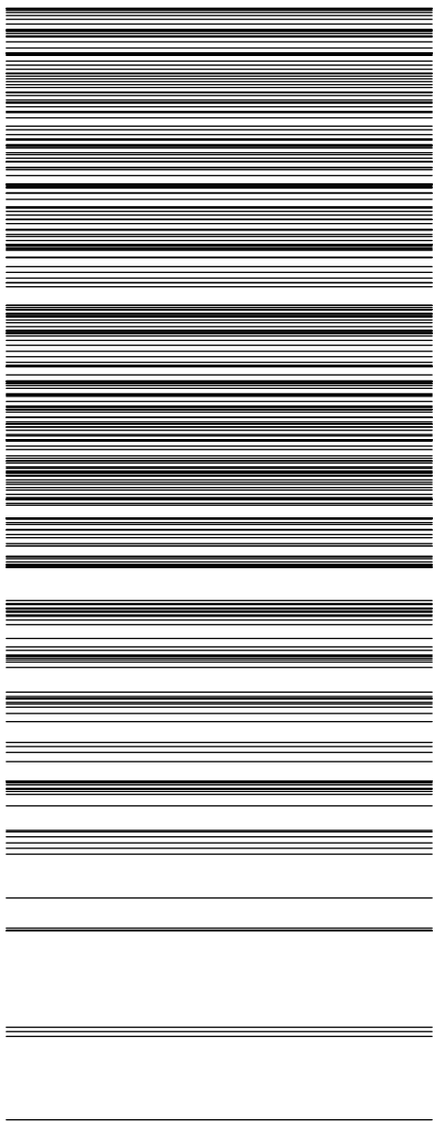
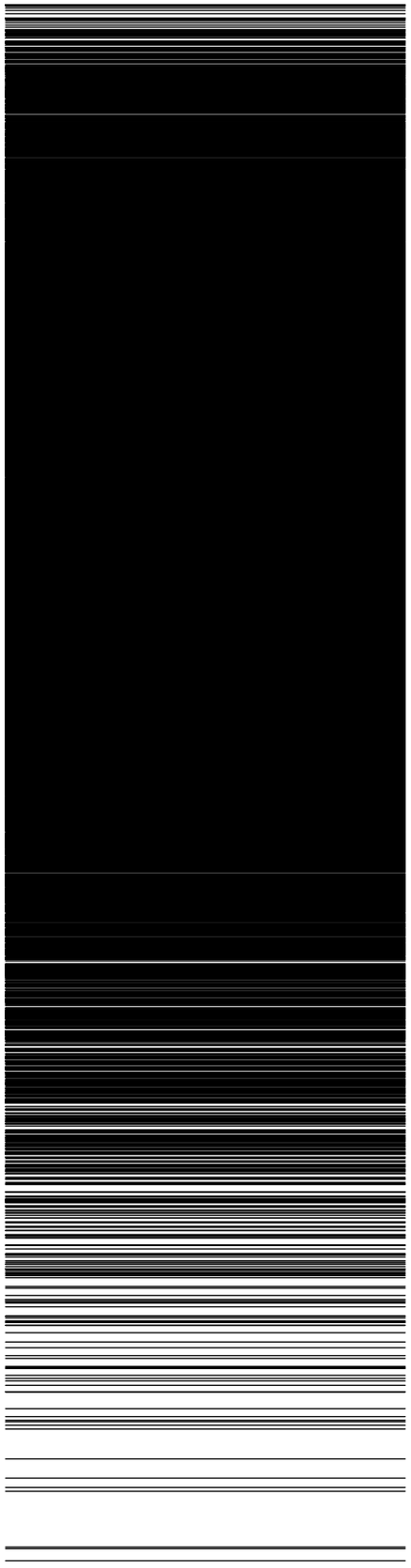
The Density Matrix Renormalization Group: Introduction and Overview

- Introduction to DMRG as a low entanglement approximation
  - Entanglement
  - Matrix Product States
  - Minimizing the energy and DMRG sweeping
- The low entanglement viewpoint versus the historical RG viewpoint
- Time evolution for spectral functions
- Some generalizations and extensions of DMRG
- Methods for 2D
  - applications to t-J model and stripes
  - Frustrated magnets and spin liquids

Software: ALPS (well developed, inflexible); itensor.org (new)



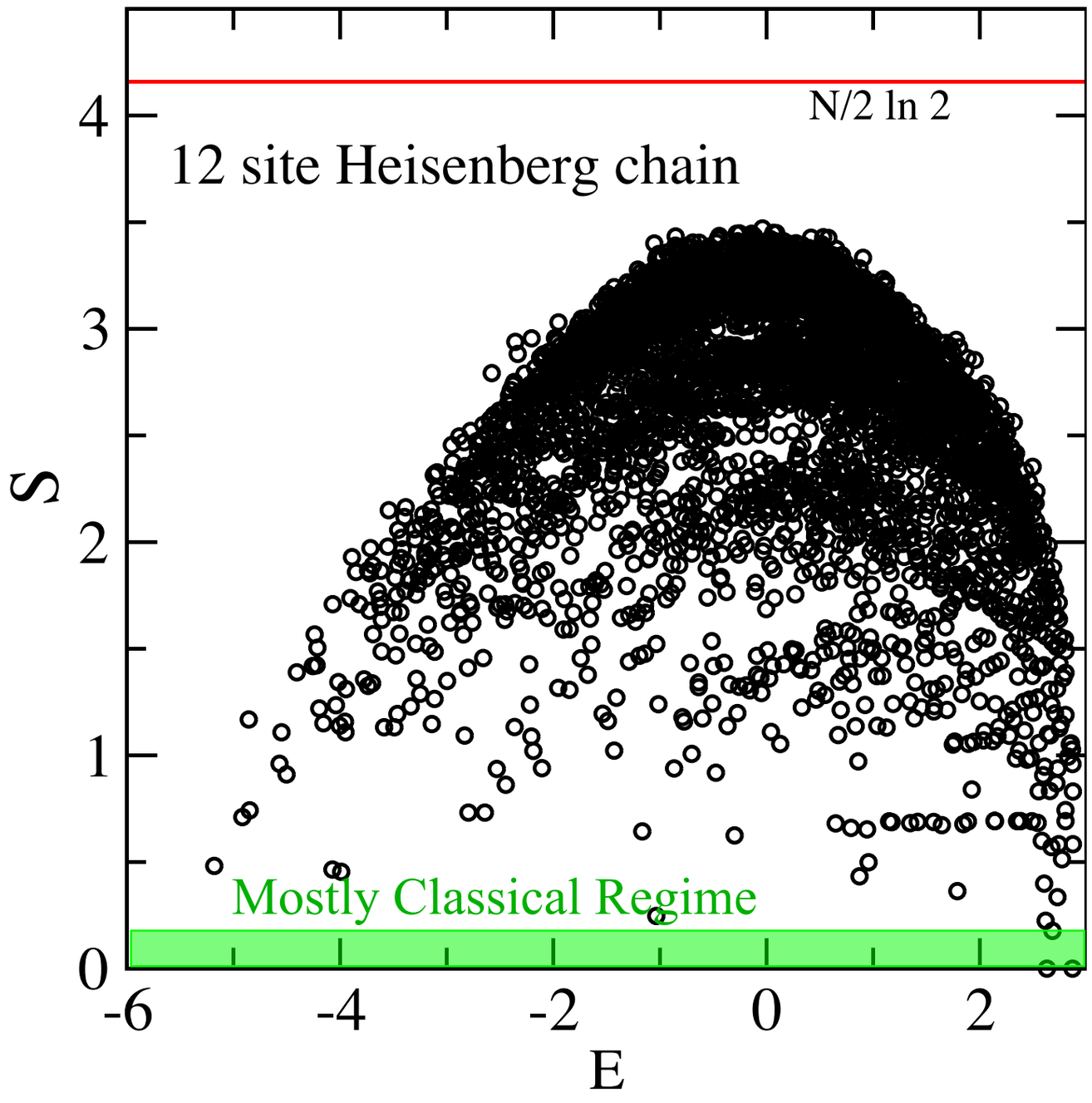
N=8



N=12

Energy levels of S=1/2 Heisenberg chains

Bulk eigenstates are "super-entangled"



Von Neumann Entanglement entropy S for every eigenstate (system divided in center)

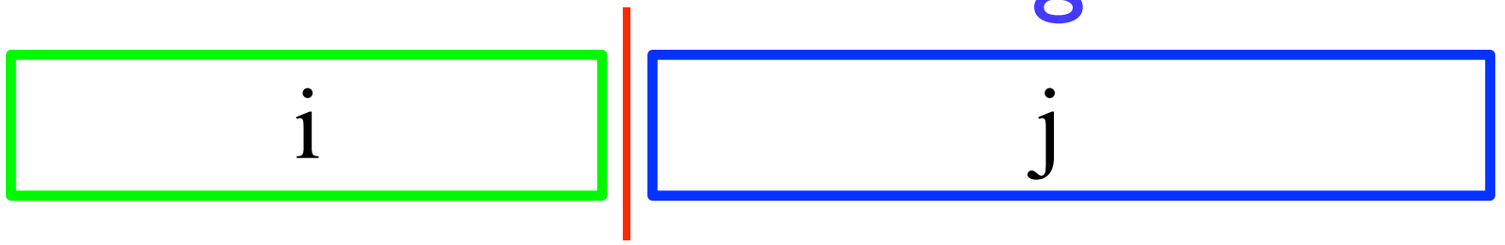
What is entanglement?

- Intuitive idea: general correlation between two parts of a system (think two separate spins: a Bell pair)
- Not always obvious: Which is more entangled?
  - 1) $|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle$ or
  - 2) $|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle$??

SVD/Schmidt Decomposition

- Let the system have two parts: left and right
  - $|\Psi\rangle = \sum \Psi_{lr} \; |l\rangle \, |r\rangle$
- Treat Ψ_{lr} as a matrix: perform the simple matrix factorization "singular value decomposition" (SVD): $\Psi = U \, D \, V$, with U and V unitary, D diagonal.
- The diagonal elements λ of D are the singular values or Schmidt coefficients. In quantum information this is called the Schmidt decomposition. The Schmidt basis vectors are given as $|\alpha\rangle = \sum_r V_{\alpha r} \; |r\rangle$, $|\tilde{\alpha}\rangle = \sum_l U_{\alpha l} \; |l\rangle$; the wavefunction is $|\Psi\rangle = \sum_\alpha \; \lambda_\alpha \; |\tilde{\alpha}\rangle \, |\alpha\rangle$ (diagonal).
- The reduced density matrix for the left side is:
  - $\rho_{ll'} = \sum_r \Psi_{lr} \Psi_{l'r}$
- If you insert the SVD, you find that U contains the eigenvectors of ρ, and the eigenvalues are $(\lambda_\alpha)^2$. Note $\sum_\alpha (\lambda_\alpha)^2 = 1$ (normalization)

Von Neumann entanglement entropy



S ~ entanglement across the cut

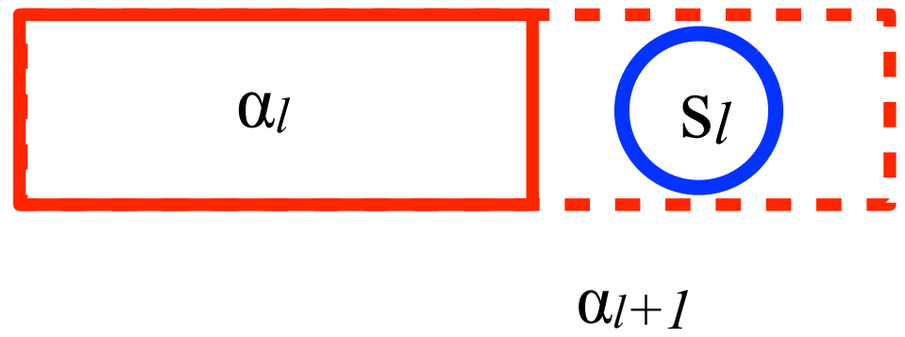
- If we think of $(\lambda_\alpha)^2$ as the probability of the state $|\tilde{\alpha}\rangle |\alpha\rangle$, then we can plug in the standard probability formula to get the von Neumann entropy
  - $S = -\sum_\alpha (\lambda_\alpha)^2 \ln (\lambda_\alpha)^2$
- There are several other entropies (different forumulas)

- Low entanglement = small S occurs when the λ_α fall off fast as the index α increases.
- Thus we have a natural *low entanglement approximation:* approximate the wavefunction by keeping a small number of α (the largest).
- In DMRG we *imagine* we do this Schmidt decomp for all positions of the dividing line between left and right.

Matrix Product States

- Insert a truncated set of density matrix/Schmidt eigenstates at every nn link (1D) (total error = sum of probabilities you've thrown away)
- The Schmidt basis states for position $l+1$ must be linear combinations of those at l

$$|\alpha_{l+1}\rangle = \sum_{\alpha_l, s_l} A[s_l]_{\alpha_{l+1}\alpha_l} |s_l\rangle|\alpha_l\rangle$$



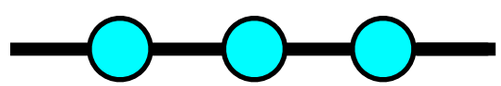
- This produces a Matrix Product State (MPS) formula for the wavefunction:

$$\Psi(s_1,s_2,..s_N) \approx A^1[s_1]\, A^2[s_2] \ldots A^N[s_N]$$

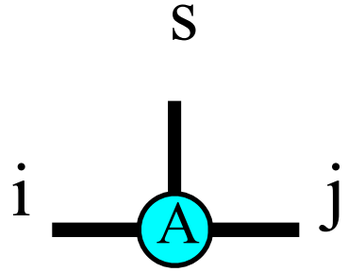
- A function is just a rule for giving a number from the inputs--here the {s} tell which matrices to multiply (first and last A's are vectors).

