Ancilla / Purification Method

View starting identity operator as a "wavefunction"

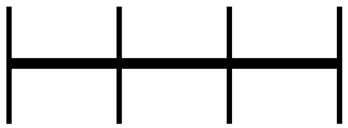
Picture is that even or "ancilla" sites thermalize odd "physical" sites

Tracing out ancillae gives infinite T density matrix for physical sites

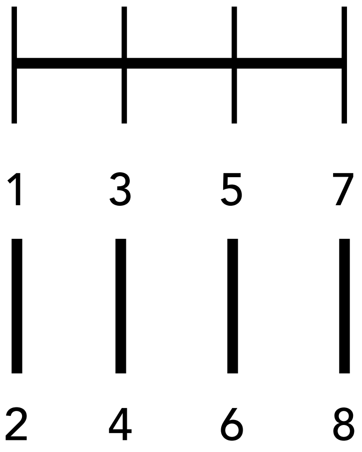
Ancilla / Purification Method

Imaginary time evolution:

Assume we have obtained $e^{-\tau H} =$



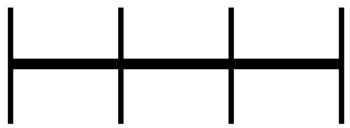
Want to apply it as



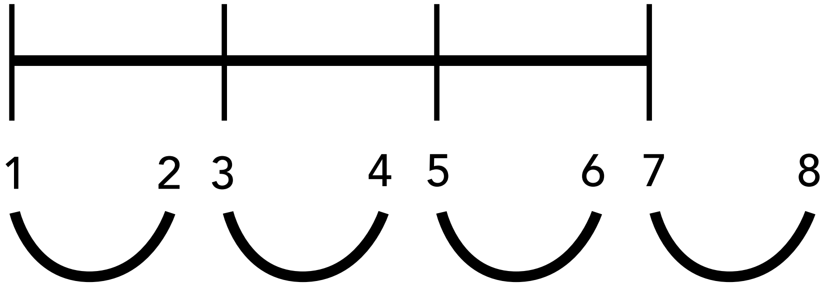
© Edwin Miles Stoudenmire, Flatiron Institute, 2018

Ancilla / Purification Method

Imaginary time evolution:

Assume we have obtained $e^{-\tau H} =$ 

Equivalent to

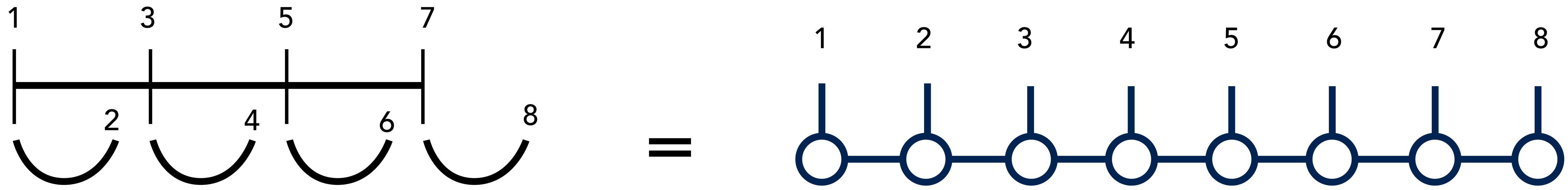


Apply to odd, or "physical" sites

Ancilla / Purification Method

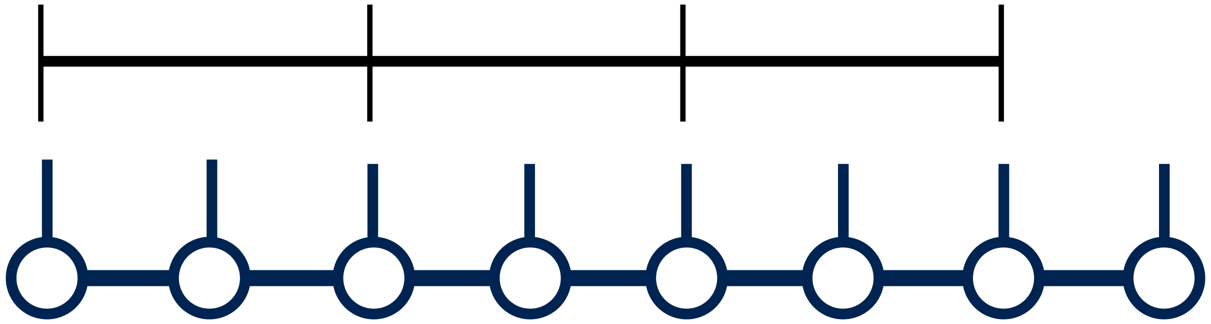
After first step, physical+ancilla "wavefunction" will be entangled

Represent as an MPS



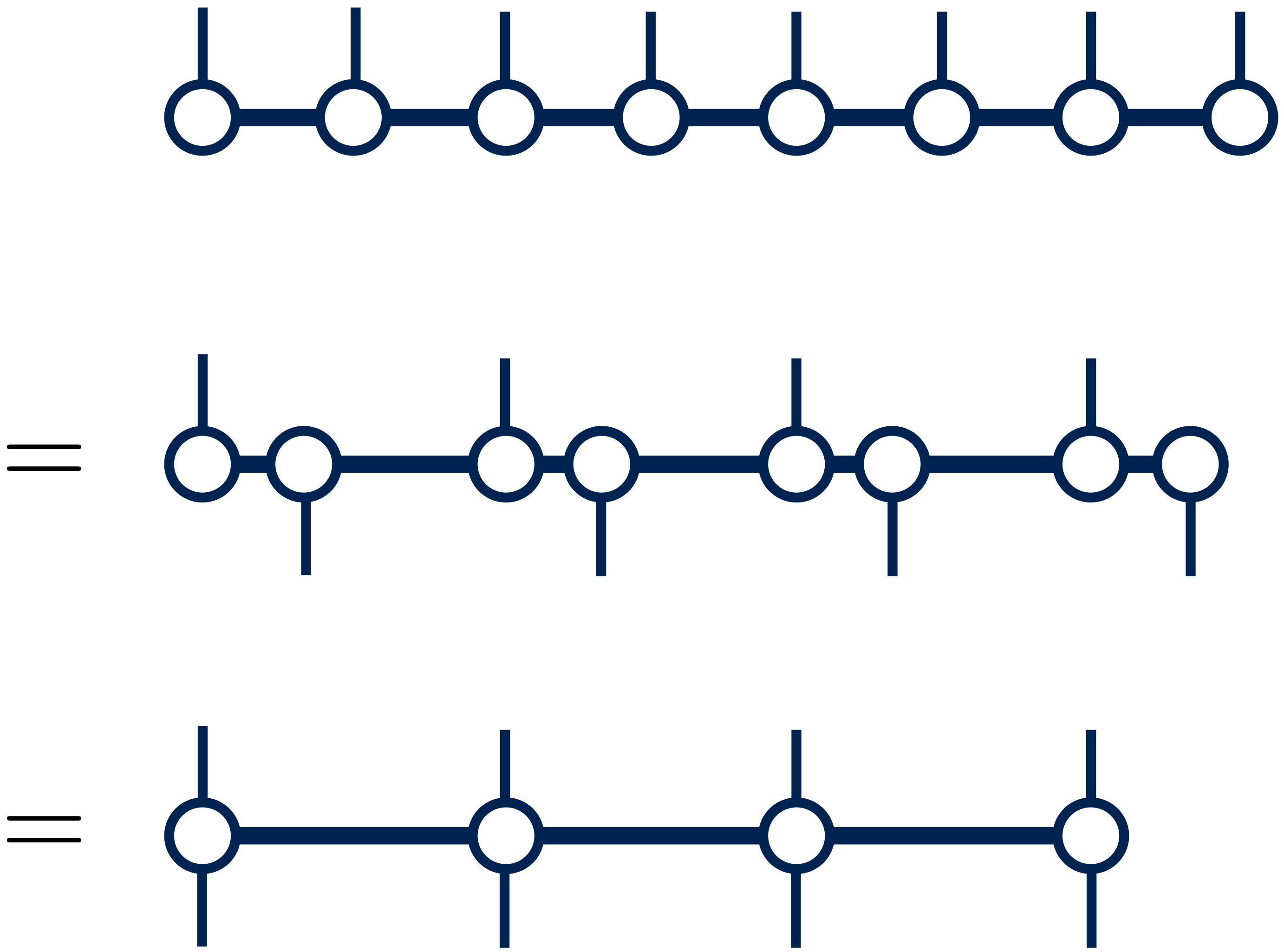
Ancilla / Purification Method

Continue time evolving physical sites (odd sites)



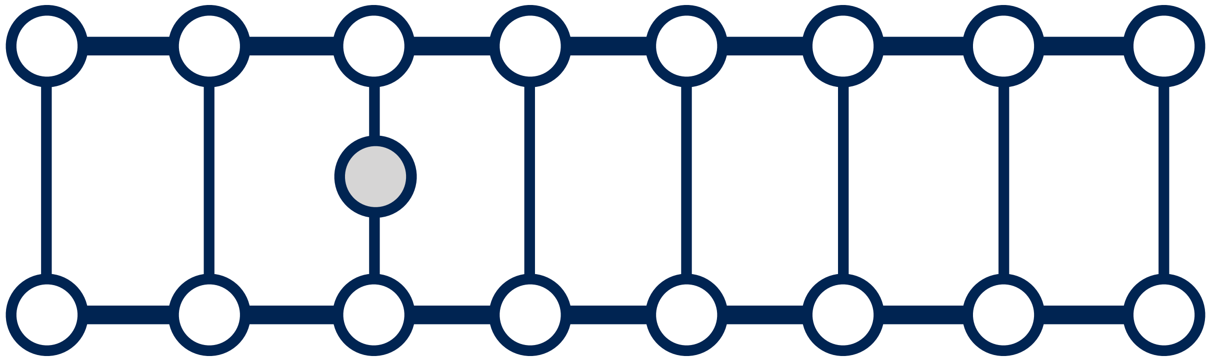
Ancilla / Purification Method

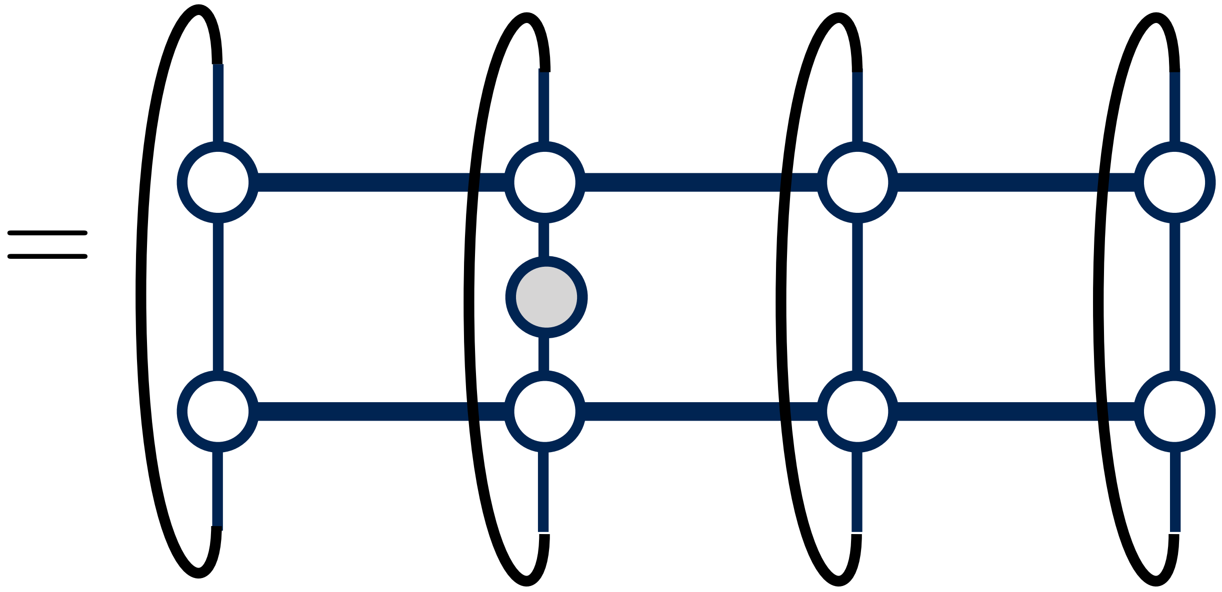
Final MPS proportional to $e^{-\beta/2H}$



Ancilla / Purification Method

Measure local operators as





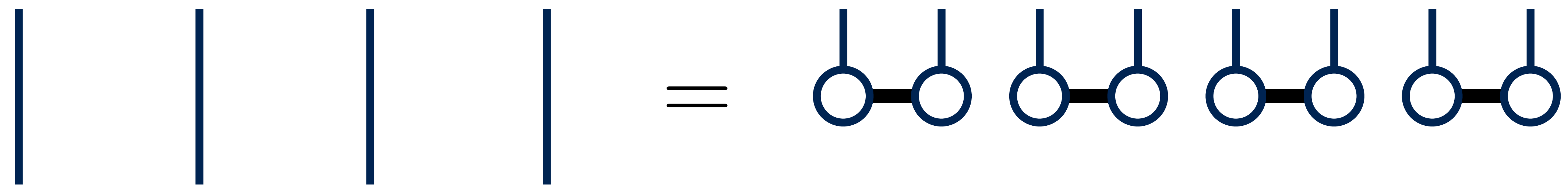
$$\mathrm{Tr}[e^{-\beta/2H}\,\hat{A}\,e^{-\beta/2H}]$$

(up to proper normalization)

Ancilla / Purification Method

How well does it work?