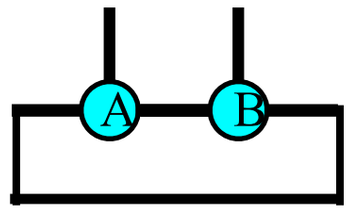
Diagrams for Matrix Product States

Vertices are matrices or tensors. All internal lines are summed over.
External lines are external indices, usually associated with states

Ordinary Matrix Multiplication: ABC = 

In an MPS, the basic unit has an extra index, like a Pauli spin matrix; or you can call it a tensor

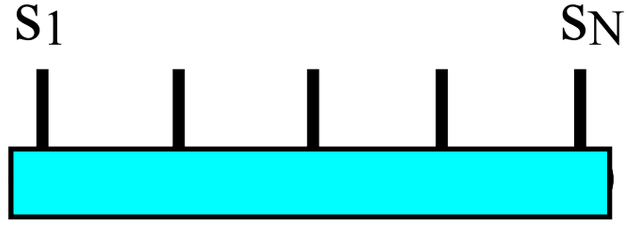
$A^{[s]}_{ij}$ = 

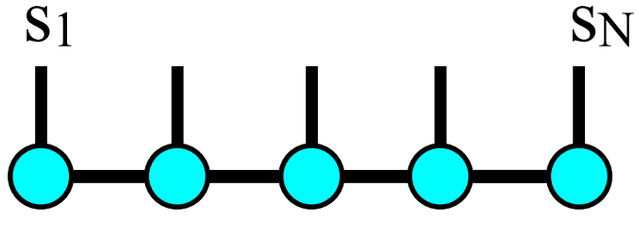
Simple diagram: $Tr[A^s B^t]$ = 

gives f(s,t)

Dimensions: i, j: m or D s: d

Matrix Product State: $\Psi(s_1,s_2,..s_N) \approx A^1[s_1]\, A^2[s_2] \ldots A^N[s_N]$

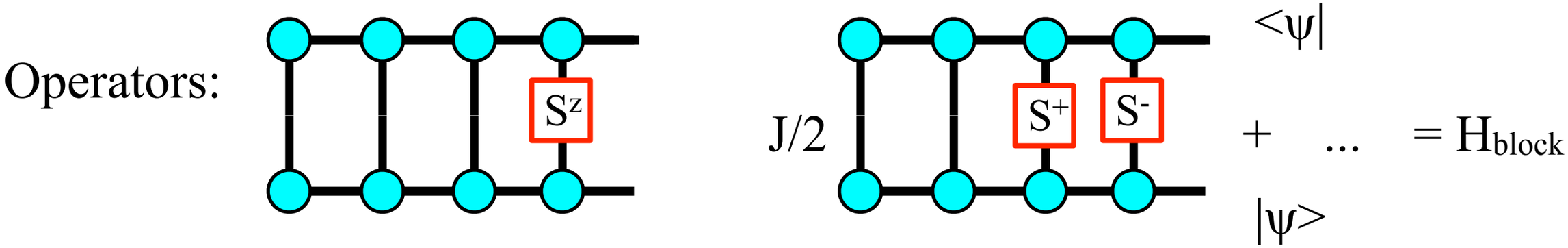
2^N 

$\approx$ 

$N\, m^2$ for m x m matrices

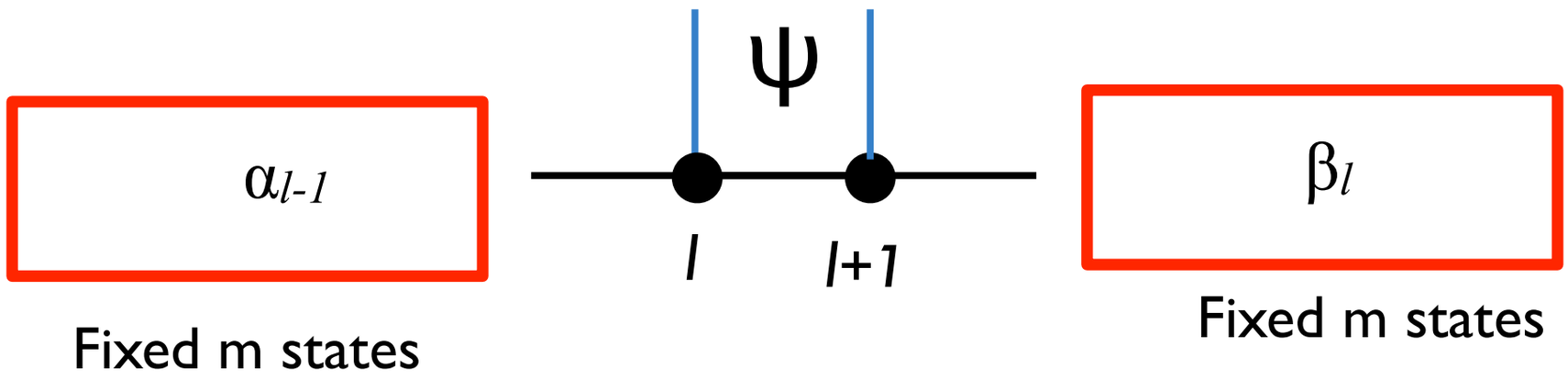
MPS as Variational states

- Two things needed:
  - Evaluate energy and observables efficiently
  - Optimize parameters efficiently to minimize energy
- Observables:



  - Working left to right, just matrix multiplies, N m^3
- Optimization:
  - General-purpose nonlinear optimization is hard
  - Lanczos solution to eigenvalue problem is one of the most efficient optimization methods (also Davidson method). Can we use that? Yes!

DMRG algorithm: one step



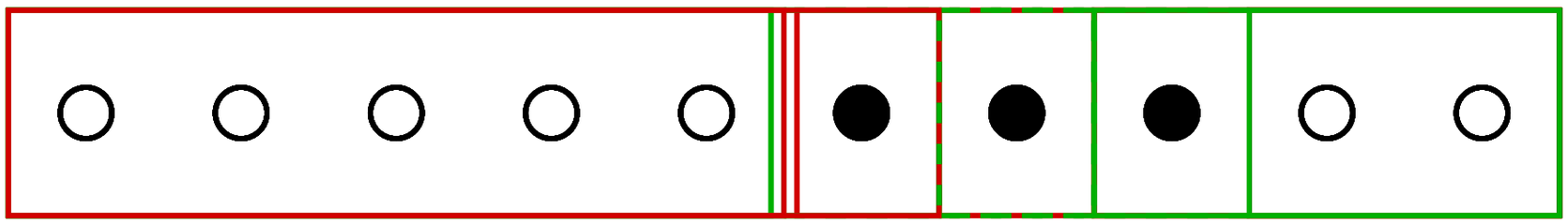
1. Use exact diagonalization to get the lowest energy $\Psi(\alpha_{l-1}, s_l, s_{l+1}, \beta_{l+2})$ within the basis of fixed block approximate Schmidt vectors α_{l-1} β_{l+2} and two sites s_l, s_{l+1}

Do an SVD on the 4 parameter wavefunction to split it up into new $A[s_l]$ and $A[s_{l+1}]$

DMRG Sweeping Algorithm

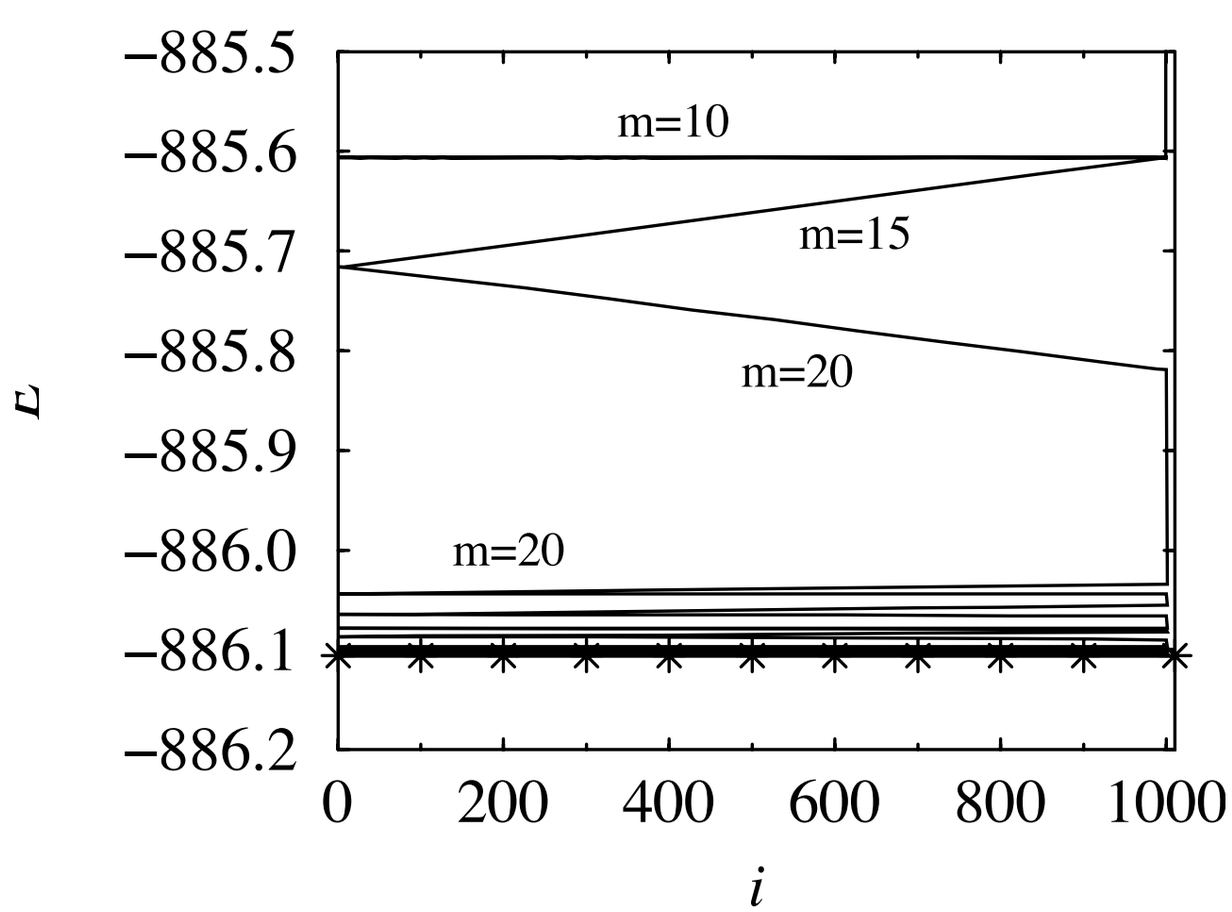
•The optimization sweeps back and forth through the system.

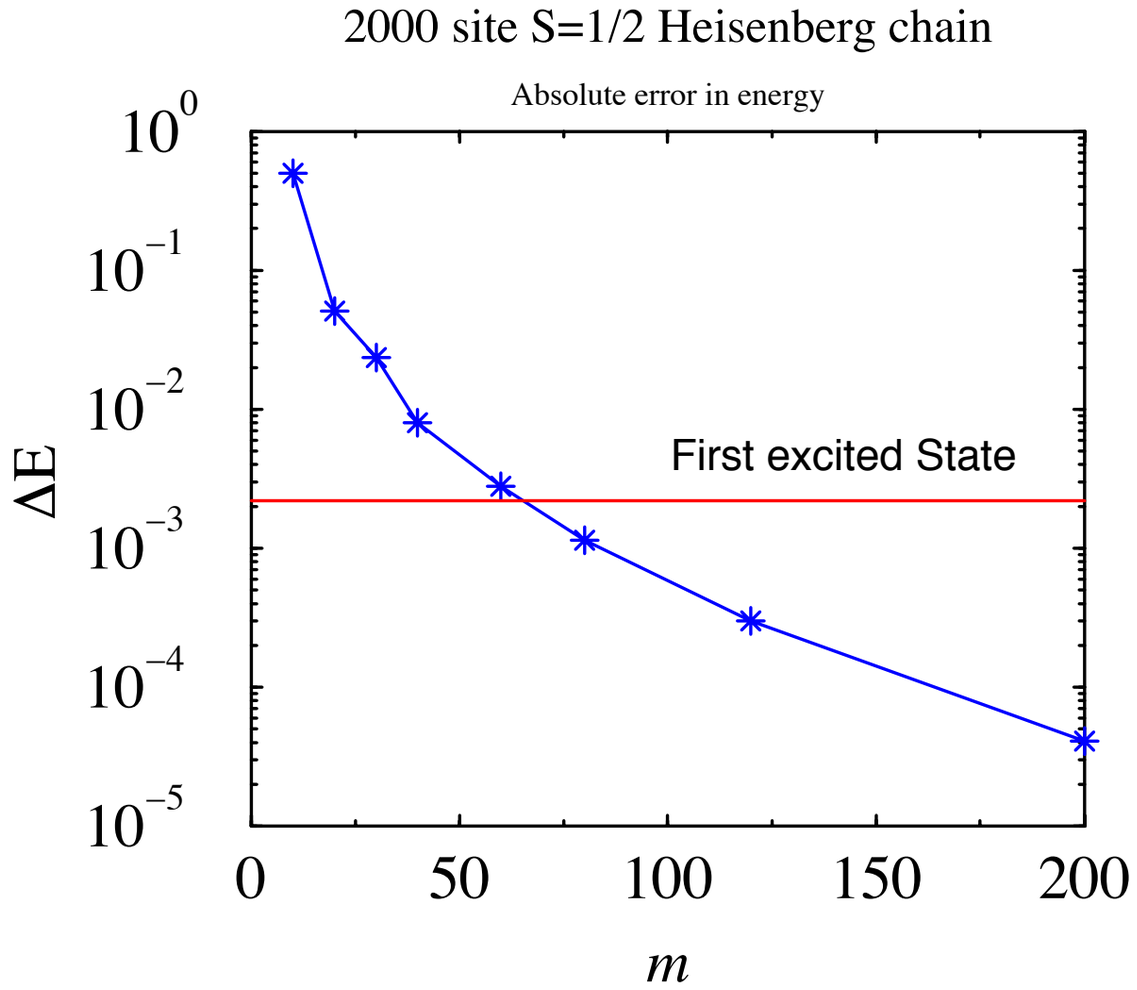
DMRG sweeps



•At each step, diagonalize approximate representation of entire system (in reduced basis)
•Construct density matrix for block, diagonalize it, keep most probable eigenstates (or SVD version)
•Transform / update operators to construct H
•Sweep back and forth, increase m

Convergence in 1D





Comparison with Bethe Ansatz

DMRG: two ways of thinking about it

- What I explained here: the MPS variational state point of view.
- The original view: Numerical RG; "Blocks" which have renormalized Hamiltonians (reduced bases) and operator-matrices in that basis
  - What is a block?
    - A block is a collection of sites (1 ... j), a matrix product basis for those sites, and the matrix representation of necessary operators in that basis.
    - We can think of a block as a renormalized system (doesn't have all its original d.o.f.) and the whole DMRG sweeping algorithm as a renormalization of the whole system (Wilson's orginal numerical RG).
    - Some things are easier to think about in each picture. DMRG practitioners should know *both pictures*!

DMRG: overview of extensions, generalizations, etc

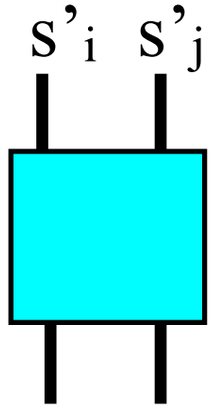
- Original two papers covered ground state energies and properties of 1D spin systems.
  - Applications to fermions and targeting several excited states was understood from the beginning and was quickly implemented.
- Application to ladder systems was also done very soon--the first steps towards 2D. Later I will cover recent 2D methods.
- Another area of strong development: dynamics. First work produced spectral functions (frequency, not time); later, work showed how to do real and imaginary time dynamics (Vidal).
- Classical Stat mech systems: developed early on; related to transfer matrices.
- Quantum chemistry: solving small molecules in a Gaussian basis. First work: White & Martin; now, most well known practitioner is Garnet Chan (Cornell --> Princeton).
- Lots of connections to quantum information--a major develpment I don't have time to do justice to.

Time Evolution (Vidal,...)