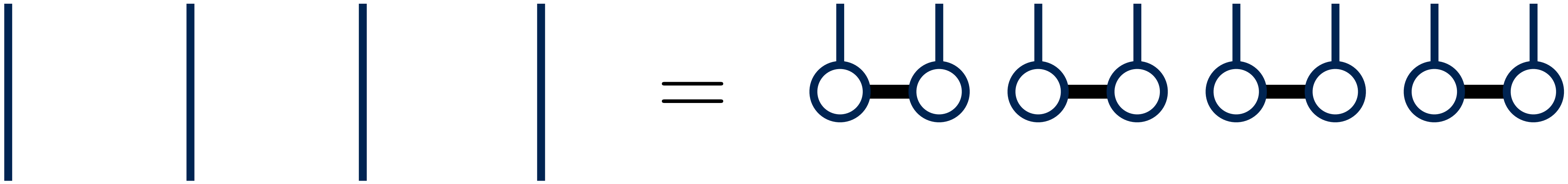
Infinite T:



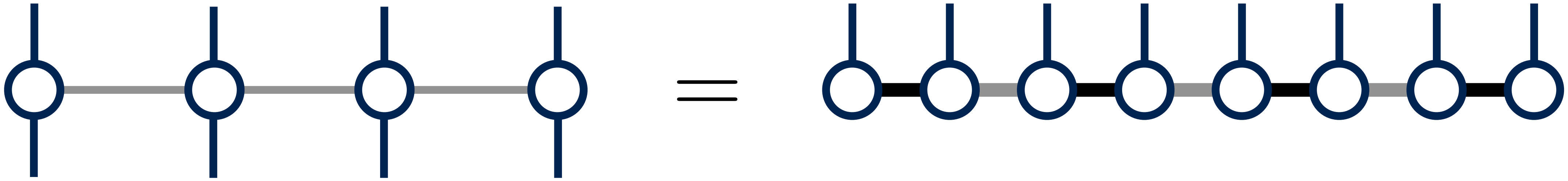
Ancilla / Purification Method

How well does it work?

Infinite T:



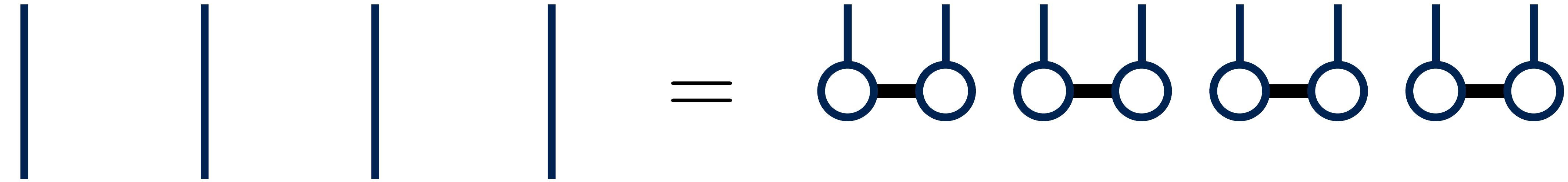
Medium T:



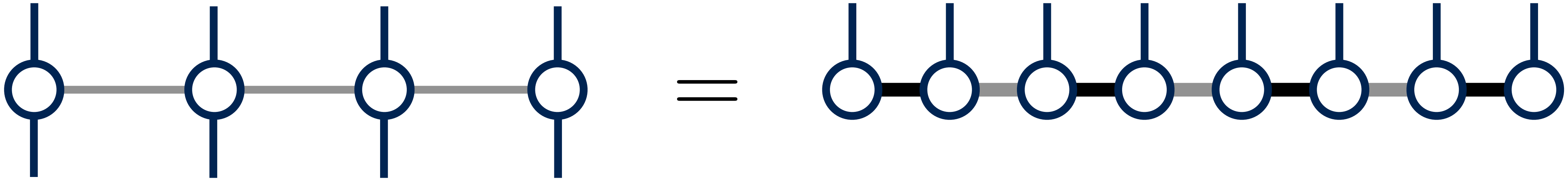
Ancilla / Purification Method

How well does it work?

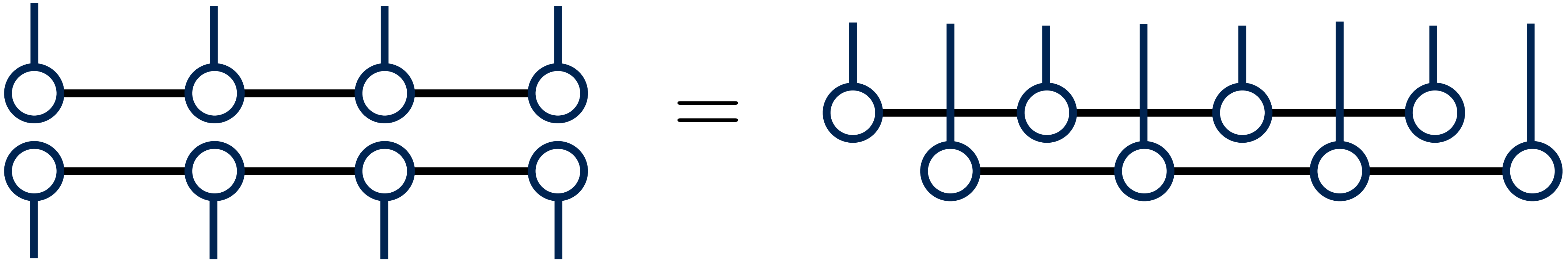
Infinite T:



Medium T:



T = 0:



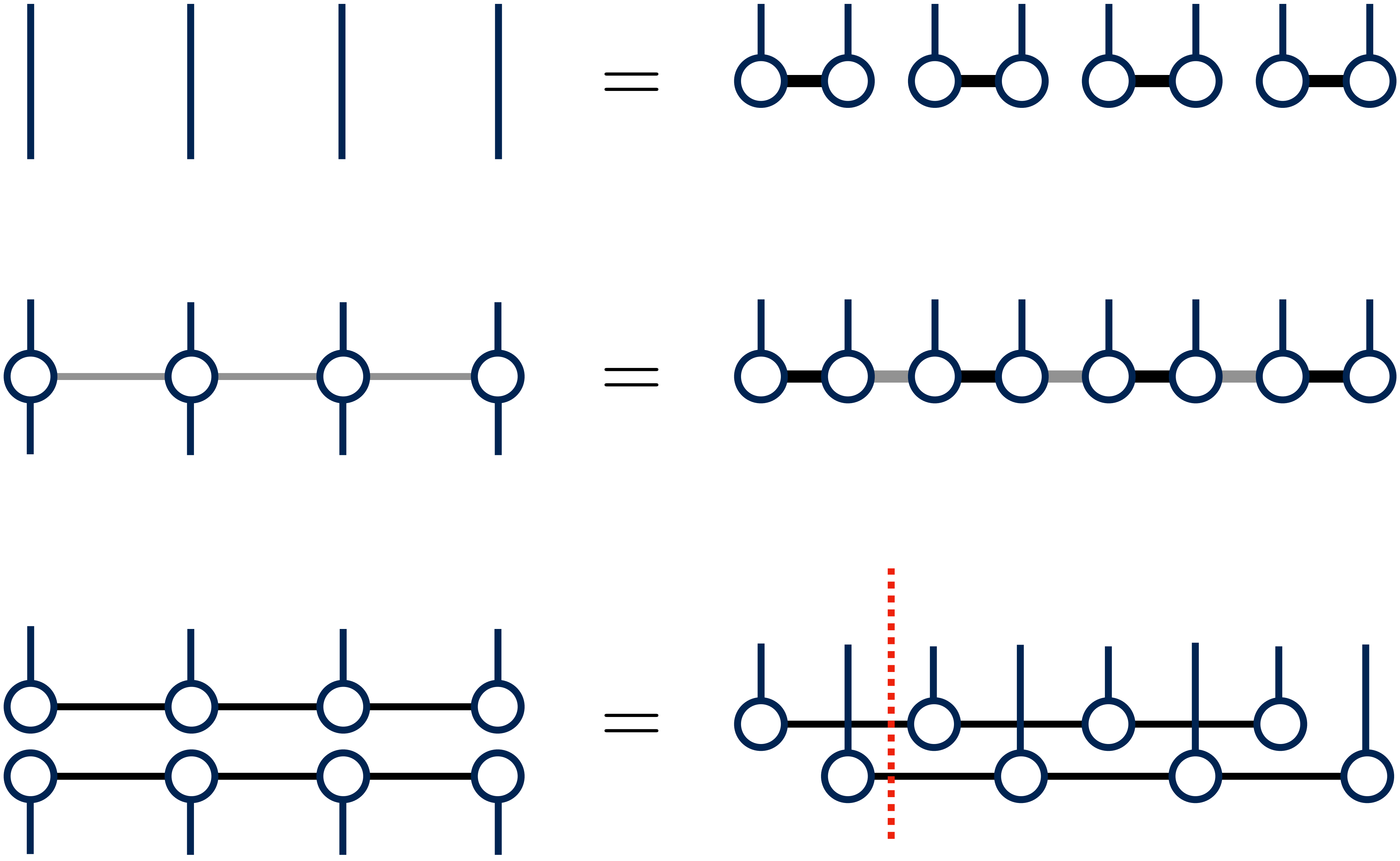
Ancilla / Purification Method

How well does it work?

Infinite T:

Medium T:

T = 0:



2x the entanglement!

Ancilla / Purification Method

How well does it work?