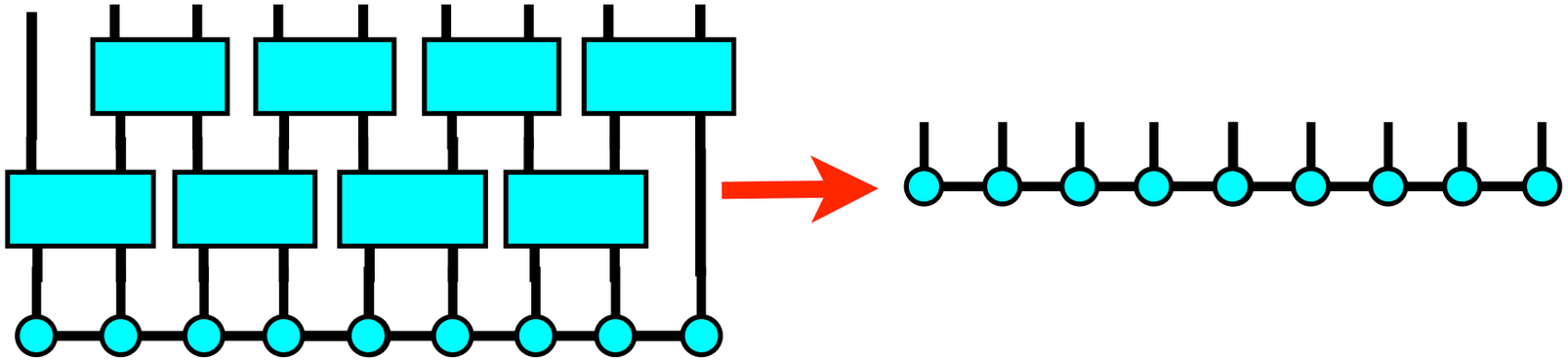
Suzuki Trotter decomposition:

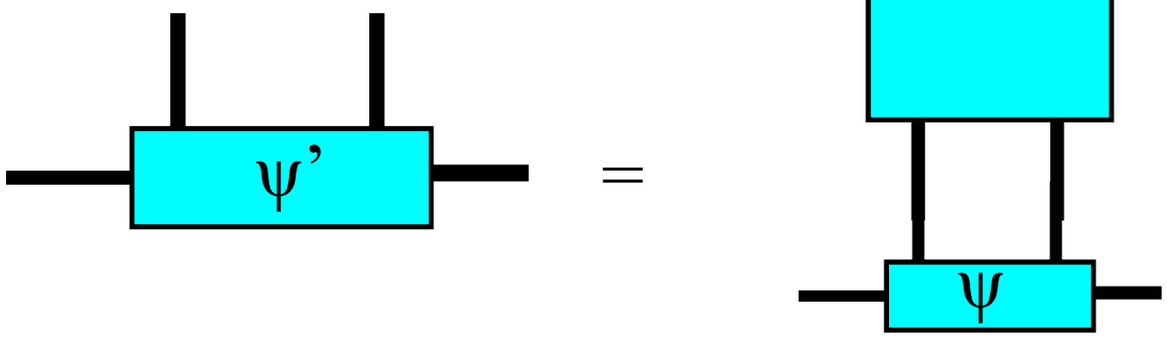
$$\exp(-iH\tau) \approx \exp(-iH_{12}\tau)\exp(-iH_{34}\tau) \ldots \exp(-iH_{23}\tau) \ldots$$

$\exp(-iH_{ij}\tau)$ =



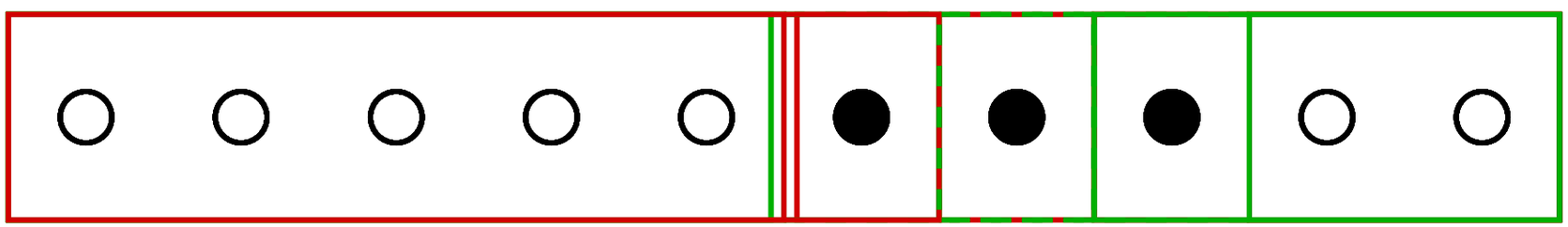
In DMRG, the bond operator for the current middle two sites is trivial to apply:





DMRG Sweeps

- Finite system method:



- During each step, instead of finding the ground state, we can apply $T_{ij} = \exp(-i H_{ij} \tau)$ (or leave ψ alone).
- When to apply T's: several versions:
  - Standard even/odd breakup:
    - 1 --- 2 --- 3 --- 4 do odd bonds in left-to-right half sweep
    - ---7 --- 6 ---5 --- do evens in right to left half sweep
  - White-Feiguin version:
    - 1 2 3 4 5 6 7 do all bonds in each half sweep
    - 14 13 12 11 10 9 8 reverse order each half sweep

Calculation of Spectral functions

- Start with standard ground state DMRG, get φ
- Apply operator to center site

$$|\psi(t=0)\rangle = S_0^+ |\phi_0\rangle$$

- Time evolve:

$$|\psi(t)\rangle = e^{-i(H-E_0)t}|\psi(0)\rangle$$

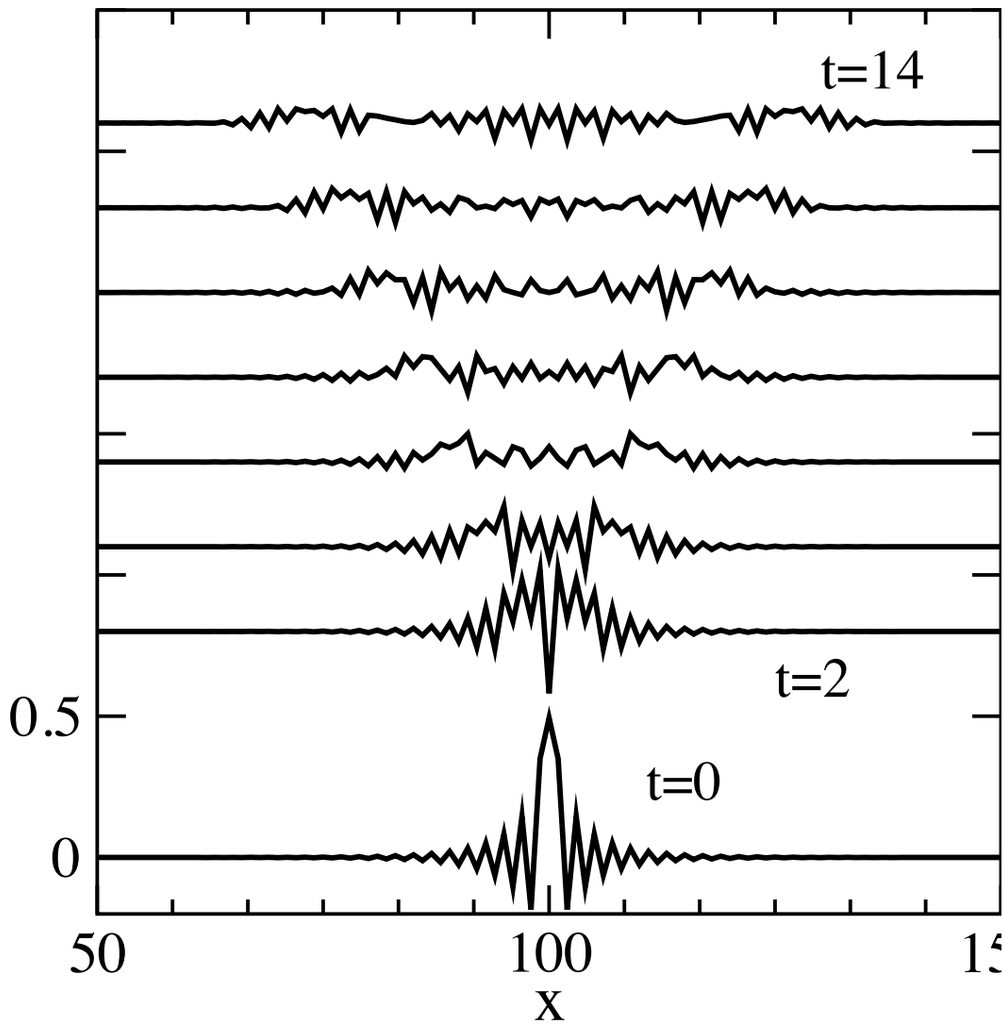
- Measure time dependent correlation function

$$G(x,t) = \langle\phi_0|S_x^-|\psi(t)\rangle = \langle\phi_0|S_x^-(t)S_0^+(0)|\phi_0\rangle$$

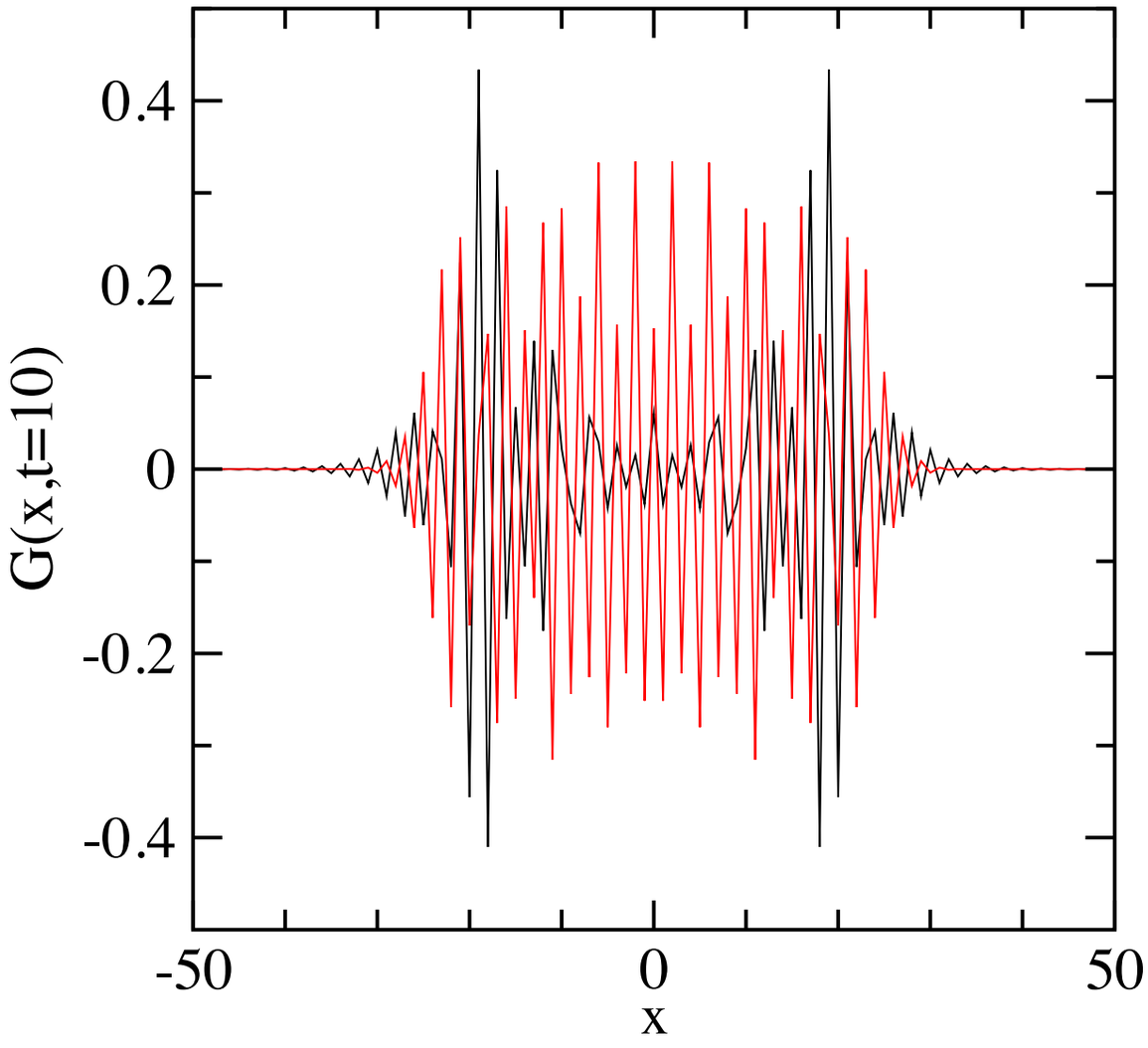
- Fourier transform with x=0 to get N(ω) or in x and t to get S(k,ω)
  - But what about finite size effects, finite time, broadening, etc??

Finite size effects: gapped systems

S=1 Heis chain



Real and Imag parts



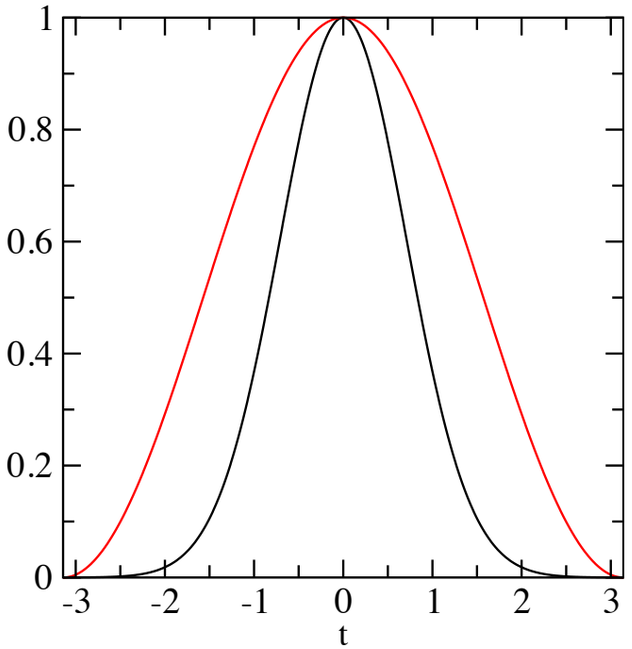
For t < L/(2 v), finite size effects are negligible.

Growth of entanglement with time

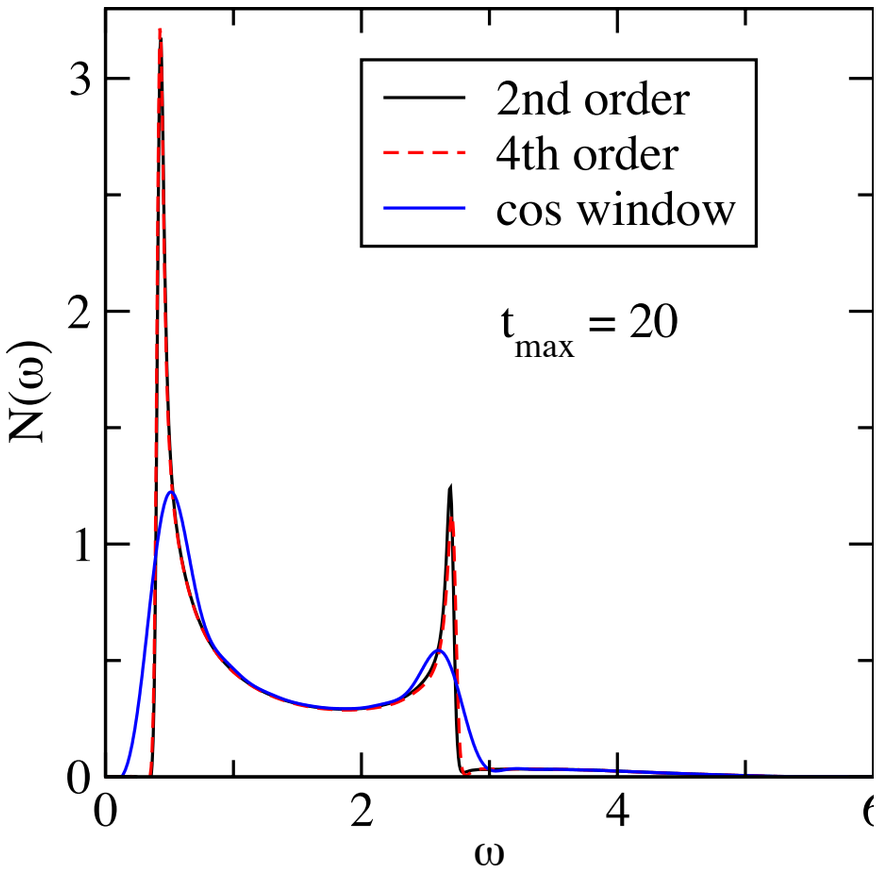
- Lots of work on growth of entanglement with time--sounds very discouraging at first
  - For "macro" changes to the wavefunction, S grows linearly (e.g. suddenly change the Hamiltonian). Then matrix dimension m must grow exponentially (and effort ~ m^3)
- Fortunately, for what we need here, one local change, growth is only logarithmic!
  - Still limited in total time we can simulate--still the key issue
- Example: for spin chains, we can go out to $t_{max} \sim 30$.
- It appears that one should be limited in frequency resolution to $\sim 1/t_{max}$
- But: the long time behavior is determined almost completely be the singularities in $A(\omega)$, and if there are just a few, we can fit them and get extremely high resolution (SRW, Affleck, Pereira)