In practice, reach $T \approx J$

At lower T, similar to DMRG with bond dimension m^2

Scaling is therefore $(m^2)^3 = m^6$

Key advantage: can measure at all T's on the way down

Ancilla / Purification Method

1D Example: S=1/2 chain (Feiguin & White 2005)

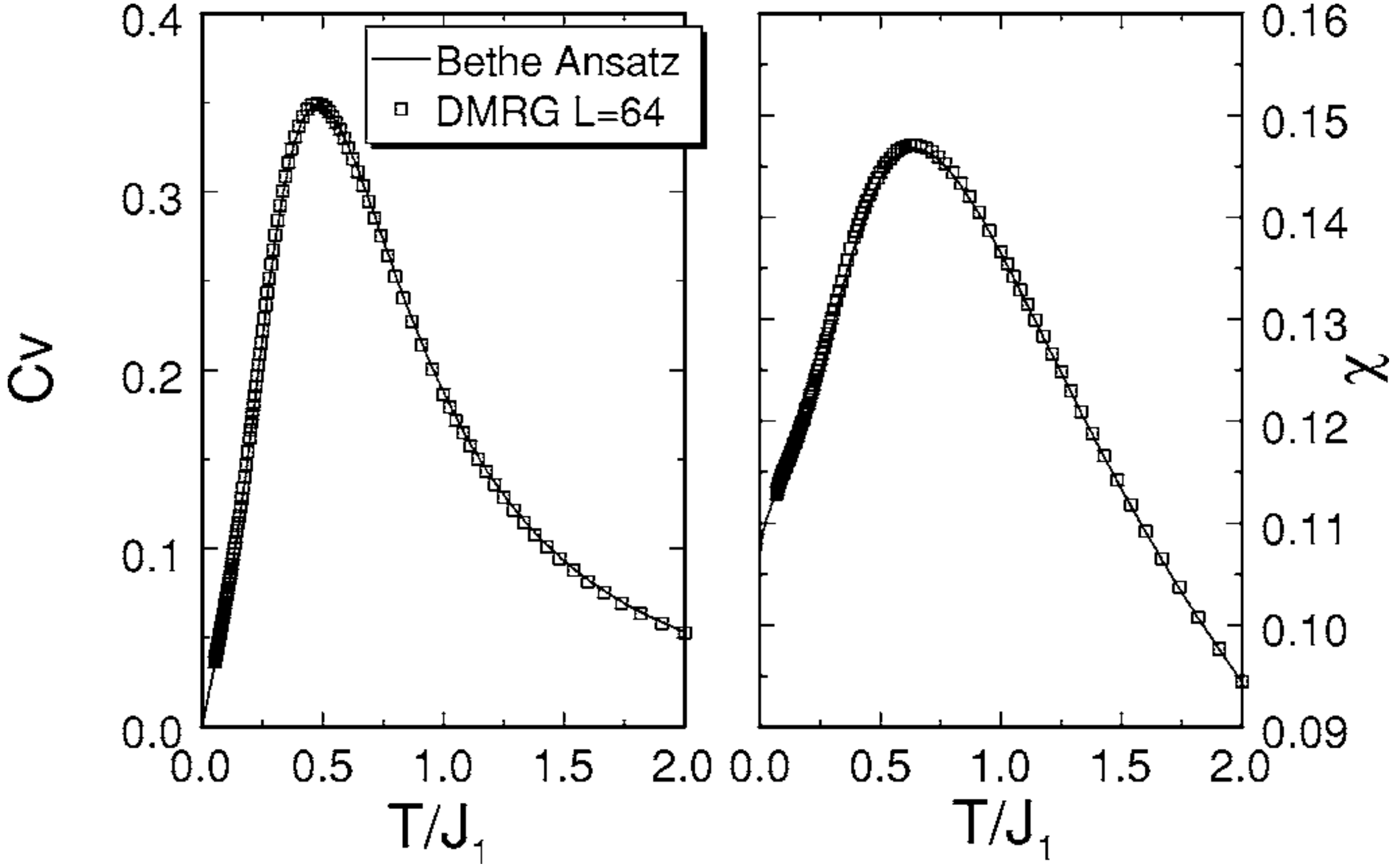
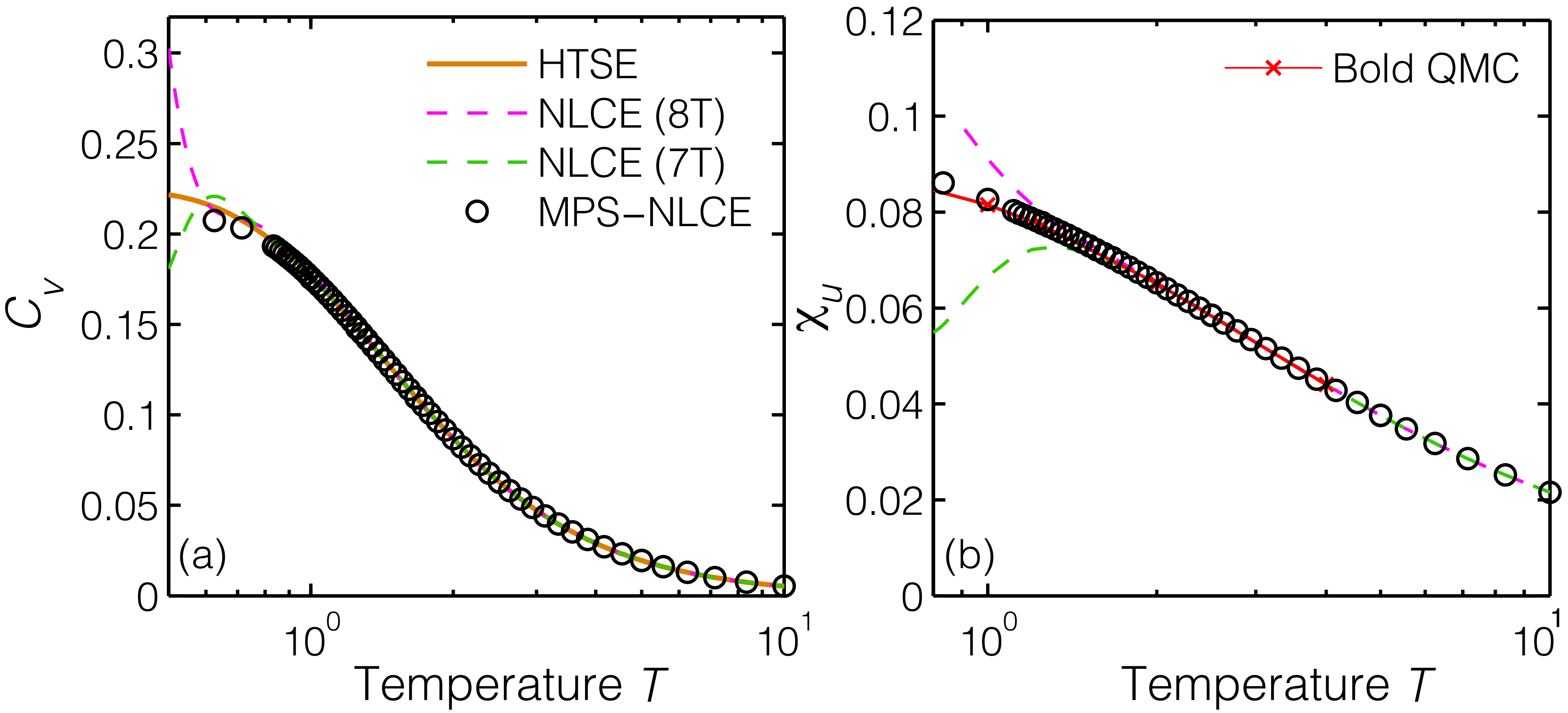


FIG. 2. Specific heat and magnetic susceptibility of a $S=1/2$ spin chain of length $L=64$, compared to exact $L=\infty$ results using the Bethe ansatz.

Feiguin, White, PRB 72, 220401(R) (2005)

Ancilla / Purification Method

2D Example: S=1/2 triangular lattice Heisenberg



Bruognolo, Zhu, White, Stoudenmire, arxiv:1705.05578

Minimally Entangled Typical Thermal States

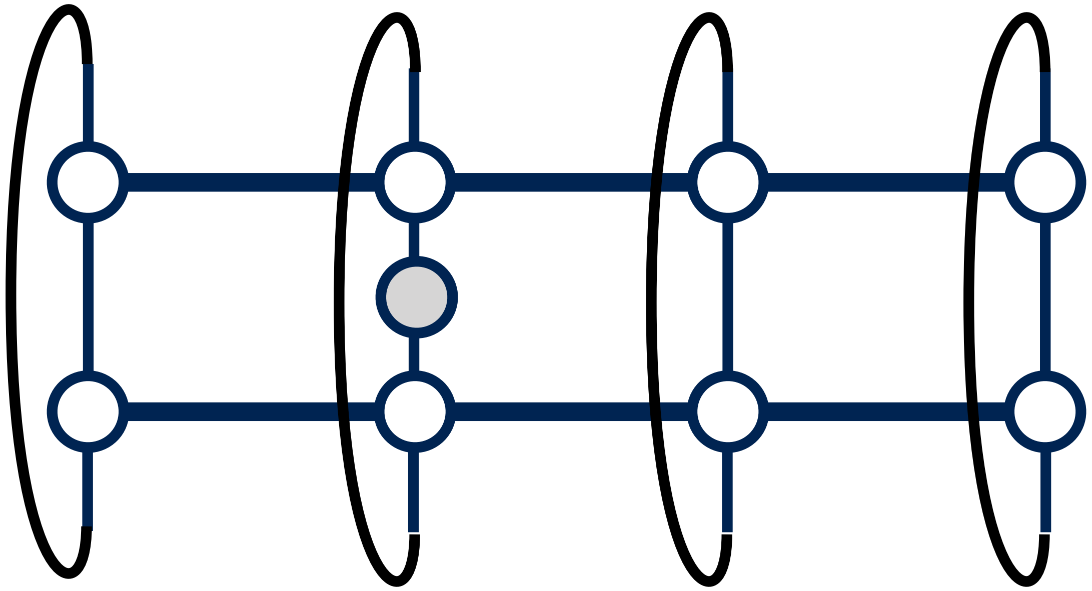
Purification has key drawback:

costly at low temperatures (even if polynomial)

➡ Minimally entangled typical thermal states (METTS) designed to get around this problem (White, 2009)

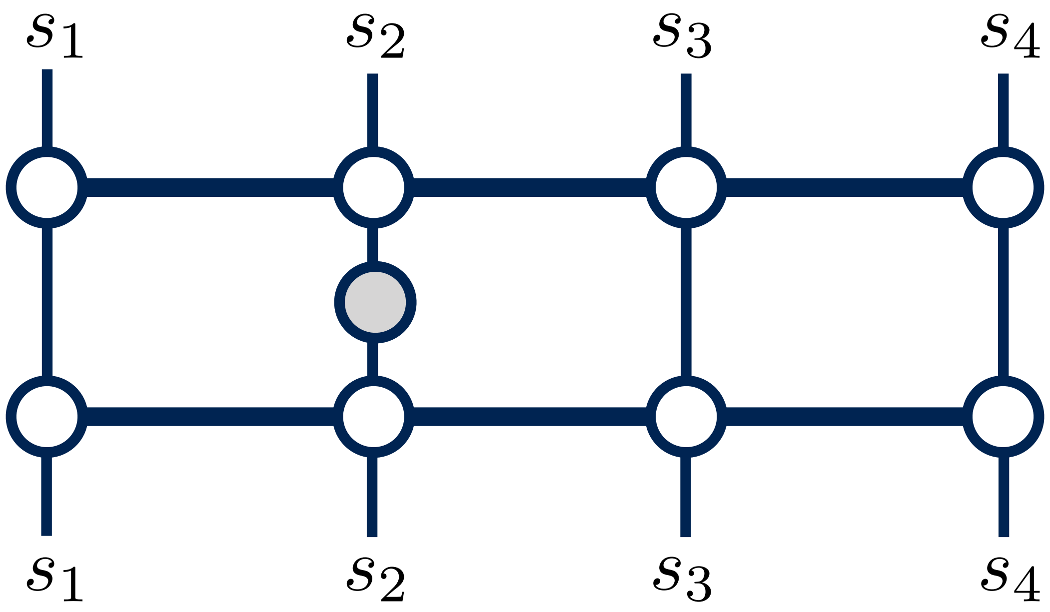
Minimally Entangled Typical Thermal States

Think of replacing trace by explicit summation



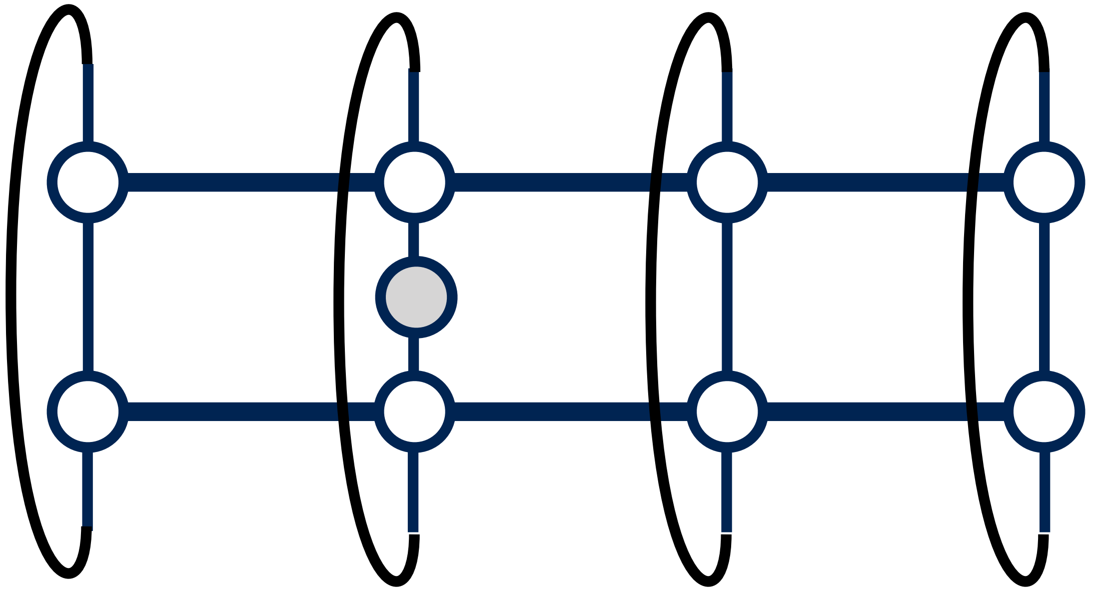
$$\frac{1}{Z}\mathrm{Tr}[e^{-\beta H/2}\hat{A}e^{-\beta H/2}]$$

$$=\frac{1}{Z}\sum_{s_1 s_2 s_3 s_4}$$



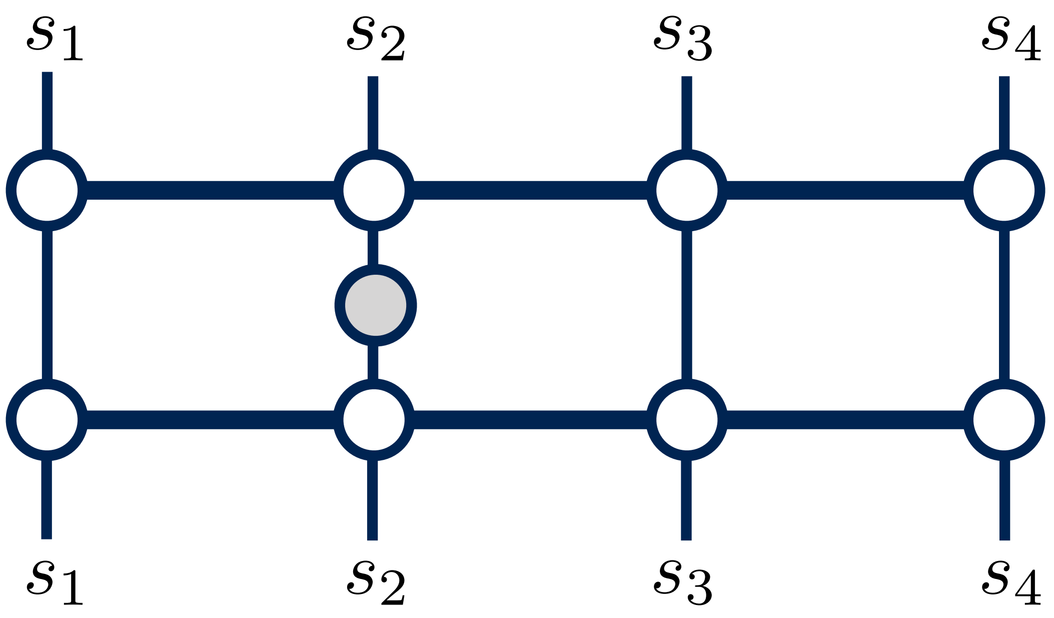
Minimally Entangled Typical Thermal States

Think of replacing trace by explicit summation



$$\frac{1}{Z}\mathrm{Tr}[e^{-\beta H/2}\hat{A}e^{-\beta H/2}]$$

$$=\frac{1}{Z}\sum_{s_1s_2s_3s_4}$$



$$\langle s_1s_2s_3s_4|$$

$$e^{-\beta H/2}$$

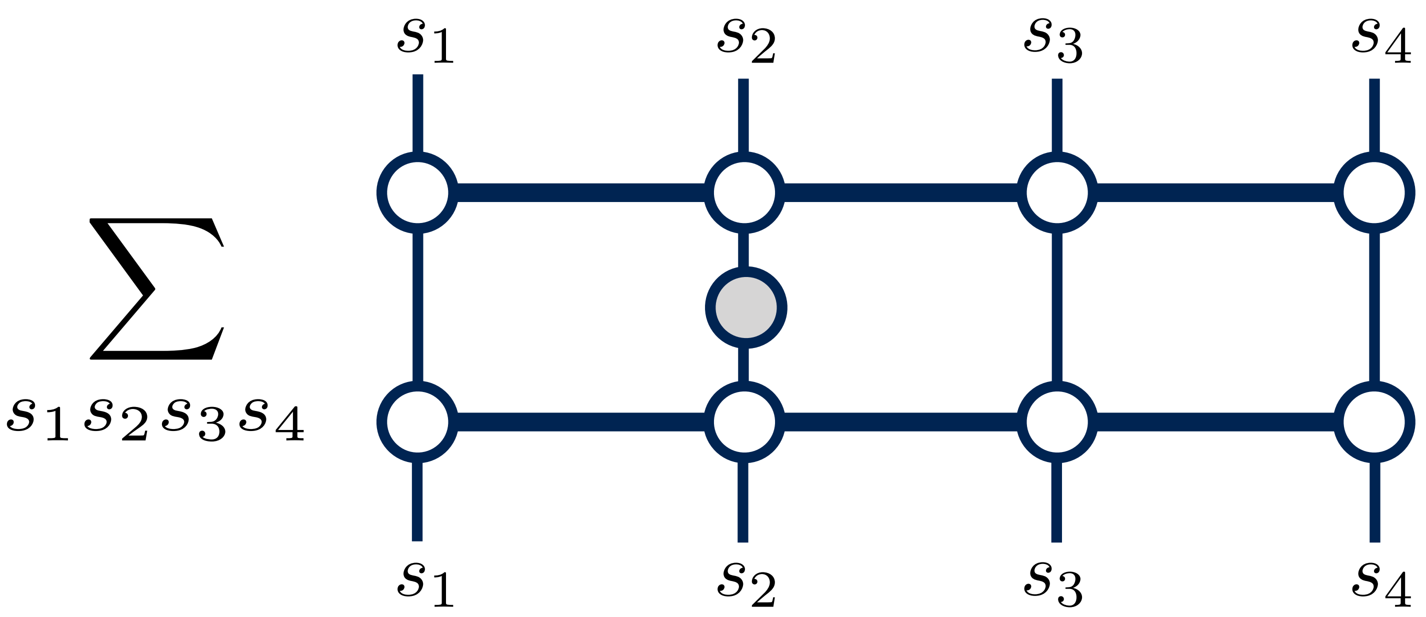
$$\hat{A}$$

$$e^{-\beta H/2}$$

$$|s_1s_2s_3s_4\rangle$$

Minimally Entangled Typical Thermal States

Rewrite in a suggestive way



$$= \frac{1}{Z} \sum_{\mathbf{s}} \langle \mathbf{s} | e^{-\beta H/2} \hat{A} e^{-\beta H/2} | \mathbf{s} \rangle$$

$$= \frac{1}{Z} \sum_{\mathbf{s}} P(\mathbf{s}) \langle \phi(\mathbf{s}) | \hat{A} | \phi(\mathbf{s}) \rangle$$

$$|\phi(\mathbf{s})\rangle = e^{-\beta H/2} |\mathbf{s}\rangle / P^{\frac{1}{2}}(\mathbf{s})$$

$$P(\mathbf{s}) = \langle \phi(\mathbf{s}) | \phi(\mathbf{s}) \rangle$$

Minimally Entangled Typical Thermal States

Suggests a sampling method

sample states $|\phi(\mathbf{s})\rangle$

$$\langle \hat{A} \rangle = \sum_{\mathbf{s}} \frac{P(\mathbf{s})}{Z} \langle \phi(\mathbf{s}) | \hat{A} | \phi(\mathbf{s}) \rangle$$

with probability $\dfrac{P(\mathbf{s})}{Z}$

Minimally Entangled Typical Thermal States

Can sample with following algorithm:

1. given product state $|\mathbf{s}\rangle$ compute $|\tilde{\phi}(\mathbf{s})\rangle = e^{-\beta H/2}|\mathbf{s}\rangle$

2. compute estimator $\langle\phi(\mathbf{s})|\hat{A}|\phi(\mathbf{s})\rangle$

3. collapse to obtain new product state $|\phi(\mathbf{s})\rangle \longrightarrow |\mathbf{s}'\rangle$
   (no rejection step vs. Metropolis)

(Step #3 should seem mysterious)

Minimally Entangled Typical Thermal States

Assuming steps #1 and #2 can be done using MPS

Why is collapse the right thing to do? $|\phi(\mathbf{s})\rangle \xrightarrow{\hspace{2cm}} |\mathbf{s}'\rangle$

Minimally Entangled Typical Thermal States

Define collapse to be selection of $|\mathbf{s}'\rangle$

with probability $|\langle\mathbf{s}'|\phi(\mathbf{s})\rangle|^2$

Markov chain with transition probability

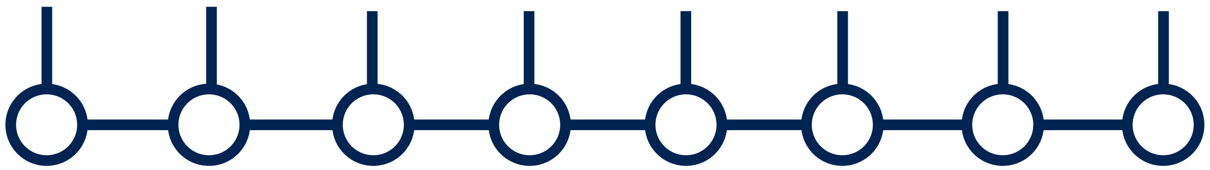
$$p_{\mathbf{s}\to\mathbf{s}'} = |\langle\mathbf{s}'|\phi(\mathbf{s})\rangle|^2$$

Can show obeys detailed balance

$$\frac{P(\mathbf{s})}{Z}p_{\mathbf{s}\to\mathbf{s}'} = \frac{P(\mathbf{s}')}{Z}p_{\mathbf{s}'\to\mathbf{s}}$$

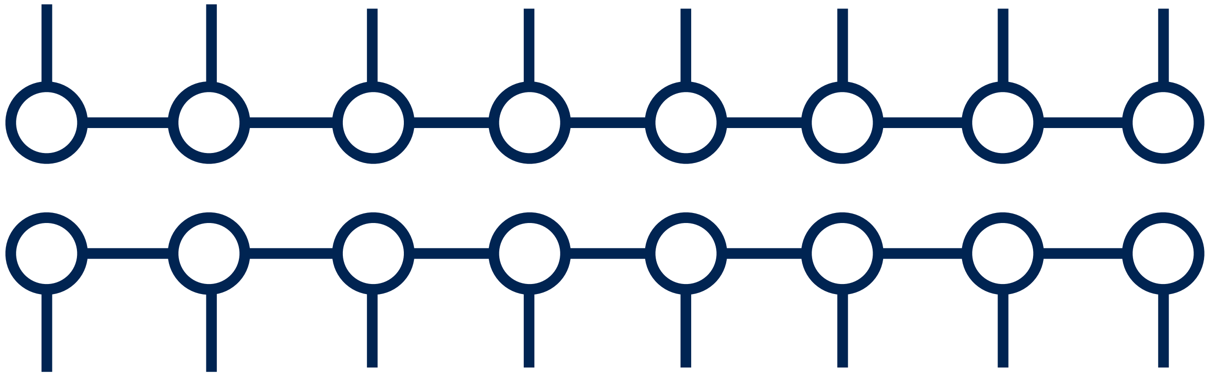
Minimally Entangled Typical Thermal States

Efficient algorithm for collapsing an MPS



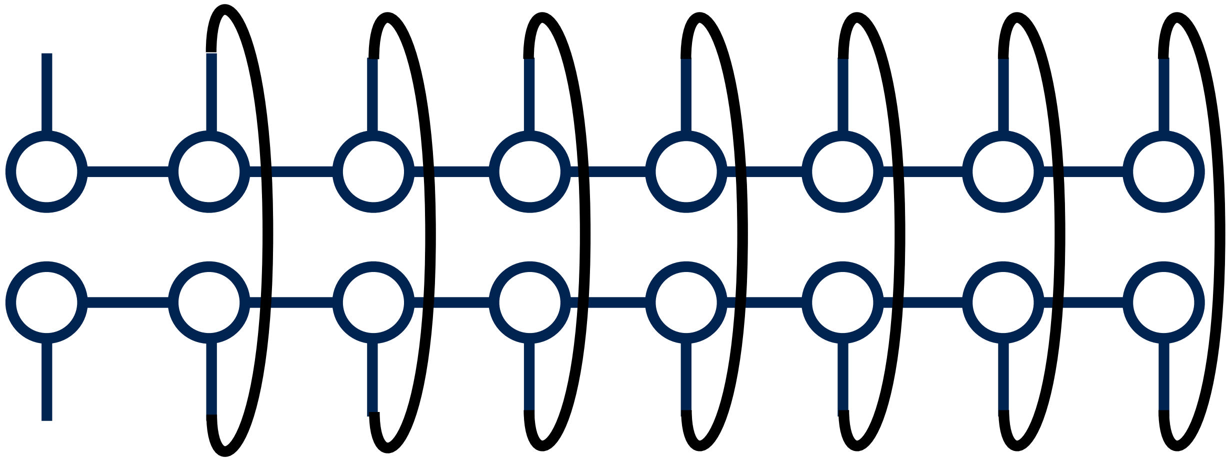
Minimally Entangled Typical Thermal States