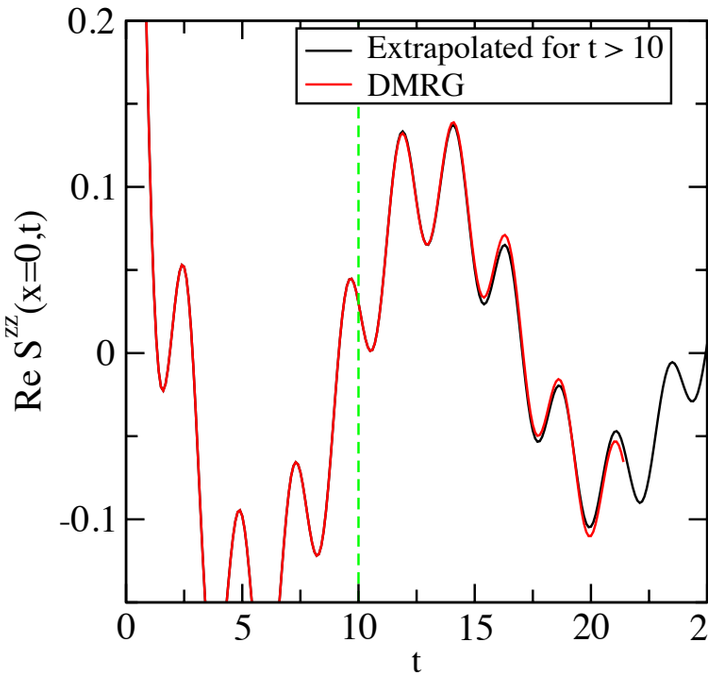
Extrapolation to large time: linear prediction

S=1 chain



Windows for time FT



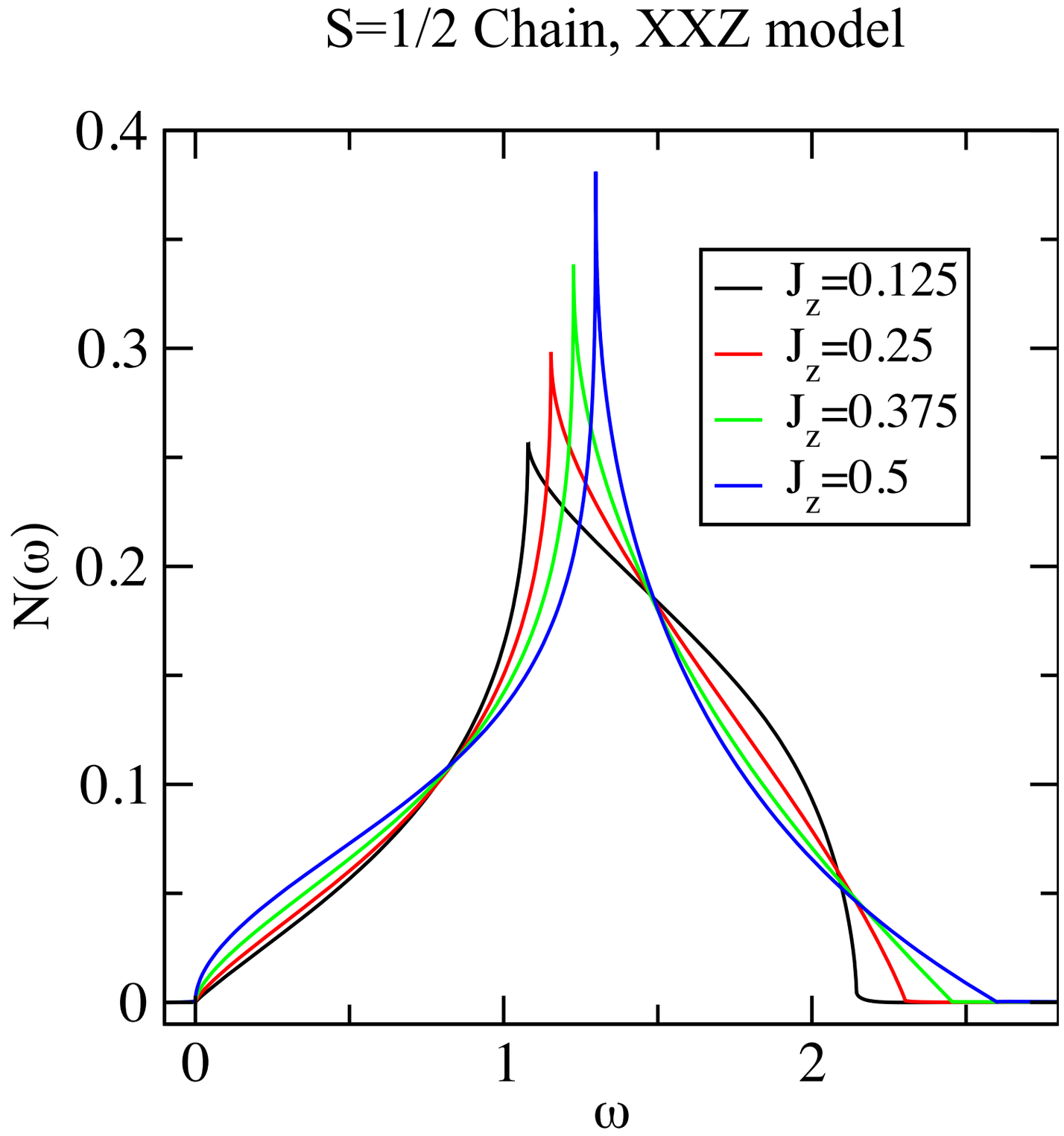


$$y_i = \sum_{j=1}^{n} d_j y_{i-j}$$

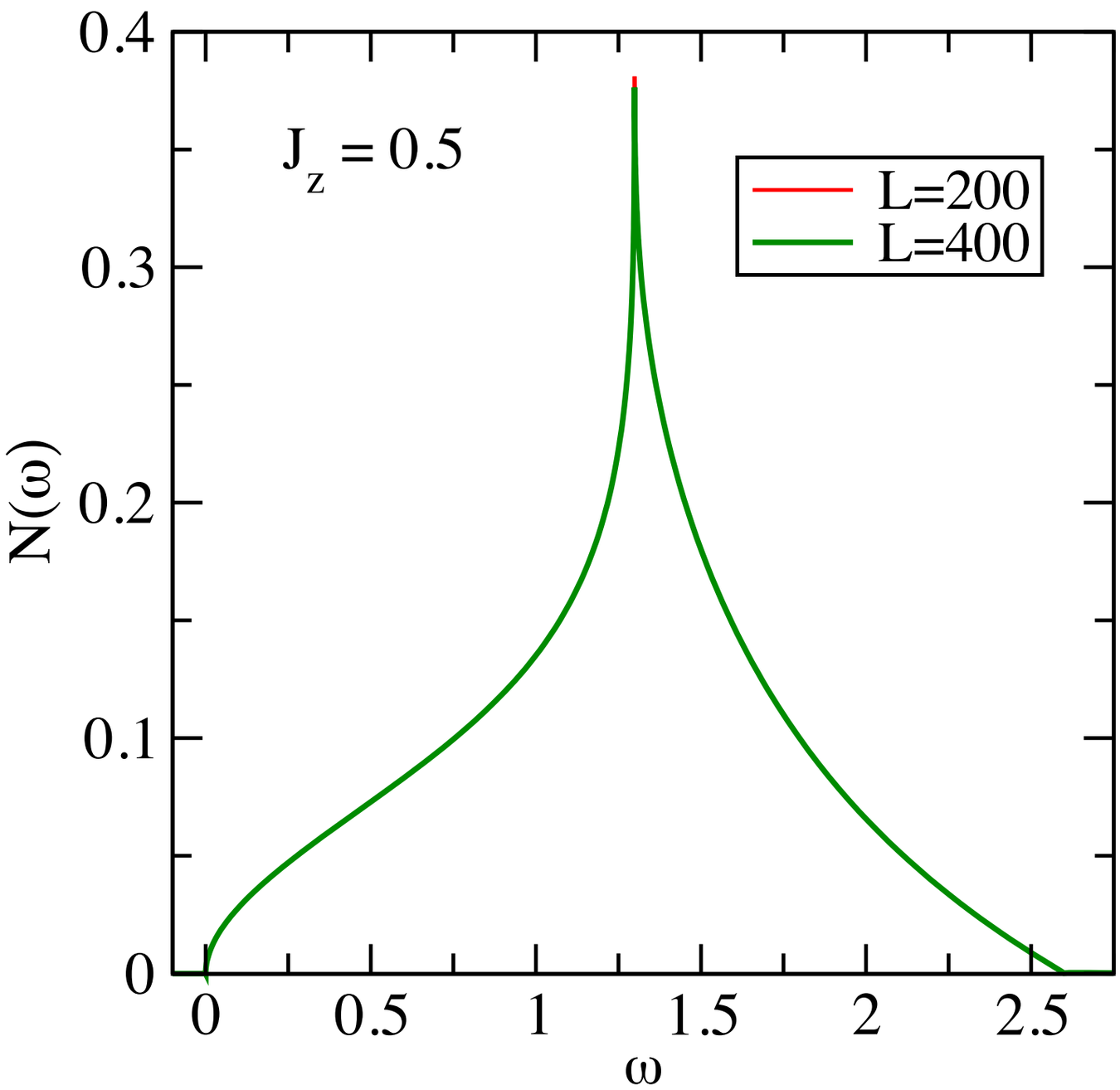
See *Numerical Recipes*

Parameters d_j determined from correlation functions of available data.

Example where singularities were fit to and used to extrapolate (less automatic than linear prediction)

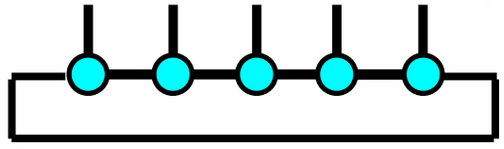


How accurate are the spectra?



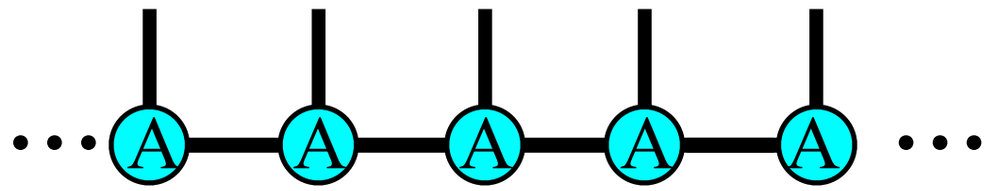
Generalizations of MPS

- Periodic BCs: long a weakness of DMRG ($m \rightarrow m^2$)
  - New variational state:

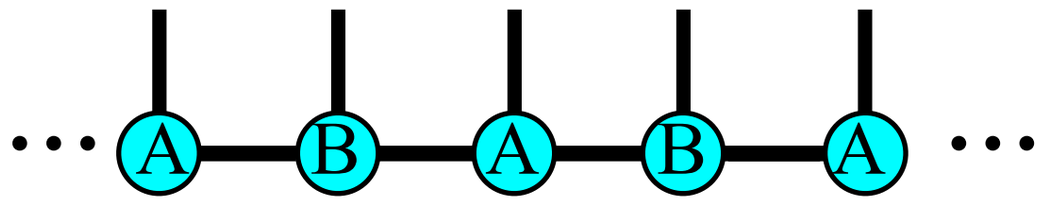


  - key issue is computational: optimization to minimize E
    - Ostlund and Rommer ('95) m = 12
    - Verstraete, Porras, Cirac (PRL 93, 227205 '04) calc time ~m^5
    - Pippan, White, Evertz (PRB 82, 024407 (2010)) calc time ~ m^3

- Infinite systems
  - Natural state:



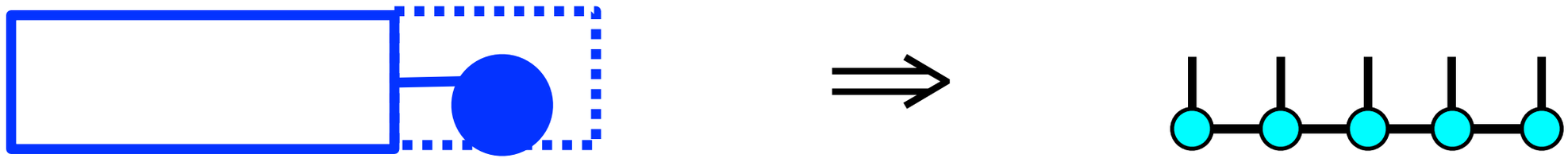
    - But: very hard to optimize A
  - Much better:



Vidal, PRL 98, 070201 (2007)

    - Trotter imaginary time evolution: odd links, then even, repeat
      - iTEBD (infinite time evolving block decimation)

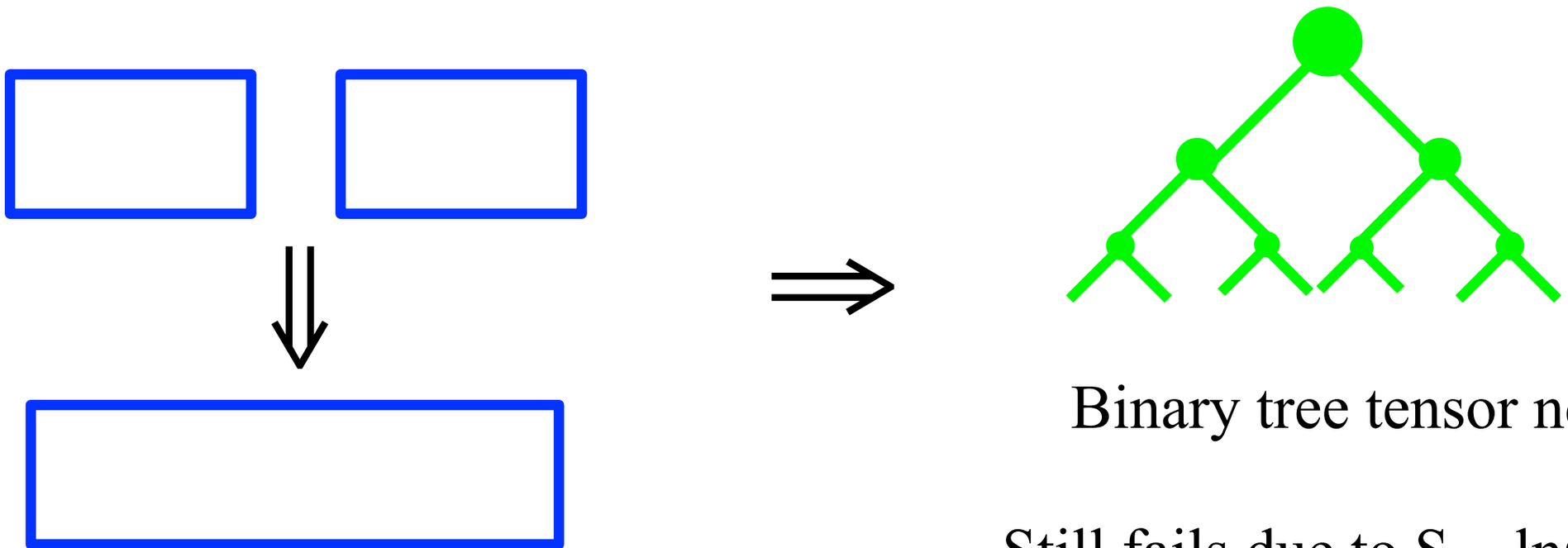
Critical 1D systems



DMRG/NRG MPS

Critical systems:

1) S ~ ln(L), so MPS eventually fails
2) MPS does not exhibit scale invariance naturally

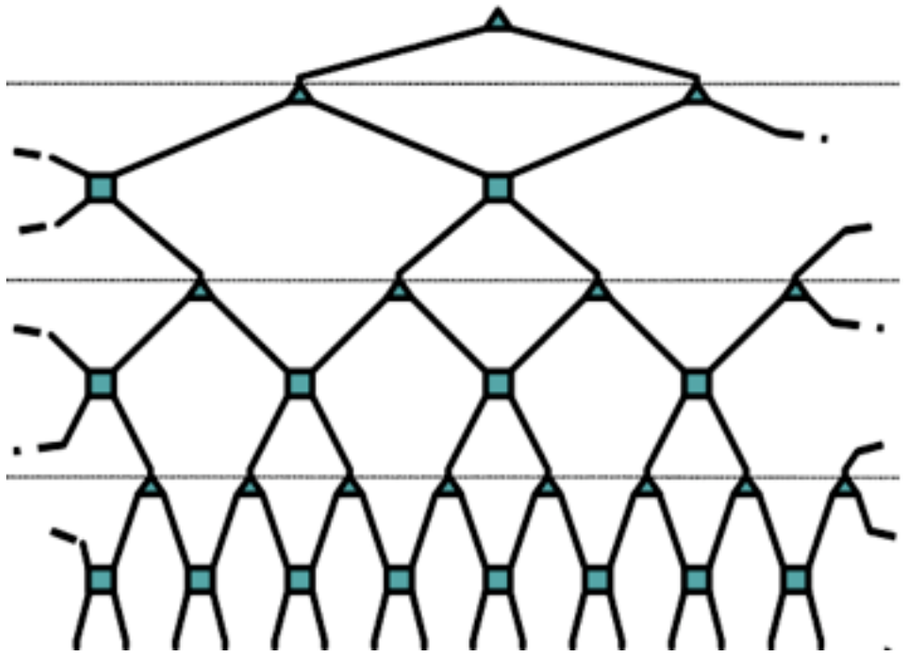


Real space RG

Binary tree tensor network

Still fails due to S ~ ln(L) !!