Efficient algorithm for collapsing an MPS



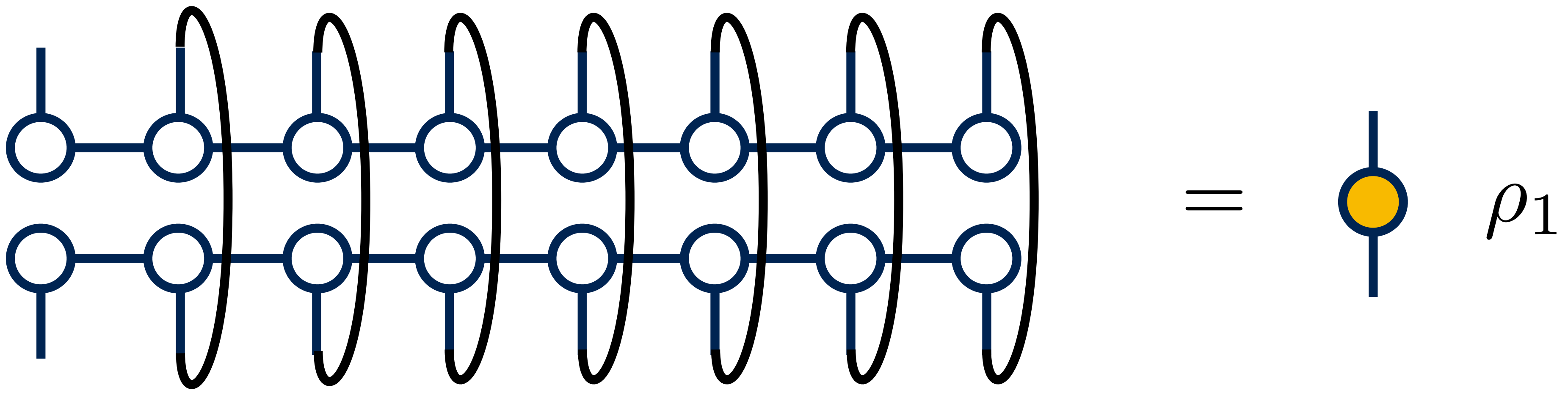
Minimally Entangled Typical Thermal States

Efficient algorithm for collapsing an MPS



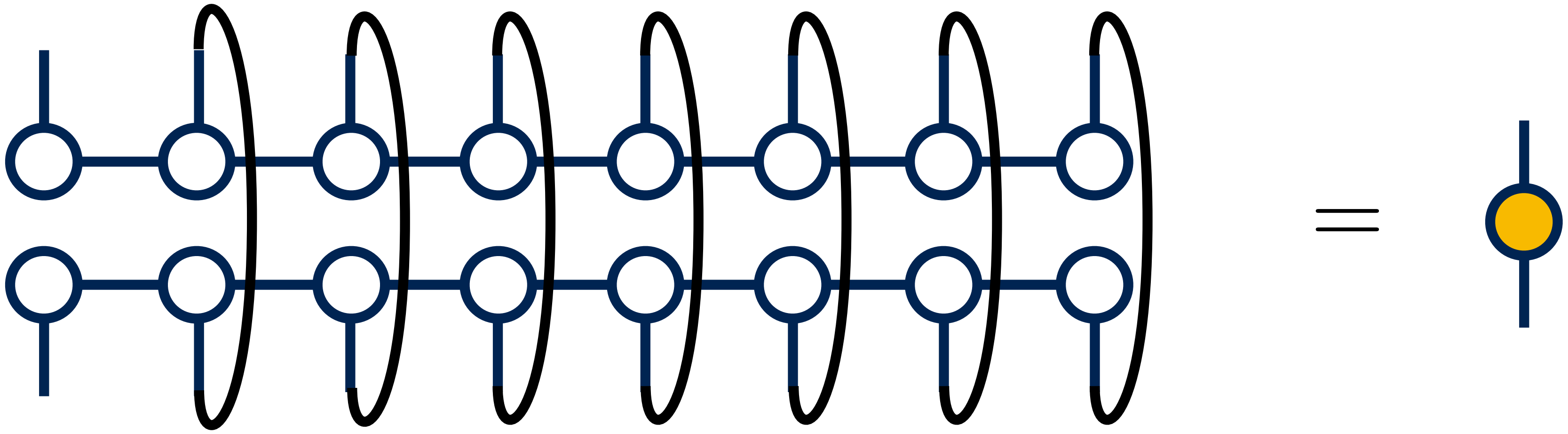
Minimally Entangled Typical Thermal States

Efficient algorithm for collapsing an MPS

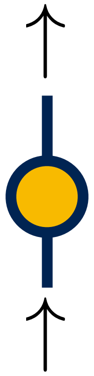


Minimally Entangled Typical Thermal States

Efficient algorithm for collapsing an MPS

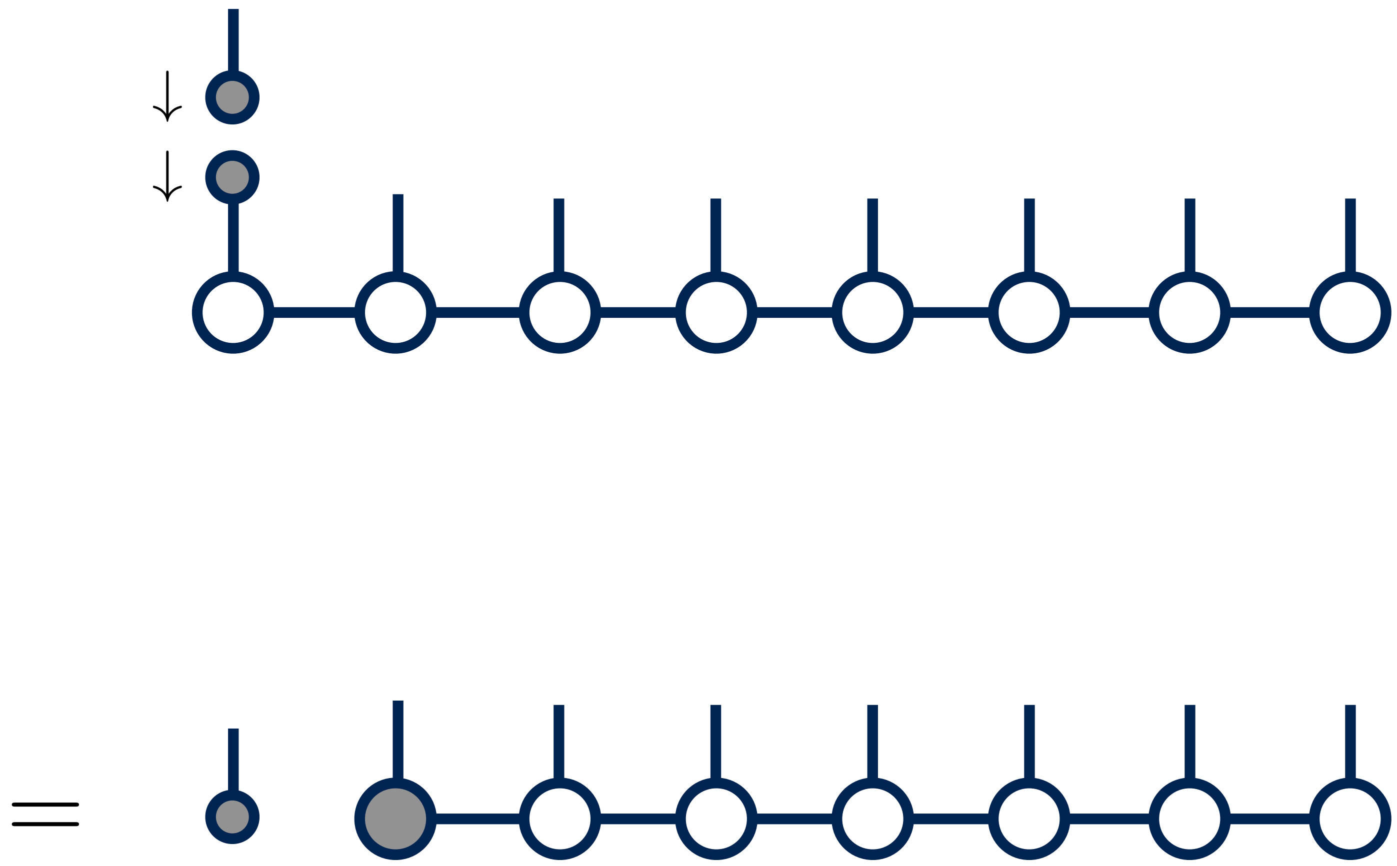


$= \quad \rho_1$

Select spin up, probability $\rho_1^{\uparrow\uparrow} =$  , or down with $\rho_1^{\downarrow\downarrow} =$ 

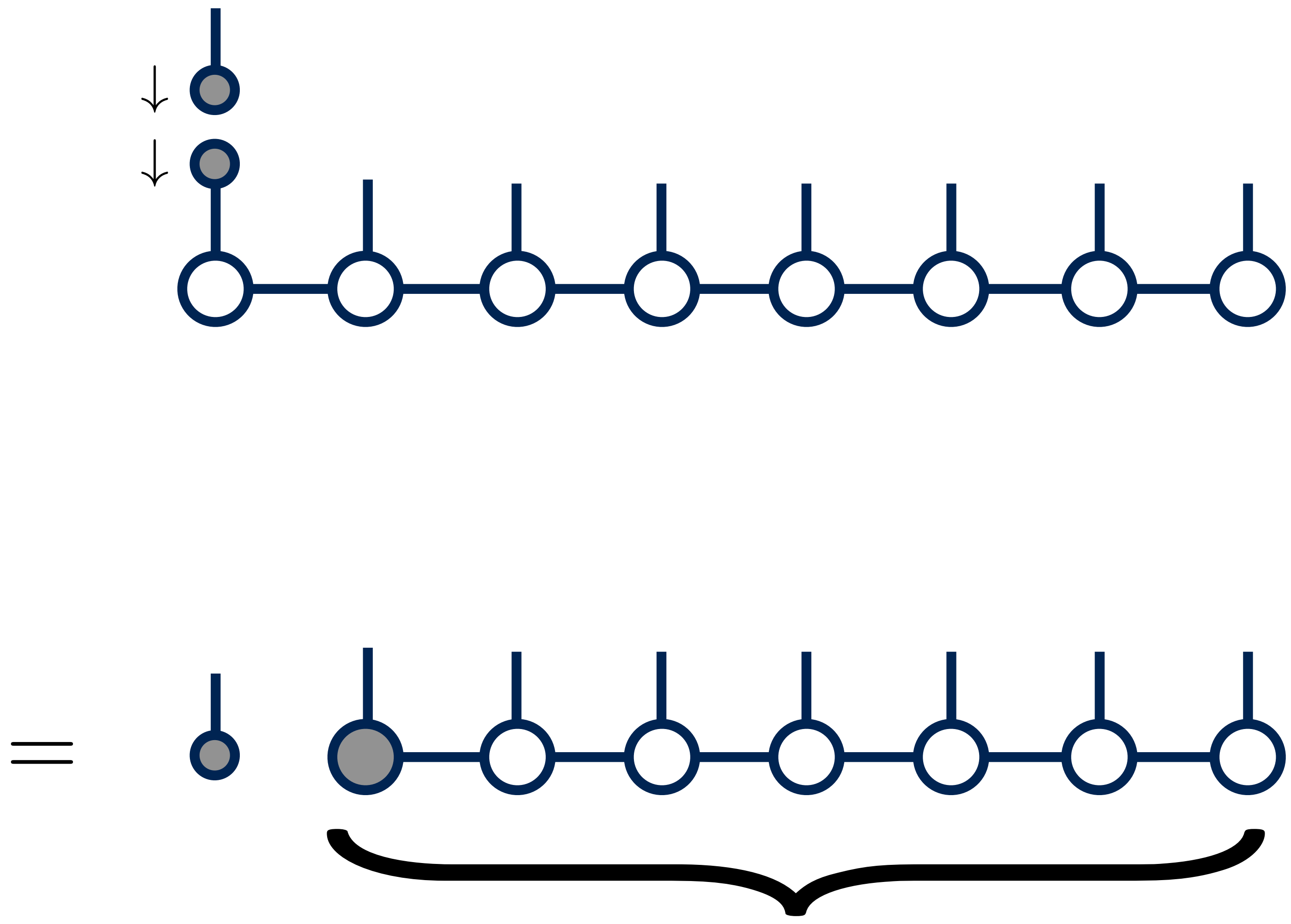
Minimally Entangled Typical Thermal States