Tensor networks for 1D critical systems



Multiscale entanglement renormalization ansatz (MERA)

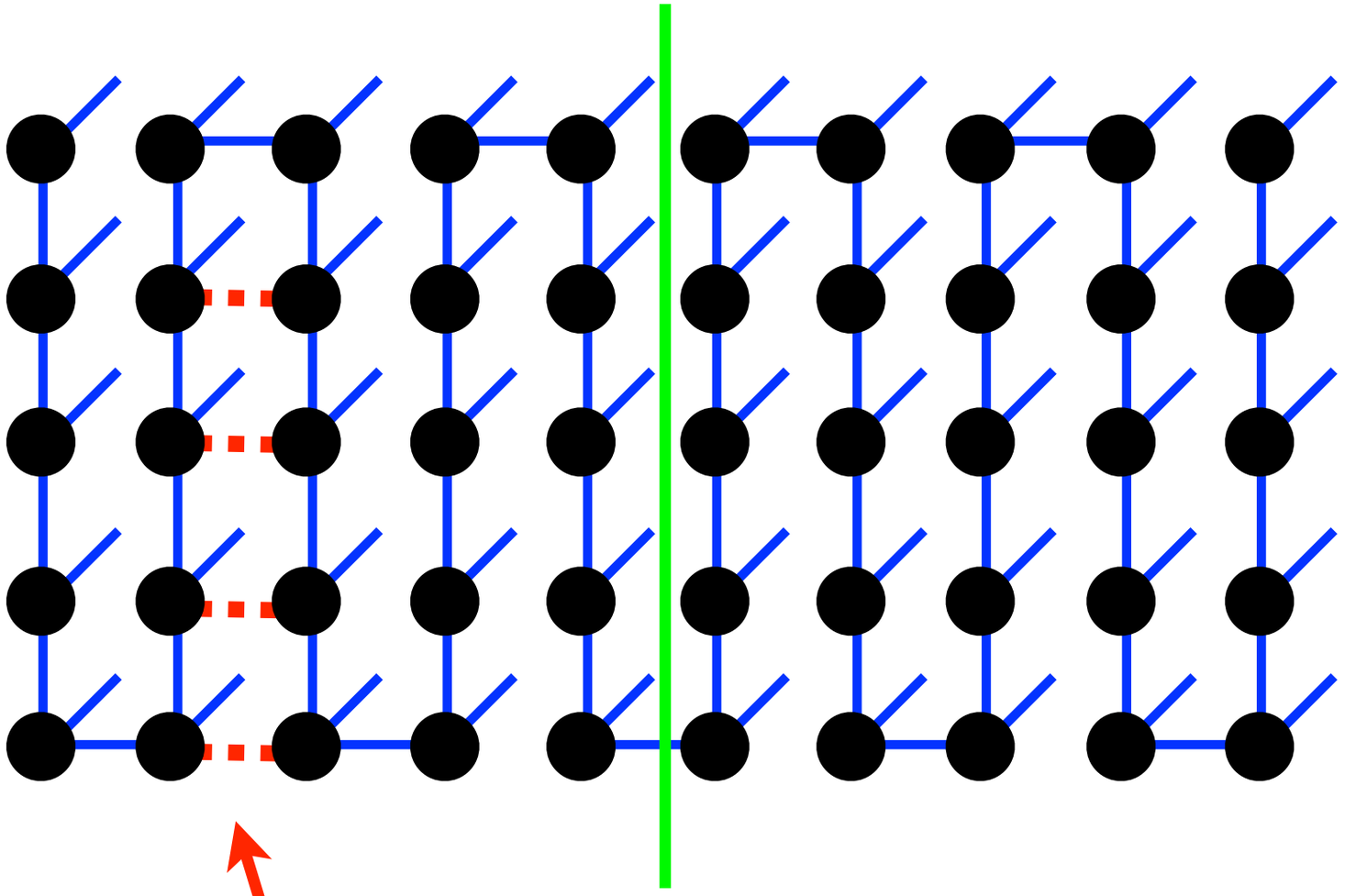
Vidal, PRL 99, 220405 (2007)

Rizzi, Montagero, Vidal PRA 77, 052328 (2008) (tMERA)
Evenbly & Vidal, arxiv:0707.1454
Pfeifer, Evenbly,&Vidal, arxiv: 0810.0580

- Entanglement gets organized at different length scales at different layers: RG
- At criticality, expect translational/scale invariance in both directions! Compression: superb
- Computation time: m^9 L ln L, or m^9 ln L for translational inv. systems, but m=6 has energy errors ~ 10^{-7} (Critical transverse field Ising model)
- State directly yields CFT central charge, scaling dims of primary fields
- Accurate correlations at large distance, e.g. r = 10^9 !!

2D algorithms

- Traditional DMRG method (MPS state)



Cut

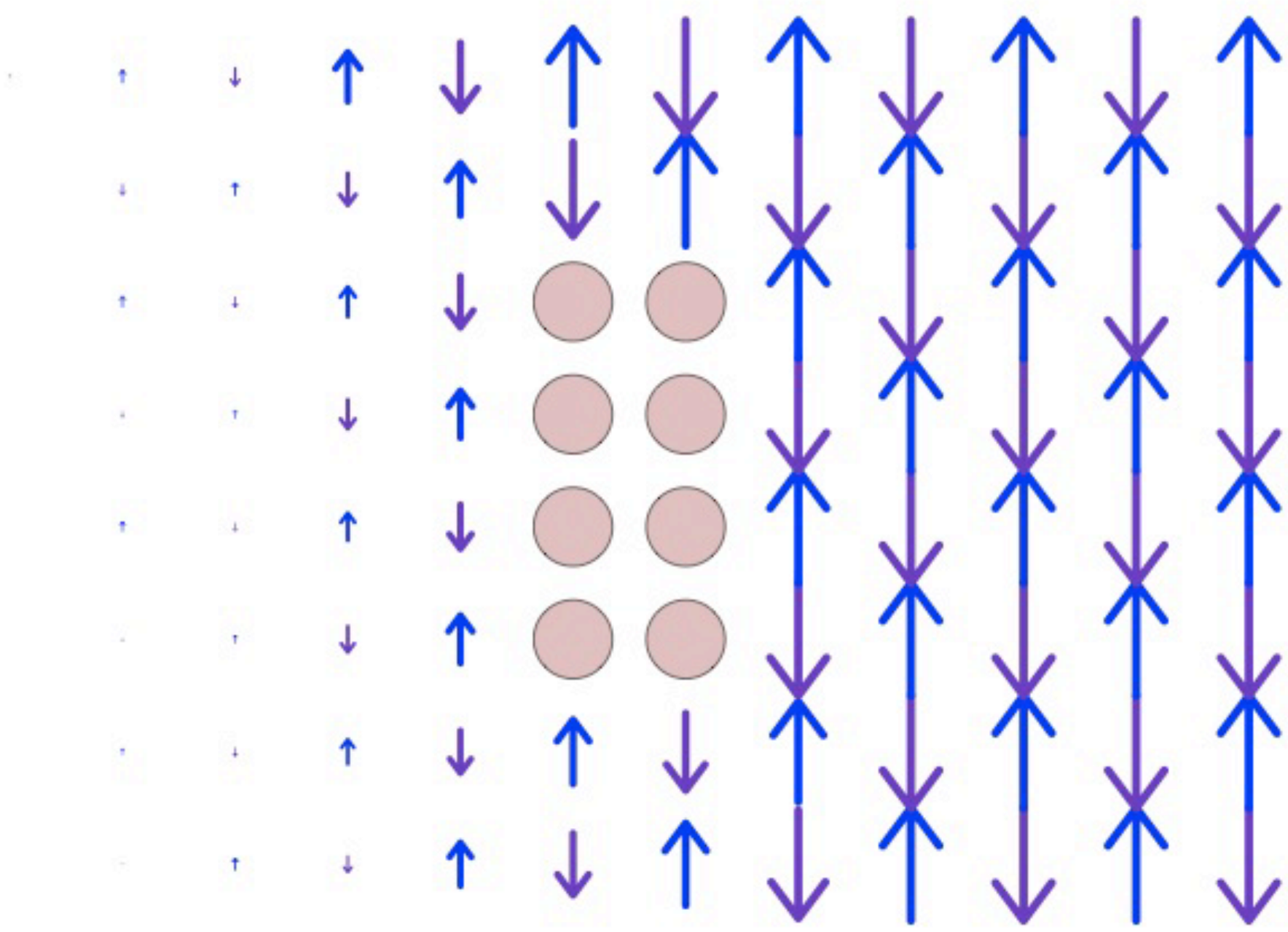
Long range bonds

$S \sim L_y$ ("area law")

$m \sim \exp(a\ L_y)$

Calc time: $L_x\ L_y^2\ m^3$; allows $m \sim 10000$, $L_y \sim 10\text{-}12$

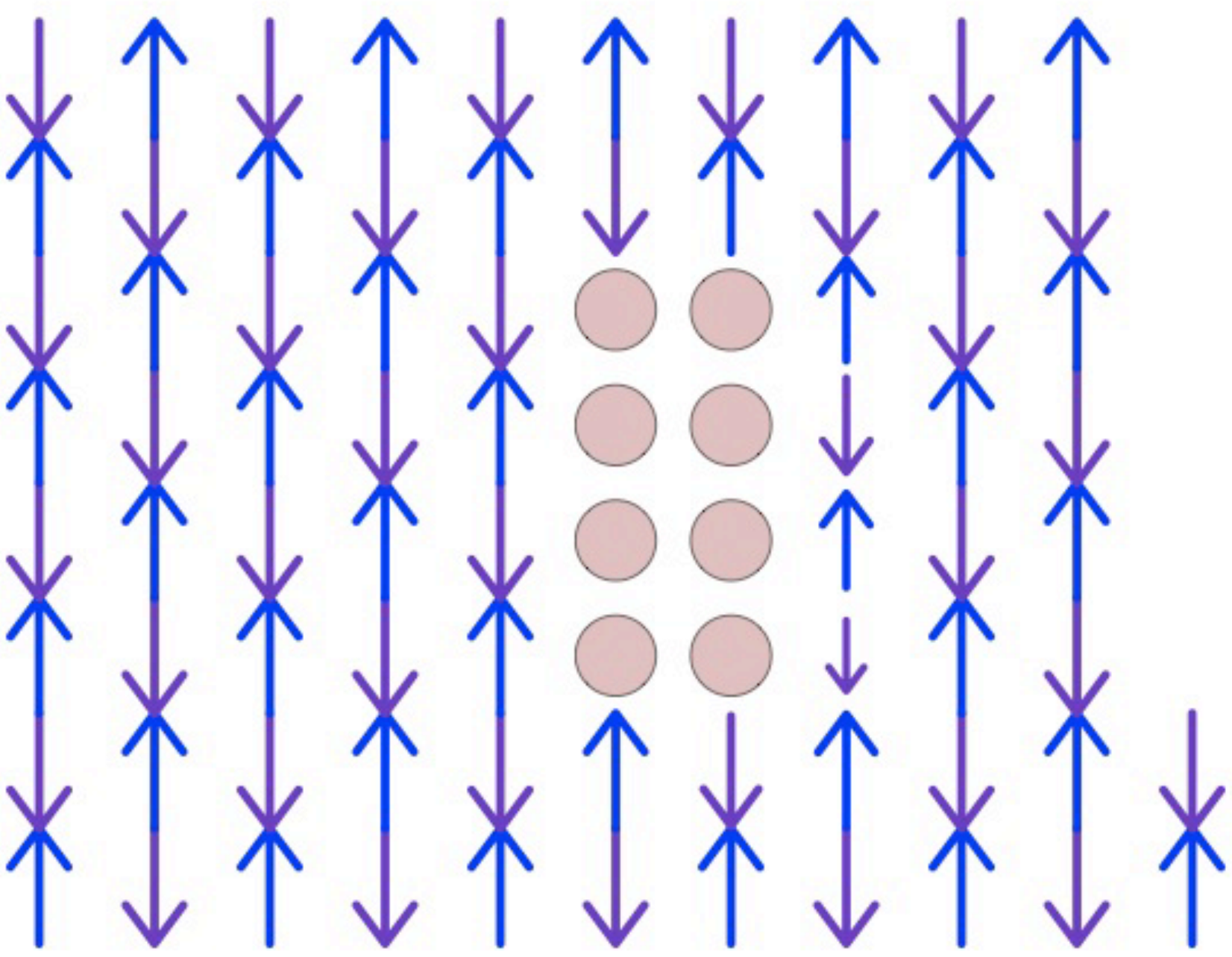
Stripes forming from a blob of 8 holes



12x8
Cylindrical BCs
t=1, J=0.35
t'=t''=0
8 holes
No pinning
fields

E = -31.7287 m = 70

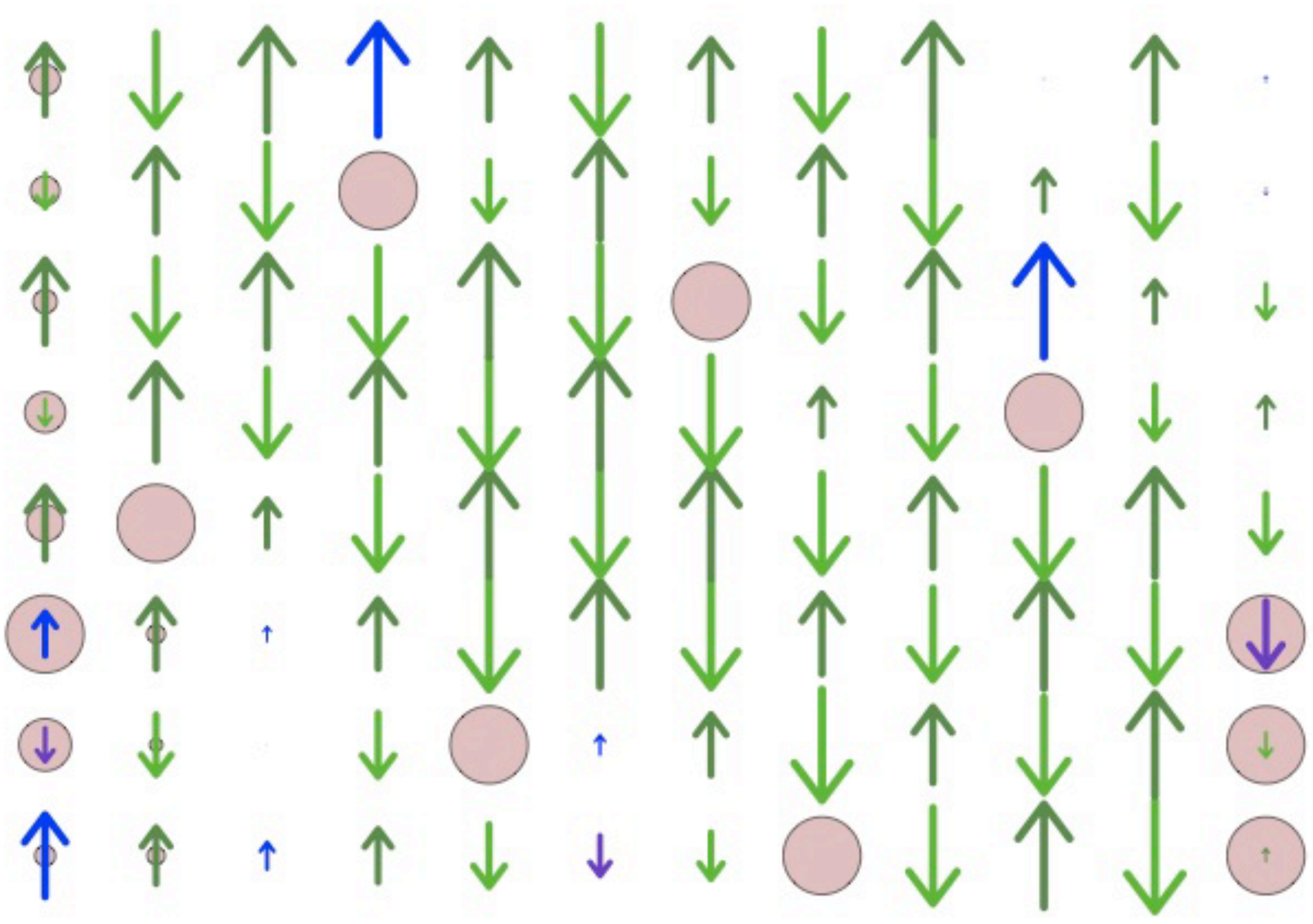
Stripes forming from a blob of 8 holes



E = -30.7350 m = 40

12x8
Cylindrical BCs
t=1, J=0.35
t'=t''=0
8 holes
AF edge
pinning fields applied for two sweeps to favor one stripe

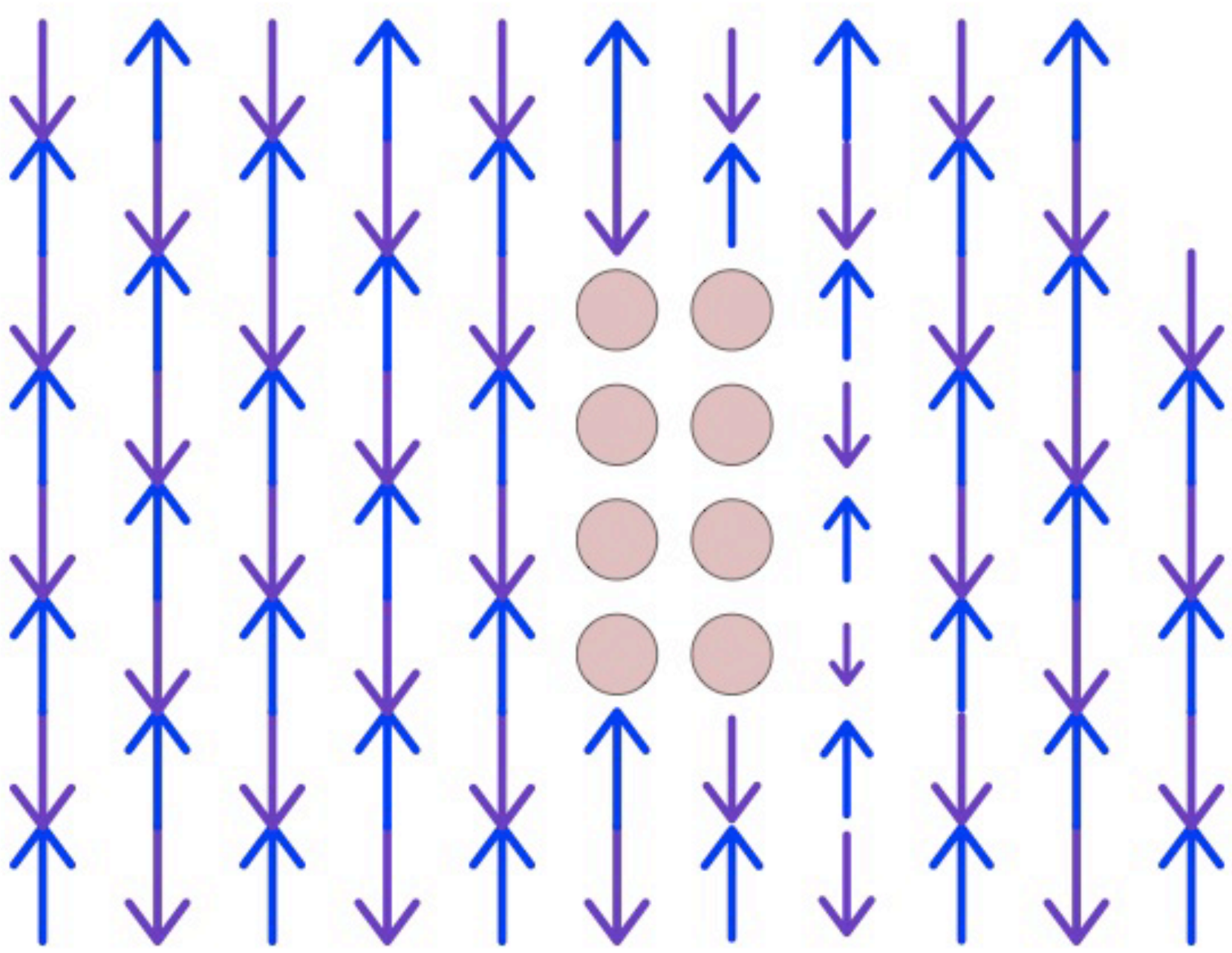
Stripes not forming from a bad initial



E = -30.6370 m = 70

12x8
Cylindrical BCs
t=1, J=0.35
t'=t''=0
8 holes
No pinning fields.
Initial state has holes spread out so favored striped state is hard to find.
Energy higher by ~0.3 t.

Curved Stripe forms due to open BCs



E = -30.8532

m = 40

12x8

Open BCs

t=1, J=0.35

t'=t''=0

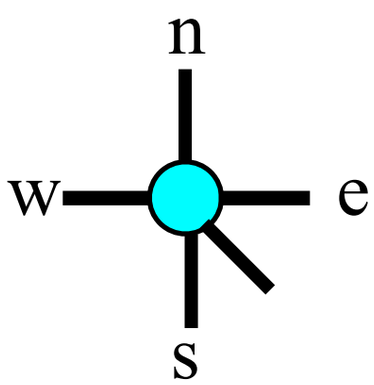
8 holes

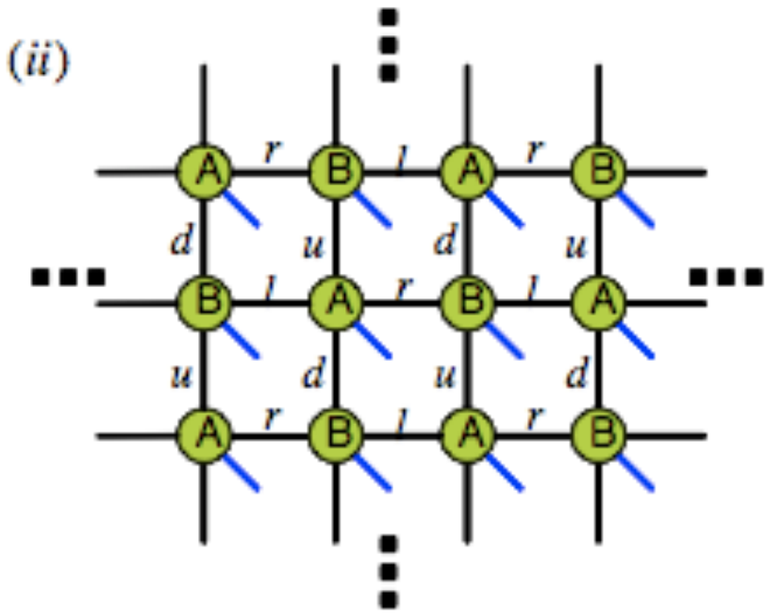
No pinning fields

Projected entangled pair states

(Nishio, Maeshima, Gendiar, and Nishino, cond-mat/0401115; Verstraete and Cirac, condmat 0407066)

- Generalize the 1D MPS ansatz to 2D:





Basic unit is 5 index tensor, blue/down index is state of a site

PEPS

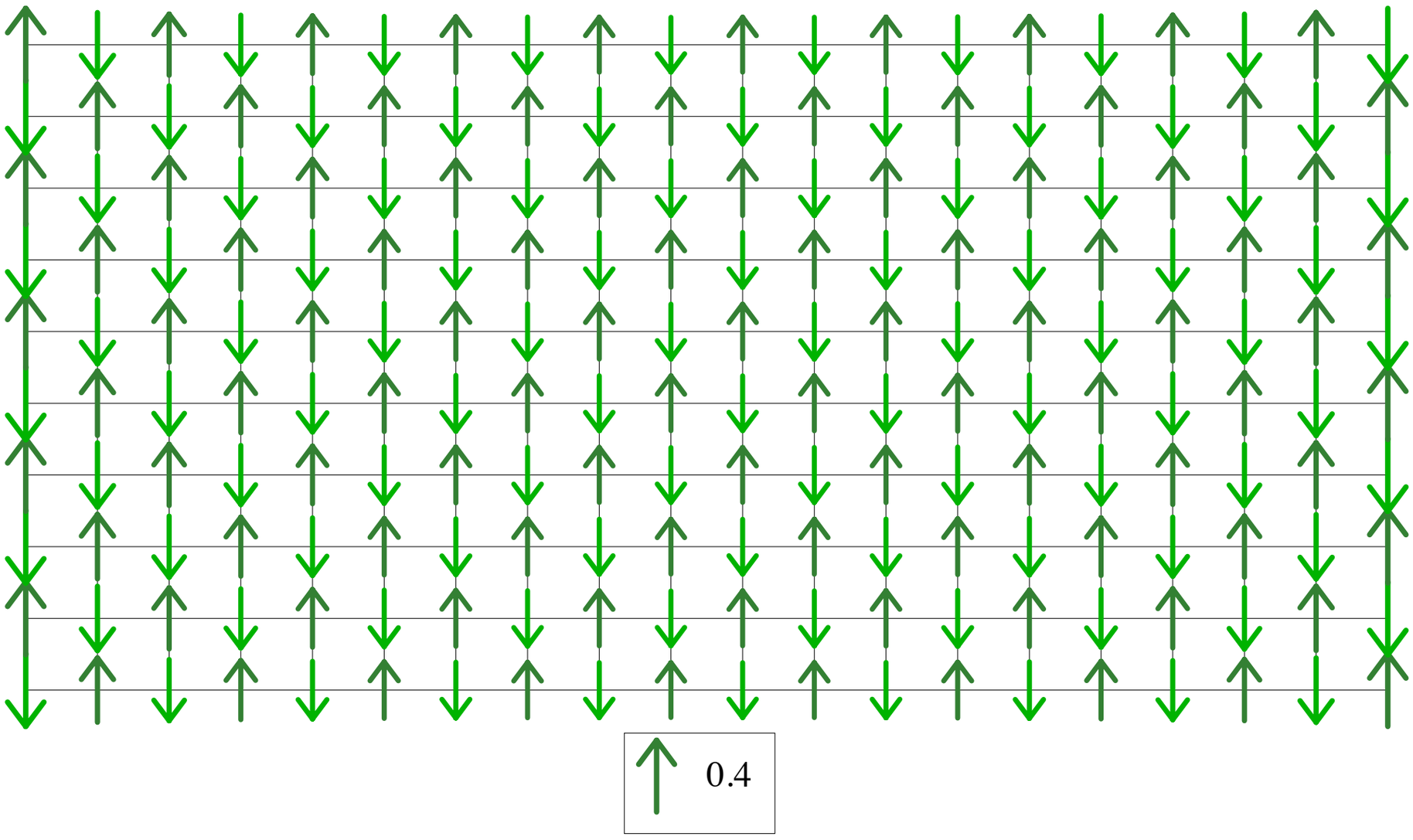
  - Much more natural representation!
  - Key issues: optimization, contraction
  - V&C approach: CPU ~ $L_x L_y m^{10}$ No exponentials!!

- MERA is another tensor network approach to 2D with similar properties
- Fermionic PEPS: simple treatment of fermionic exchange

Some Practical aspects of DMRG for hard systems and Applications to 2D

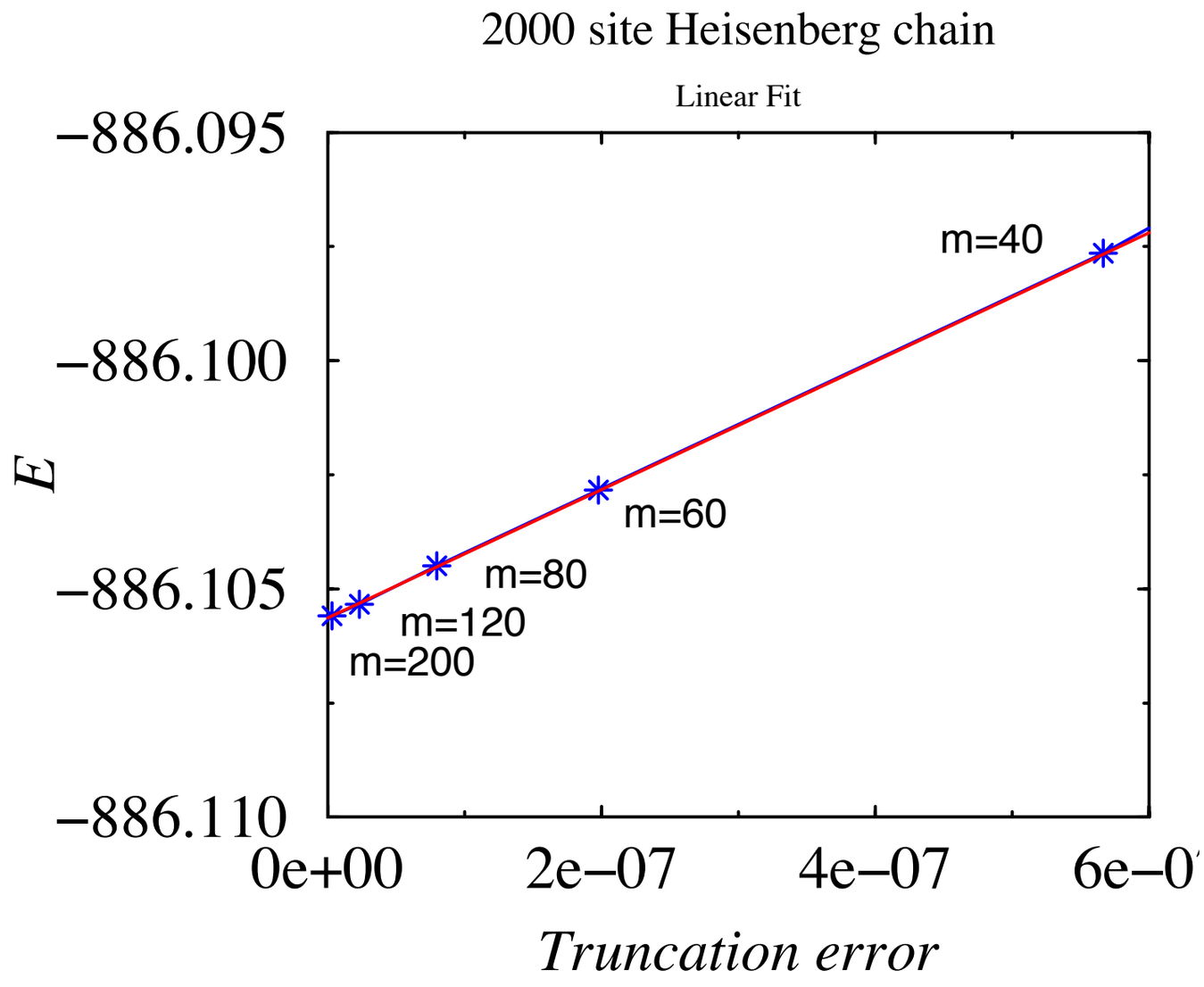
- Extrapolation in truncation error for energy and observables
- Tips for very efficient calculations
- Some results for 2D Heisenberg models:
  - Square lattice
  - Triangular lattice
  - Kagome lattice