Say 'down' is picked for site 1



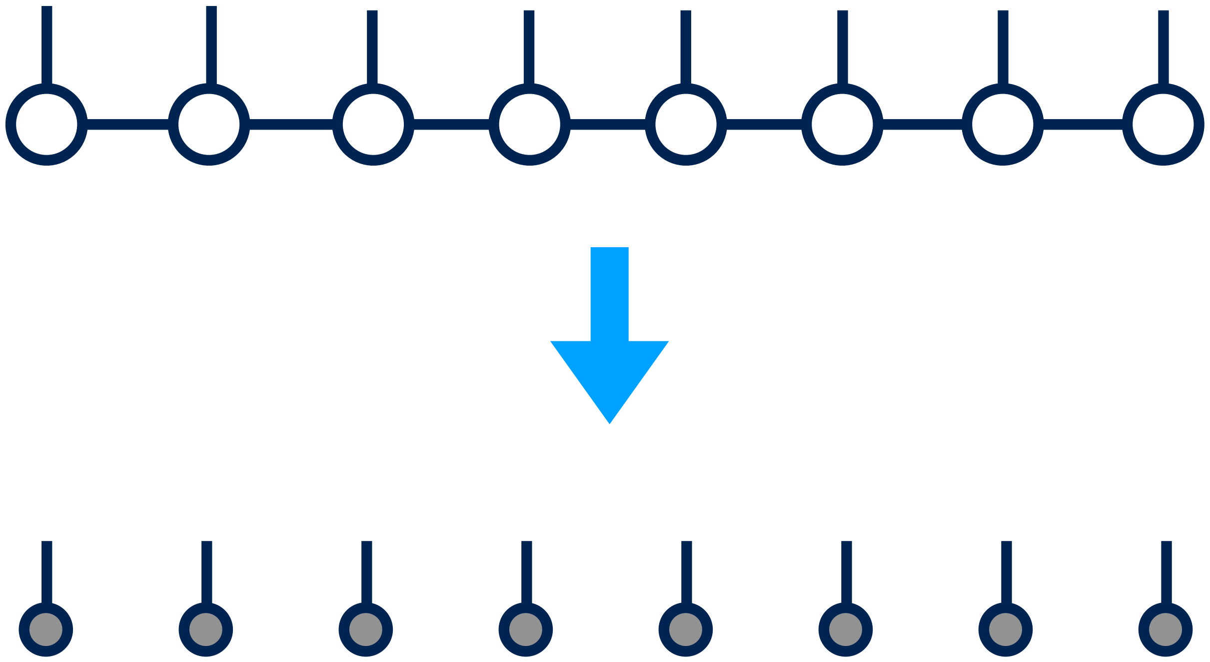
Minimally Entangled Typical Thermal States

Say 'down' is picked for site 1



repeat for sites 2,3,...,N

Minimally Entangled Typical Thermal States



Interestingly, will draw product state $|\mathbf{s}\rangle$ with probability $|\langle\mathbf{s}|\psi\rangle|^2$ in a single run of the algorithm

Minimally Entangled Typical Thermal States

Having computed N states $|\phi(\mathbf{s}_1)\rangle,\ |\phi(\mathbf{s}_2)\rangle,\ |\phi(\mathbf{s}_3)\rangle, \ldots$

Estimate expectation values as

$$\langle \hat{A} \rangle \simeq \frac{1}{N} \sum_{j=1}^{N} \langle \phi(\mathbf{s}_j) | \hat{A} | \phi(\mathbf{s}_j) \rangle$$

Minimally Entangled Typical Thermal States

Computing each state $|\phi(\mathbf{s})\rangle$ has similar cost to ground state DMRG for low T

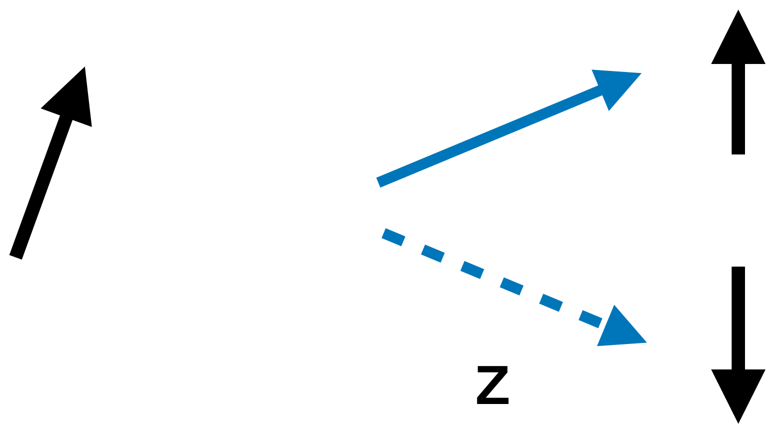
For higher T even cheaper

Overall scaling like $\beta N m^3$

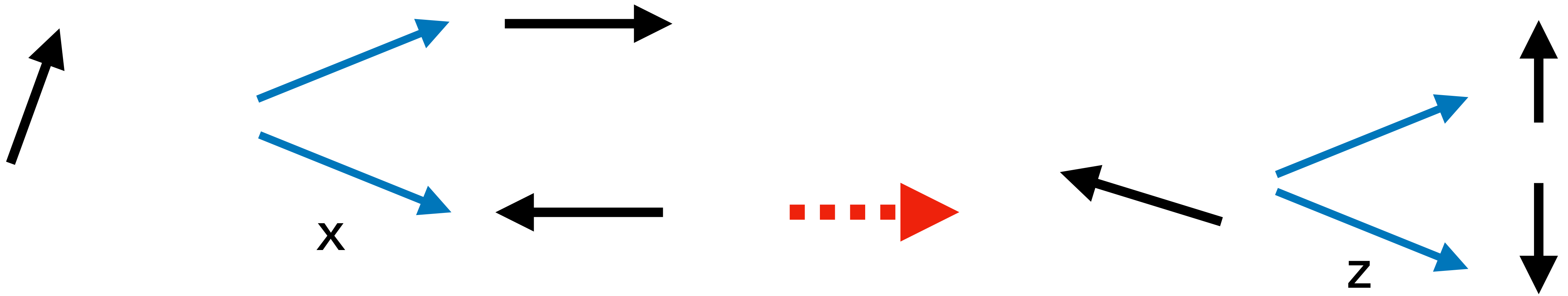
Minimally Entangled Typical Thermal States

Can also choose collapse basis to minimize autocorrelation:

Collapsing into z basis at high T "remembers" prev. state



Switching to x and z basis on alternating steps mixes rapidly



Minimally Entangled Typical Thermal States

States $|\phi(\mathbf{s})\rangle$ are called METTS

$$|\phi(\mathbf{s})\rangle \propto e^{-\beta H/2}|\mathbf{s}\rangle$$

"Typical" because

- averaging them gives correct thermal result
- product states at high T, ground state at low T
- if system breaks symmetries, METTS do too
- expectation values converge rapidly

Example of a METTS:

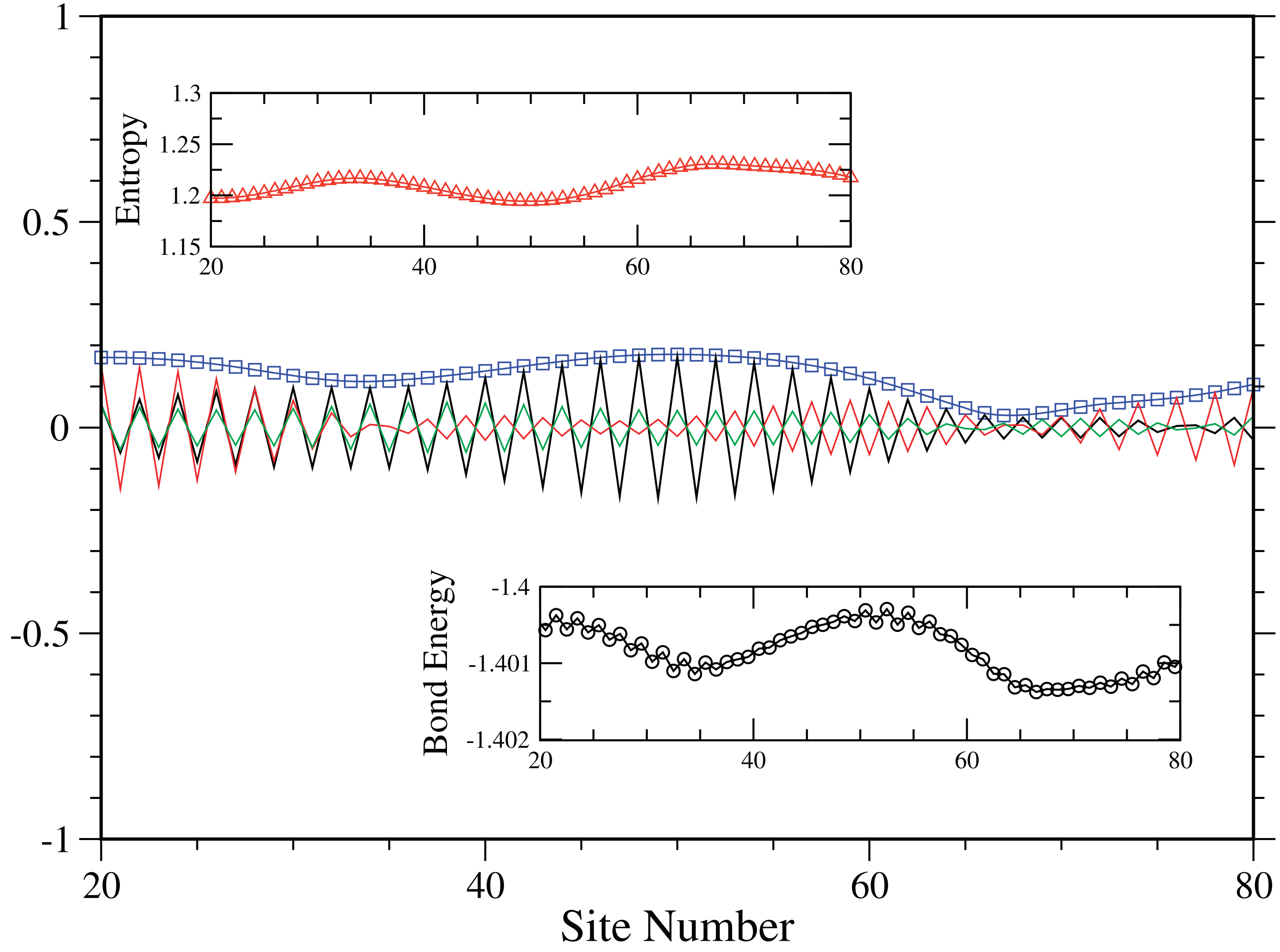
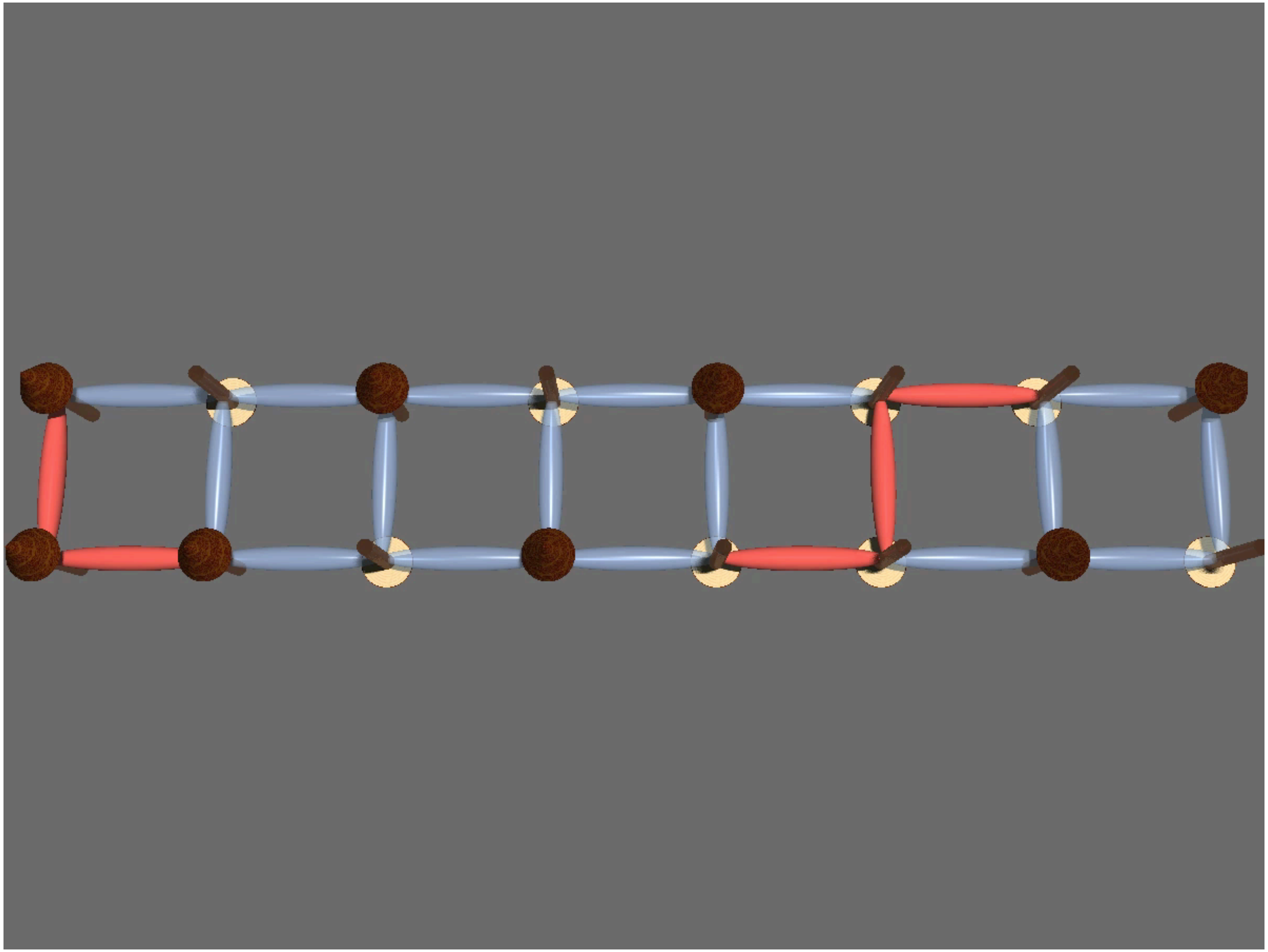