Square lattice: benchmark against



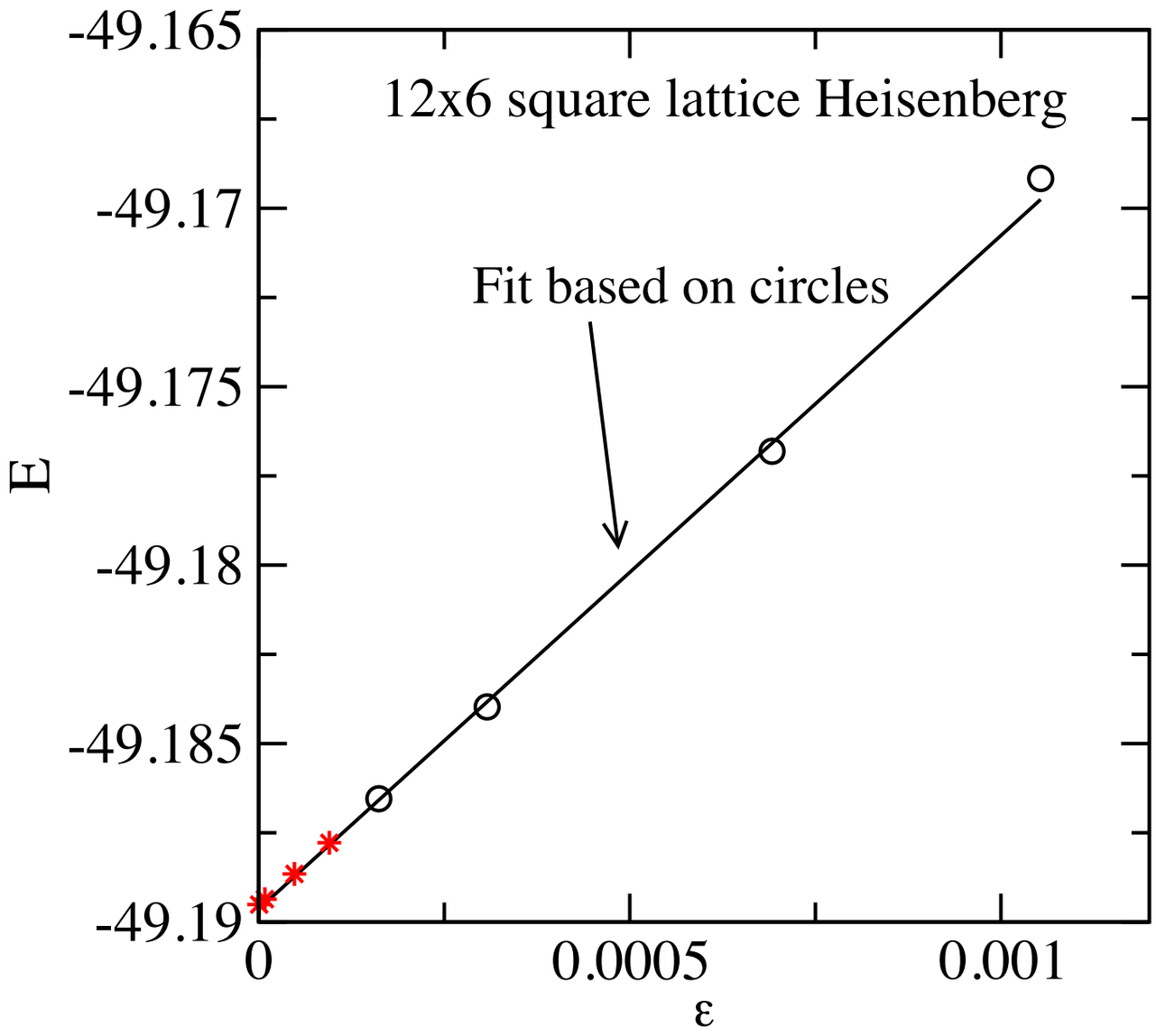
- Cylindrical BCs: periodic in y, open in x
- Strong AF pinning fields on left and right edges
- 21 sweeps, up to m=3200 states, 80 hours

Extrapolation of the energy



Extrapolation improves the energy by a factor of 5-10 and provides an error estimate.

Energy extrapolation



Probability of states thrown away
= truncation error (function of m)

Assign error bars to result: if the fit is this good, assign (extrapolation from last point)/5

(no derivation, just experience that this works on lots of systems)

If the fit looks worse, increase the error bar (substantially) or don't use that run/keep more states or smaller size system.

Extrapolation of local observables(ref: White and Chernyshev, PRL 99, 127004 (2007))

- Standard result for a variational state

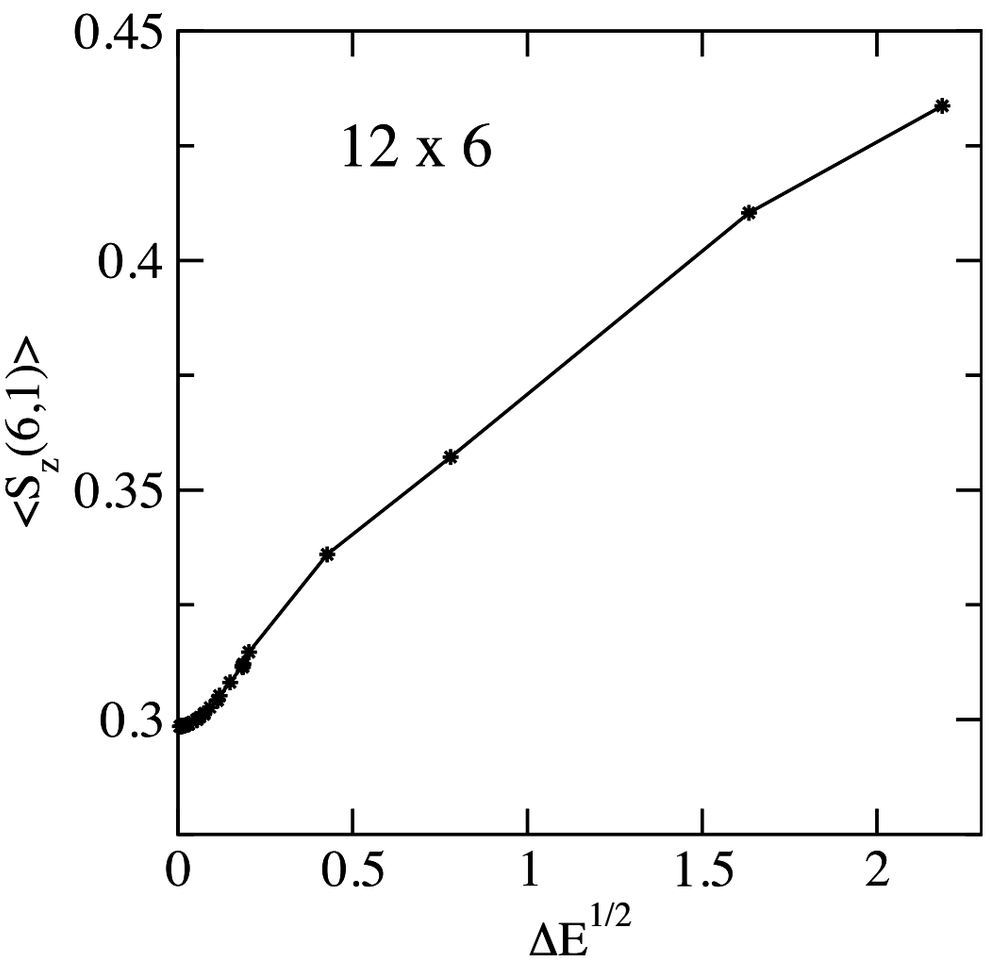
$$|\psi\rangle = |G\rangle + |\delta\rangle, \quad \langle G|\delta\rangle = 0,$$

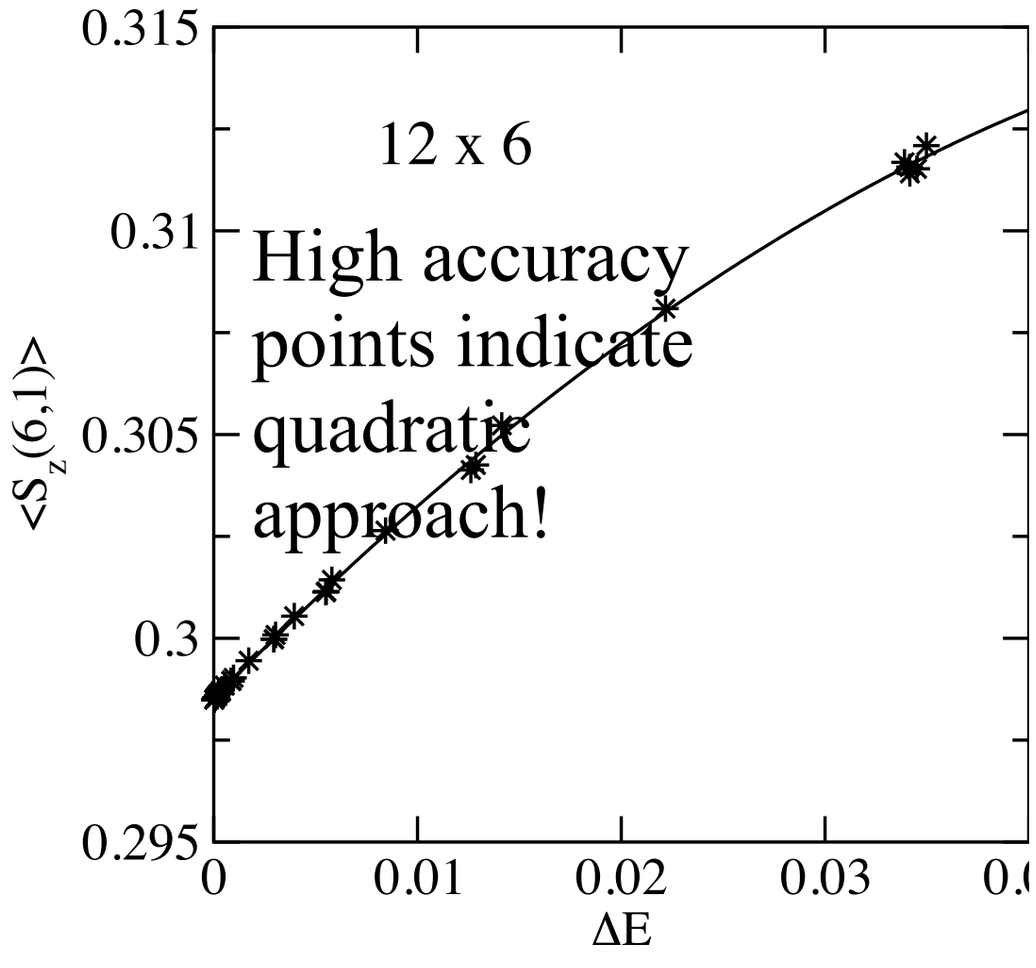
$$A = (1 + \langle\delta|\delta\rangle)^{-1}(A_G + 2\langle G|\hat{A}|\delta\rangle + \langle\delta|\hat{A}|\delta\rangle)$$

$$E = (1 + \langle\delta|\delta\rangle)^{-1}(E_G + \langle\delta|\hat{H}|\delta\rangle)$$

- Consequences:
  - Variational calculations can have excellent energies but poor properties
  - Since DMRG truncation error $\varepsilon \sim \langle\delta|\delta\rangle$, $E \sim \varepsilon$, but otherwise extrapolations vary as $A \sim \varepsilon^{1/2}$
- These $\varepsilon^{1/2}$ extrapolations have never worked well.