Figure 9. Properties of a METTS produced for the 100 site $S = 1$ Heisenberg chain at $T = 0.1$, central 80 sites. In the main plot, the solid lines (red, green and black) show the three components of $\langle \vec{S} \rangle$, while the (blue) boxes show $|\langle \vec{S} \rangle|$. The entanglement entropy on each bond is shown in the top inset, while the expectation value of each Hamiltonian bond term is shown in the bottom inset.

Minimally Entangled Typical Thermal States

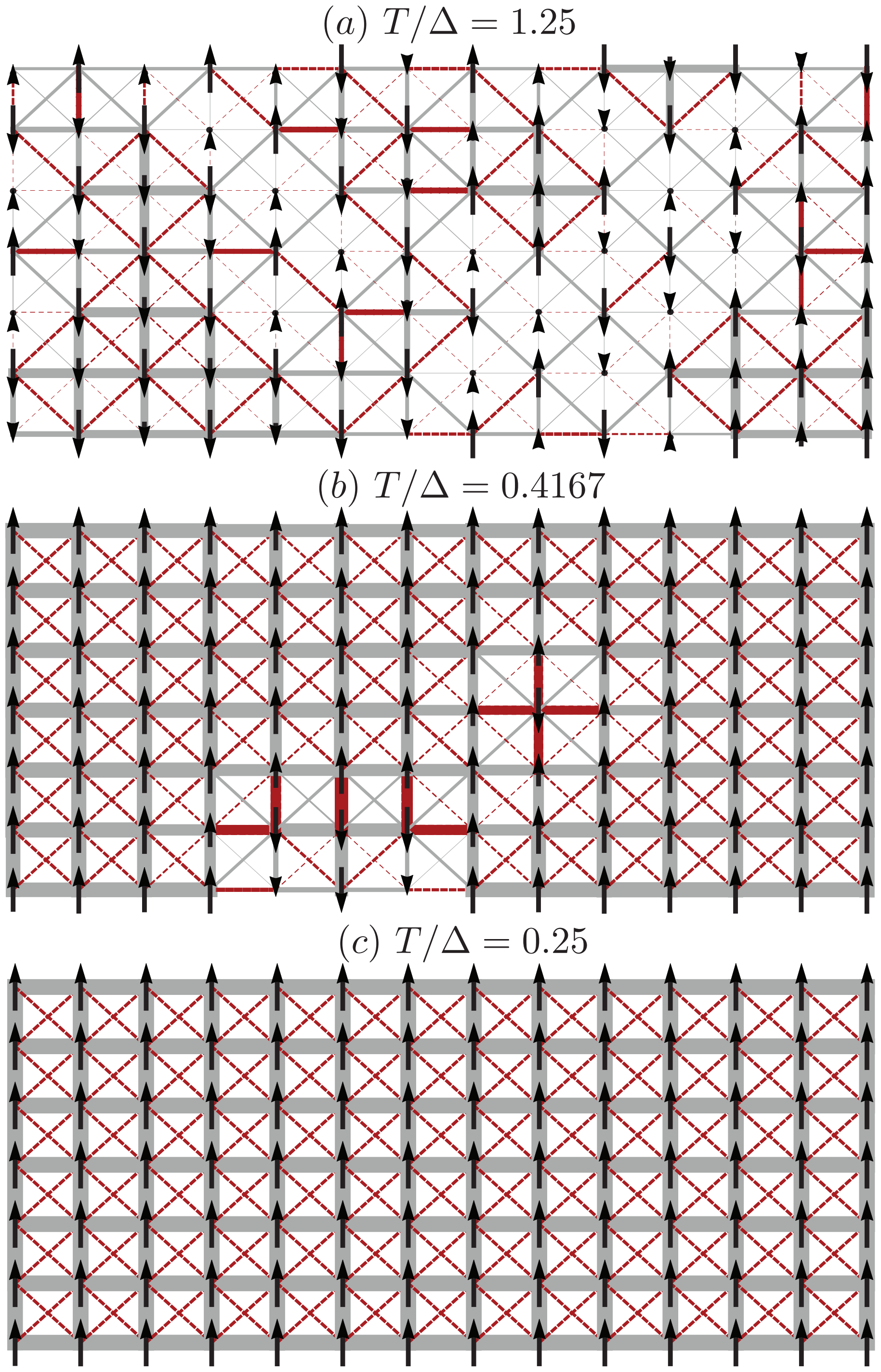
Movie of the METTS algorithm (S=1/2 Heisenberg ladder, $\beta = 5$)



Minimally Entangled Typical Thermal States

METTS produced for the *frustrated* XXZ model ($\Delta = 5$)

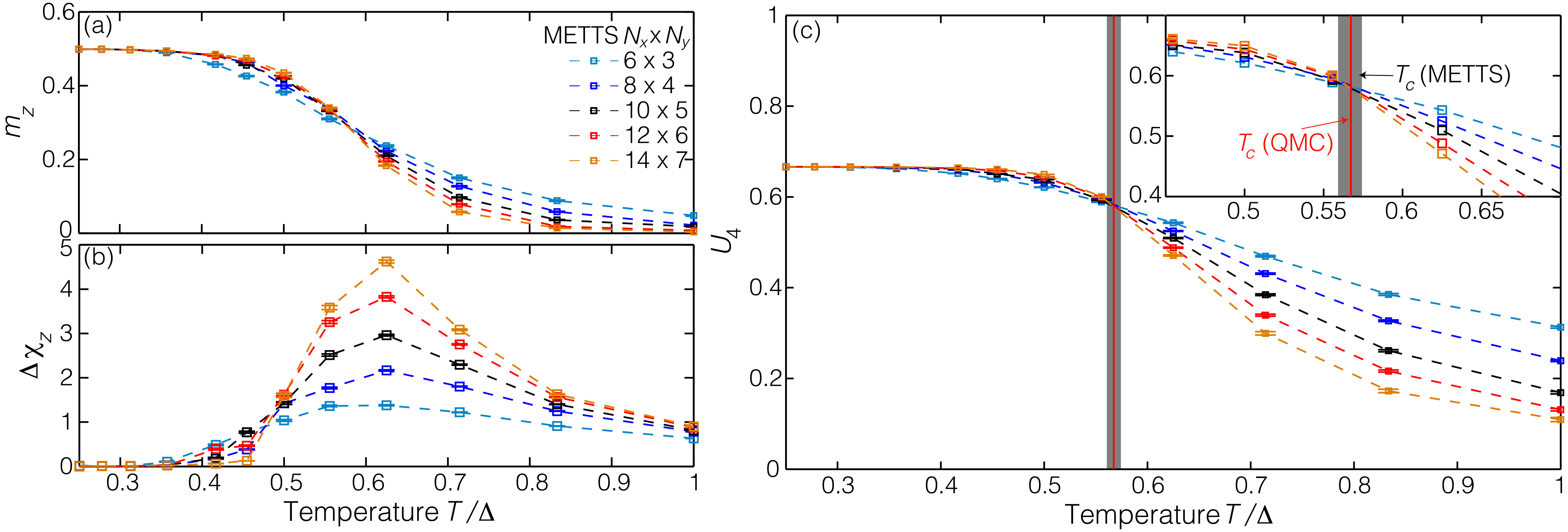
$$\begin{aligned}\hat{H} &= J_1 \sum_{\langle i,j \rangle} \left(\hat{S}^x_i \hat{S}^x_j + \hat{S}^y_i \hat{S}^y_j + \Delta \hat{S}^z_i \hat{S}^z_j\right) \\ &\quad + J_2 \sum_{\langle\langle i,j \rangle\rangle} \left(\hat{S}^x_i \hat{S}^x_j + \hat{S}^y_i \hat{S}^y_j + \Delta \hat{S}^z_i \hat{S}^z_j\right),\end{aligned}$$



Minimally Entangled Typical Thermal States

Using METTS to obtain thermodynamic properties

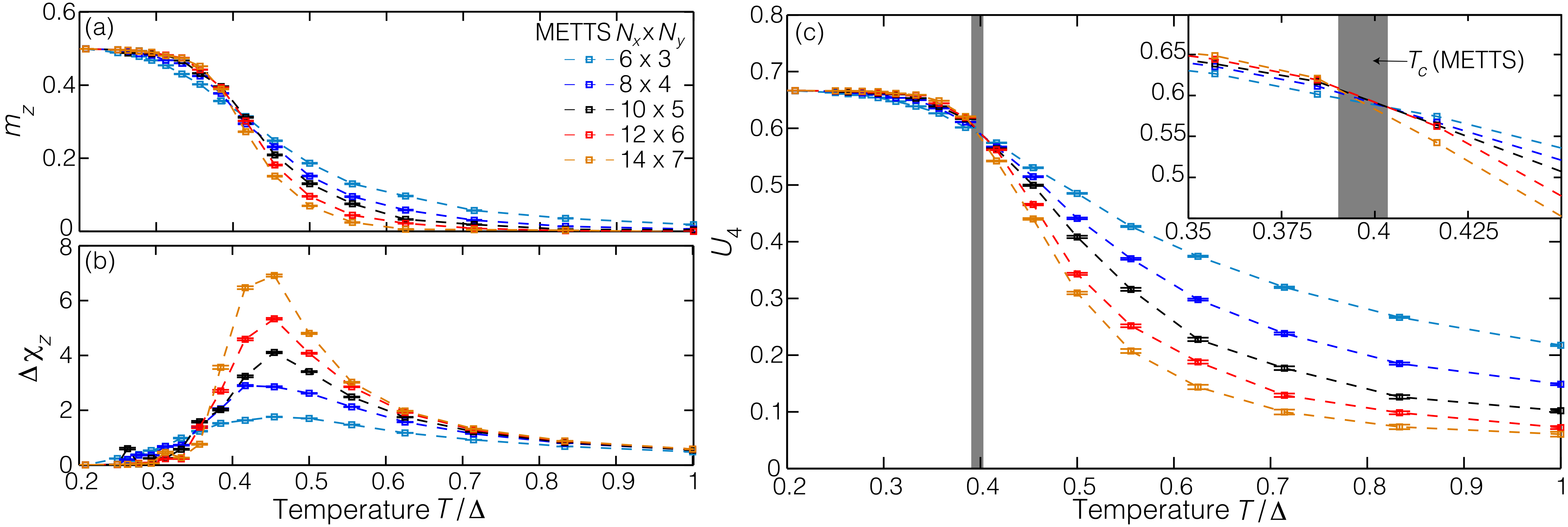
Test on unfrustrated case ($J_2 = 0$)



Minimally Entangled Typical Thermal States

Using METTS to obtain thermodynamic properties

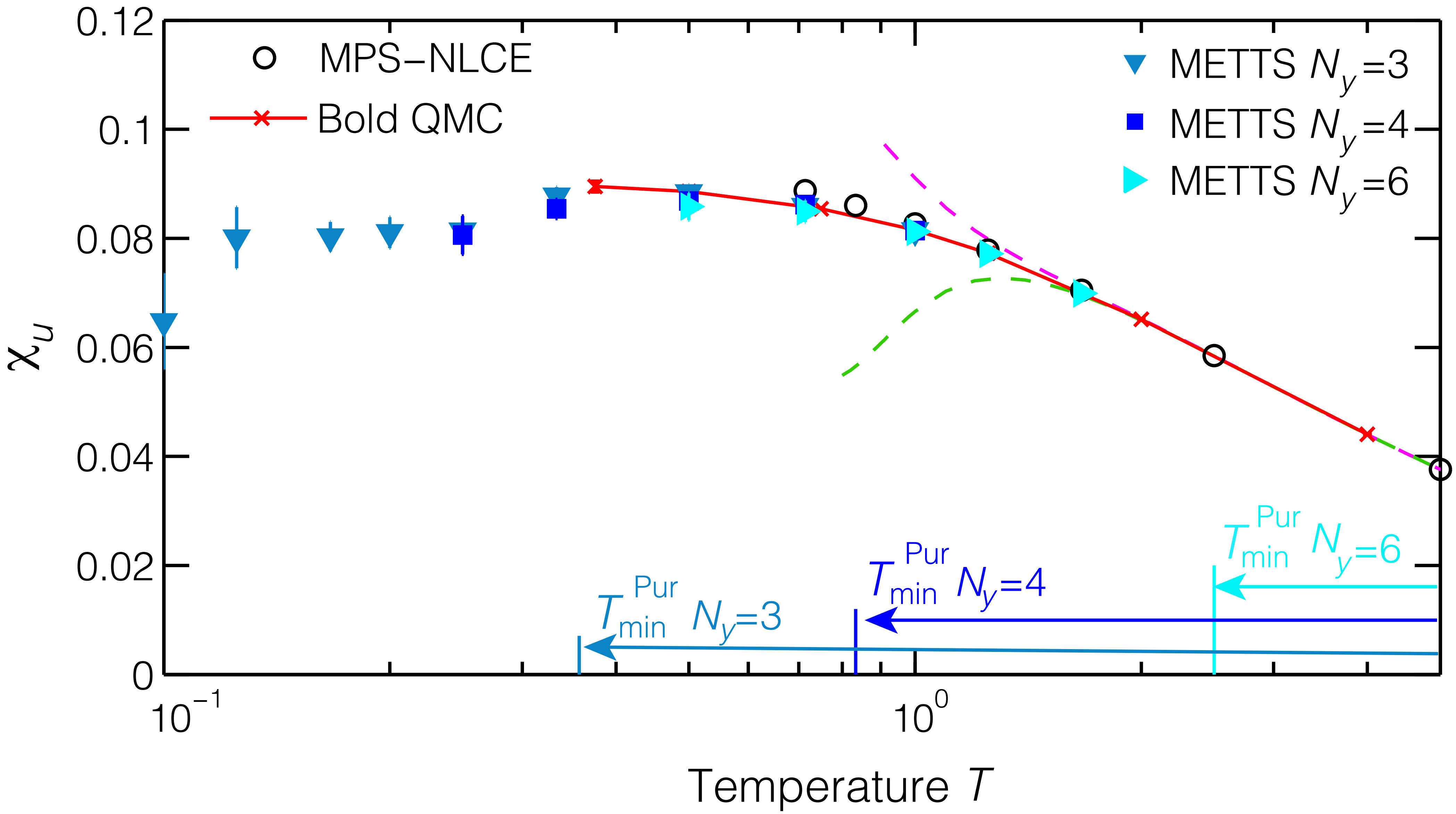
Apply to frustrated case ($J_2/|J_1| = 0.2$)



Minimally Entangled Typical Thermal States

Using METTS to obtain thermodynamic properties

Susceptibility of triangular lattice Heisenberg model

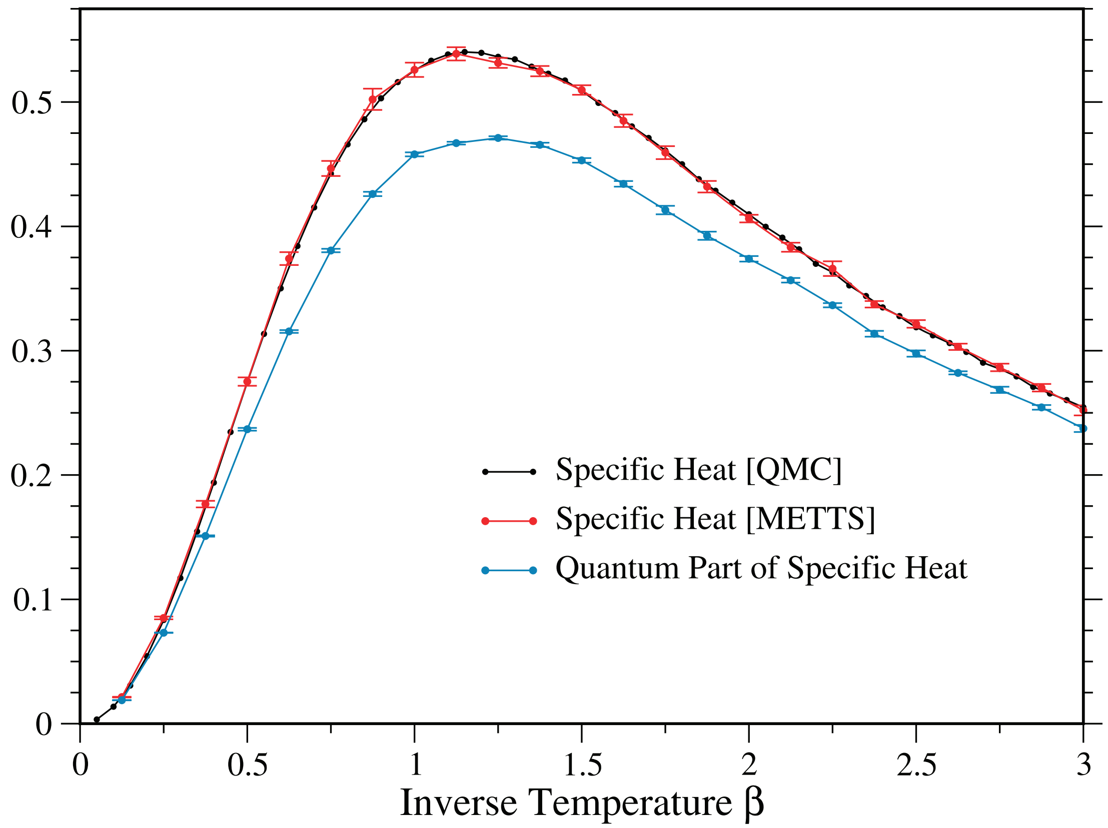


Minimally Entangled Typical Thermal States

Interesting separation of quantum vs. classical fluctuations

"Quantum specific heat" $C^q = \frac{\beta^2}{N} \sum_i \frac{P(i)}{\mathcal{Z}} \left[\langle H^2 \rangle_i - \langle H \rangle_i^2 \right]$

Classical part of specific heat $(C - C^q)$ is precisely the sample variance of the energy



N=100, S=1 chain

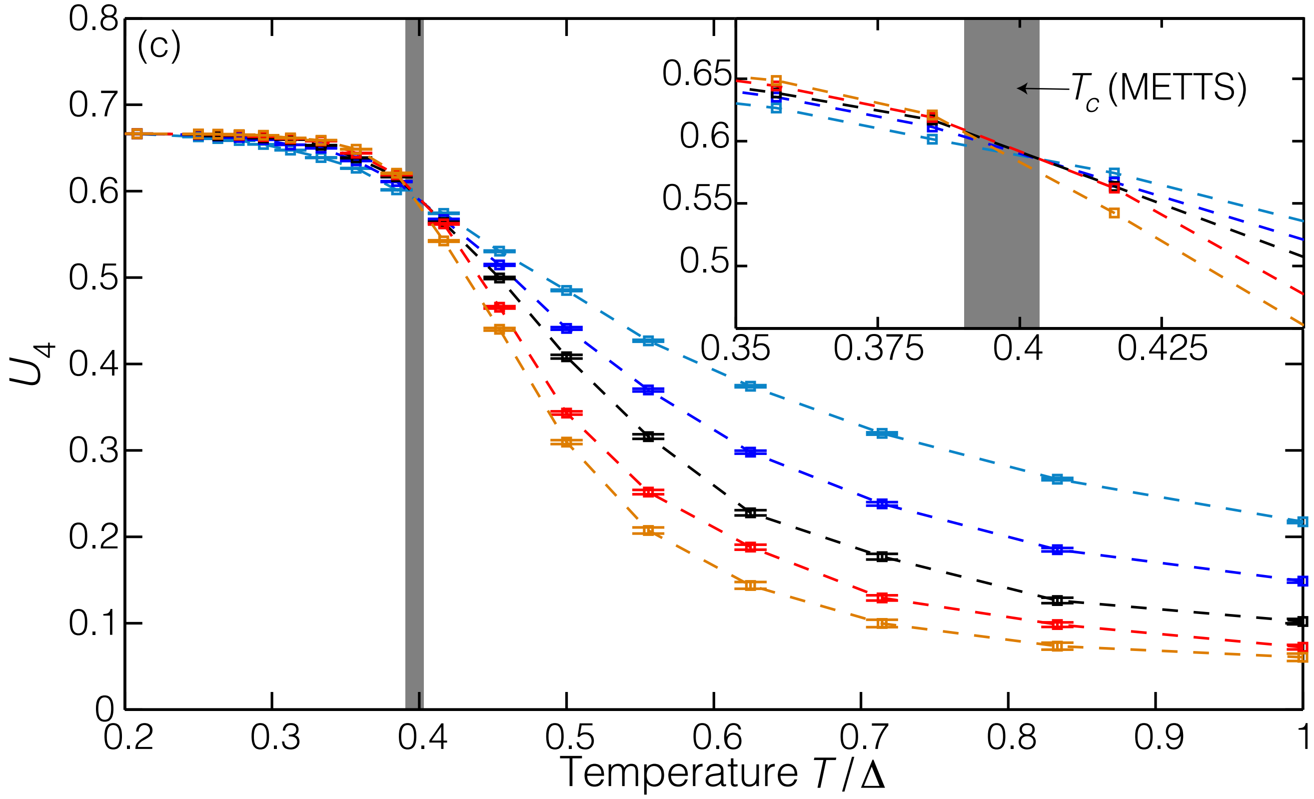
Stoudenmire, White, NJP 12, 055026 (2010)

Summary of Finite Temperature Methods

Purification effective for higher T

METTS sampling efficient for lower T

Can apply both to 2D and frustrated / fermionic systems



Future Directions

Many interesting ideas to improve both methods

- exploit freedom to apply arb. unitary to collapse basis (entangled collapses)

- mix ancilla (traced) and METTS (collapsed) sites

- make better use of symmetries [Binder, Barthel, PRB **95**, 195148]

- apply mixed state truncation protocol for dynamics
  [White, Mong, Zaletel, Refael arxiv:1707.01506]