Typical extrapolation of magnetization

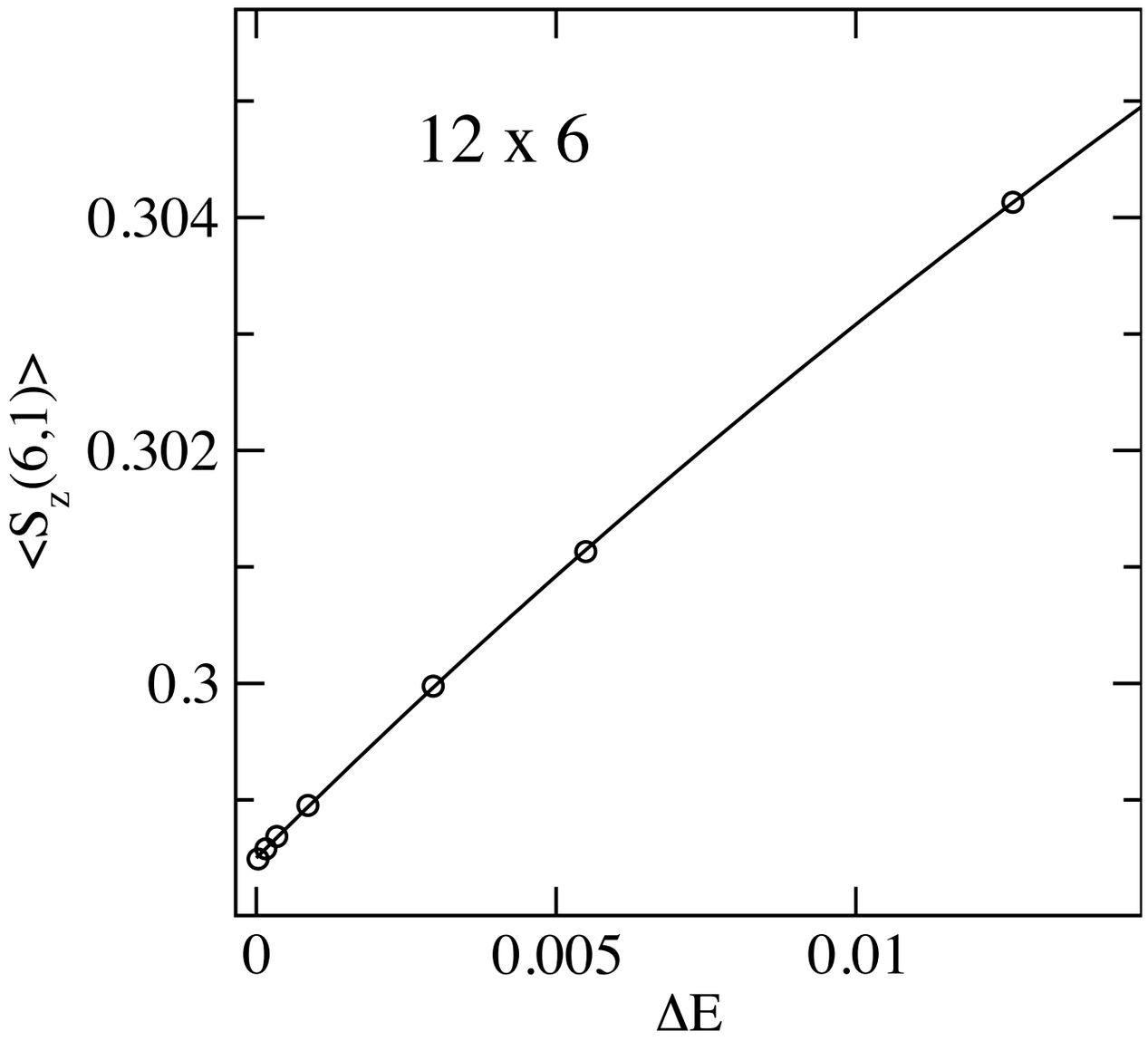




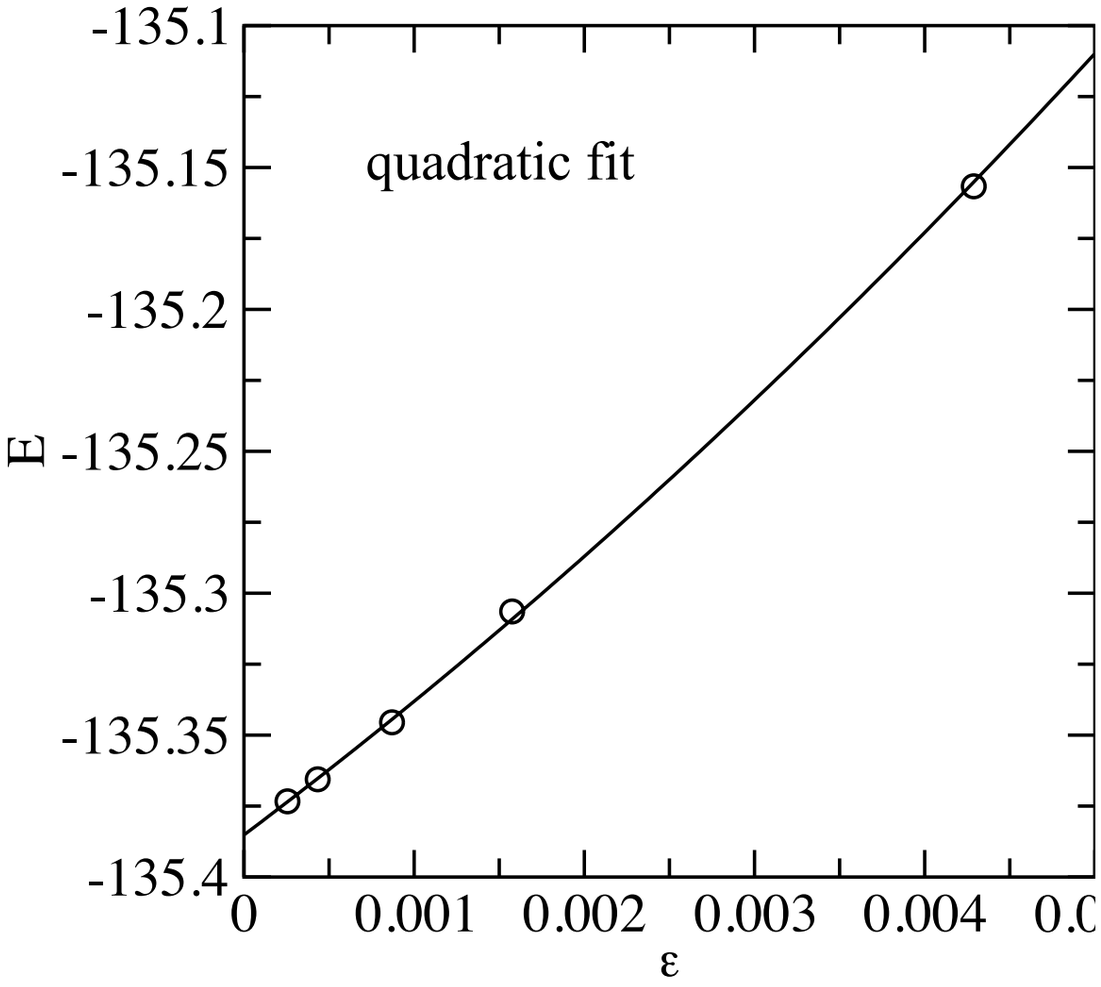
Pinning AF fields applied to edges, cylindrical BCs

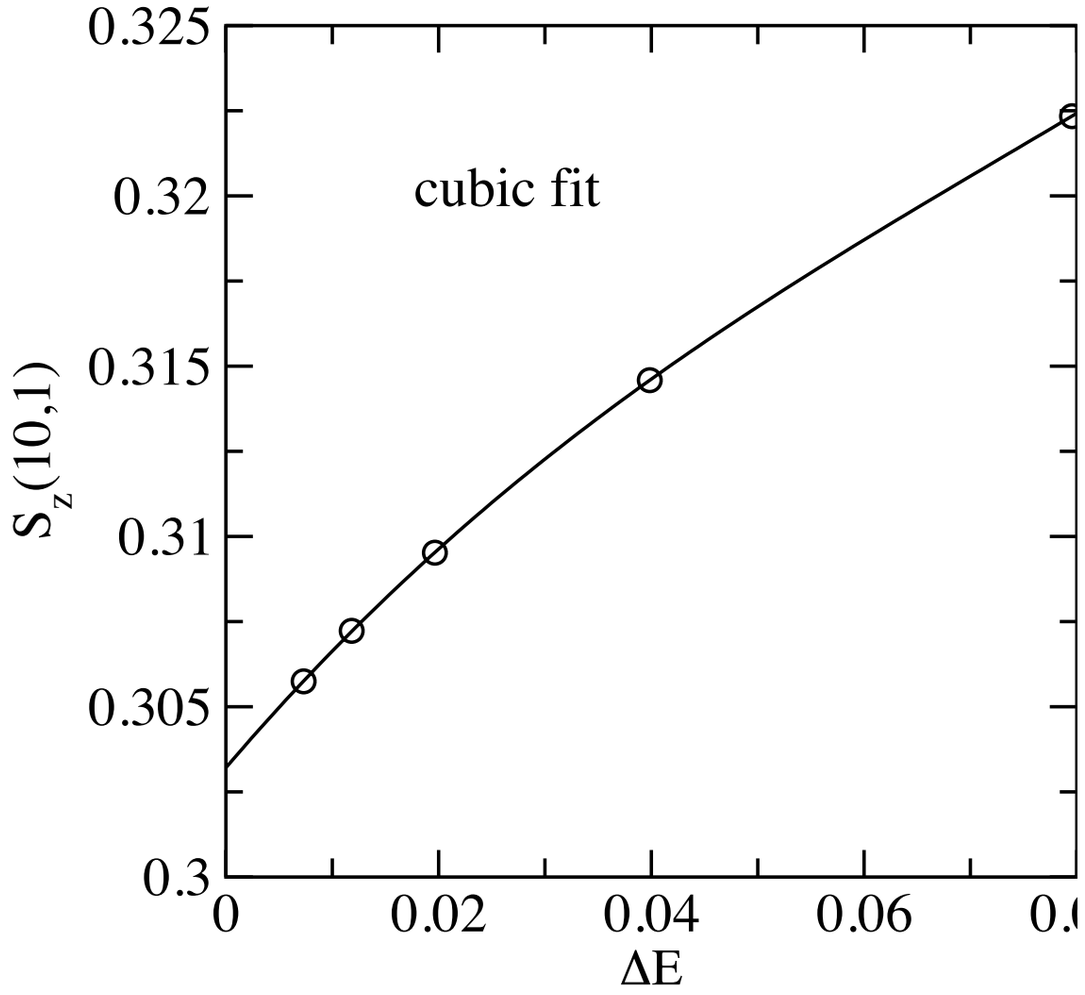
Now we understand why the local measurements converge fast; see White & Chernyshev

Cubic fit to well-converged measurements



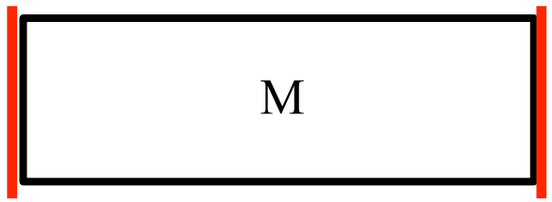
20x10 square lattice Heisenberg



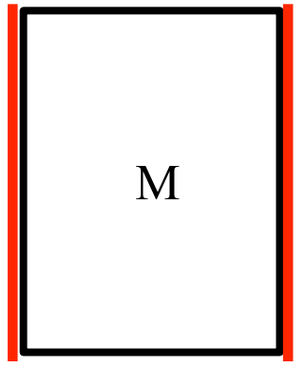


Result: central M = 0.3032(9)

Improved finite size scaling: choosing aspect ratios to reduce finite size effects



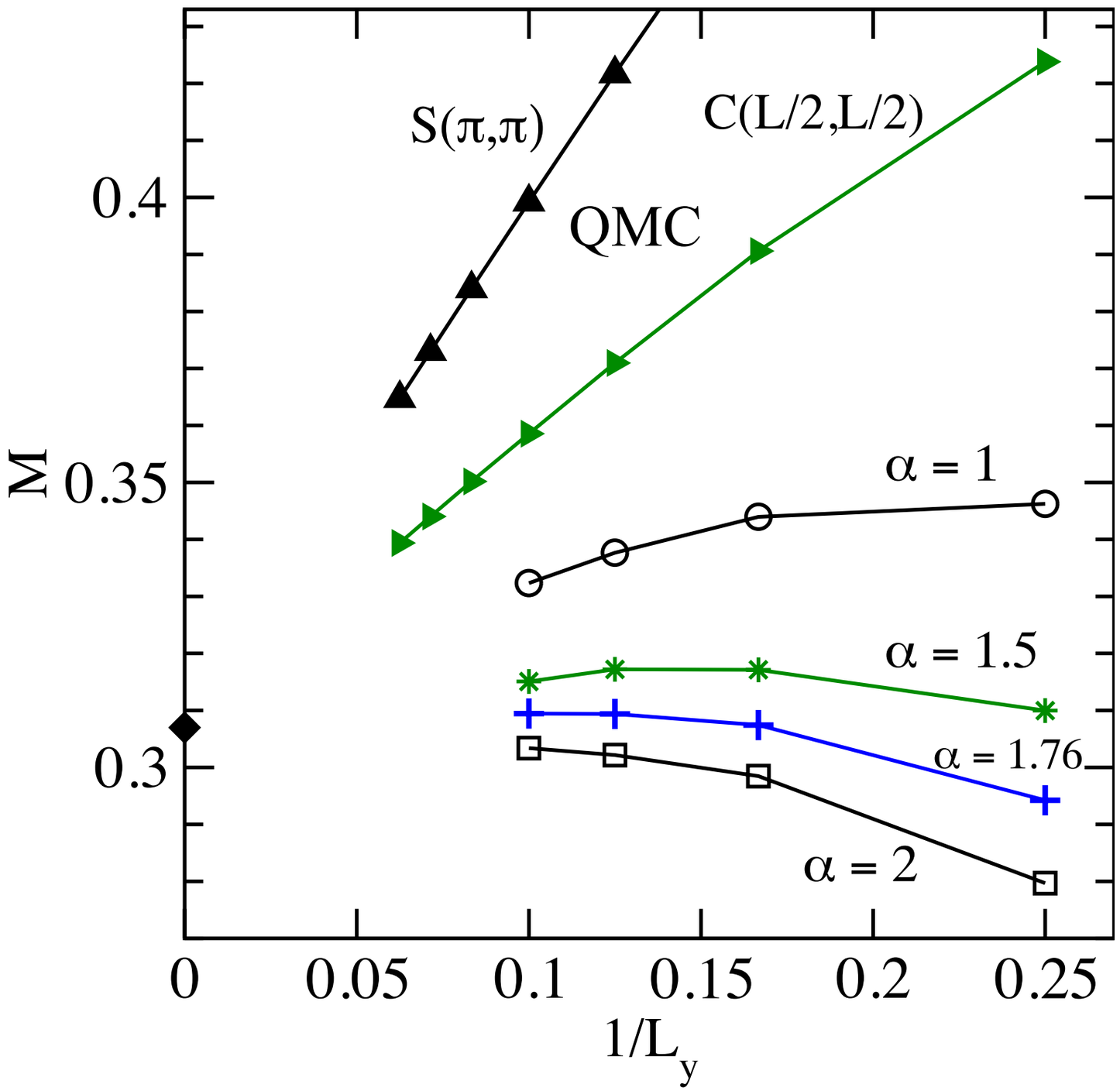
Long: 1D makes M small



Short: proximity to strong pinning makes M large

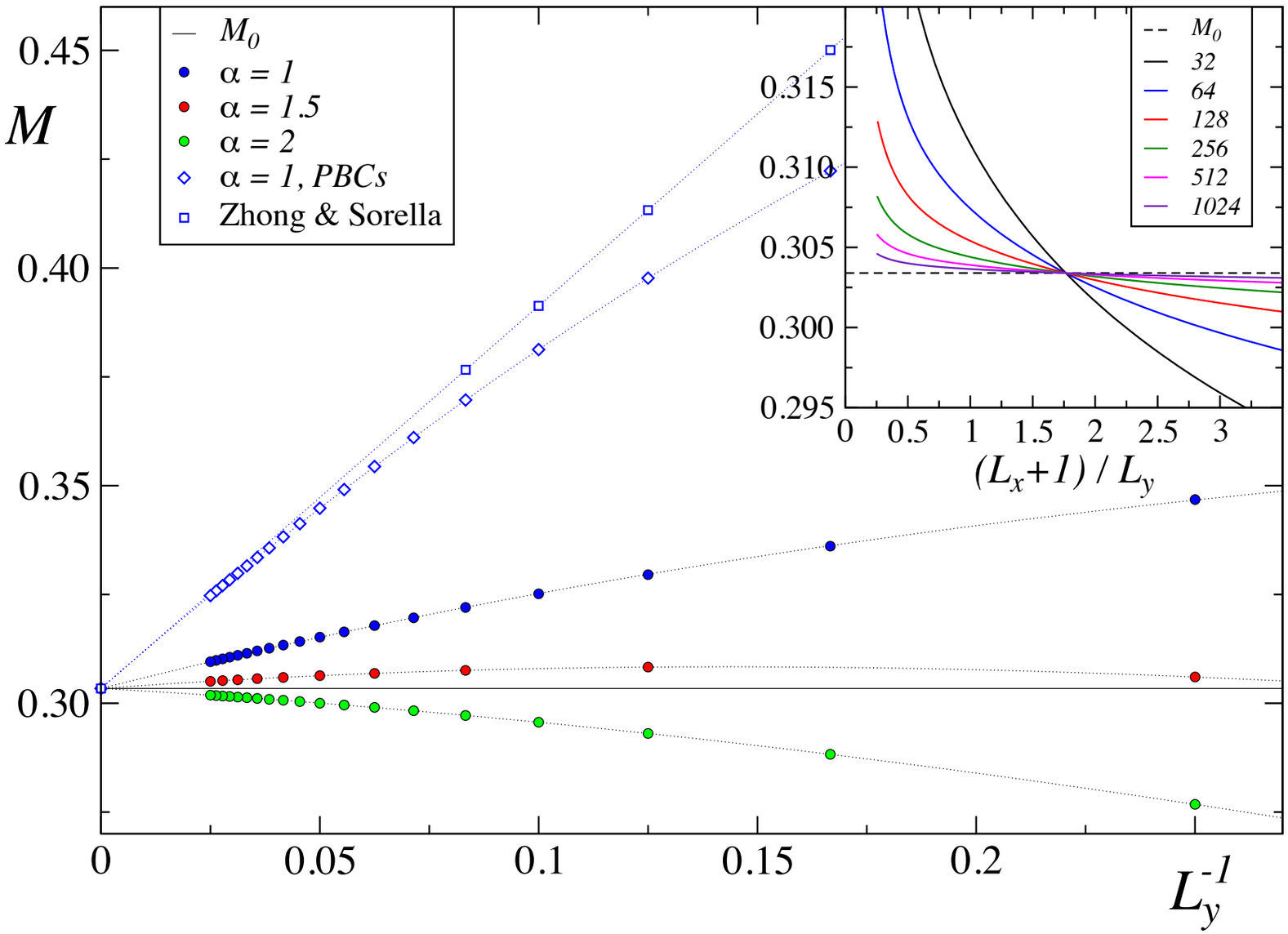
- "Standard" measurements in QMC estimate $\bar{M}^2$ using correlation functions and have large finite size effects $O(1/L_y)$
- Can one choose a special aspect ratio to eliminate $O(1/L_y)$ term?
- What is behavior at large length scales? Use finite system spin wave theory as a guide.

Square lattice



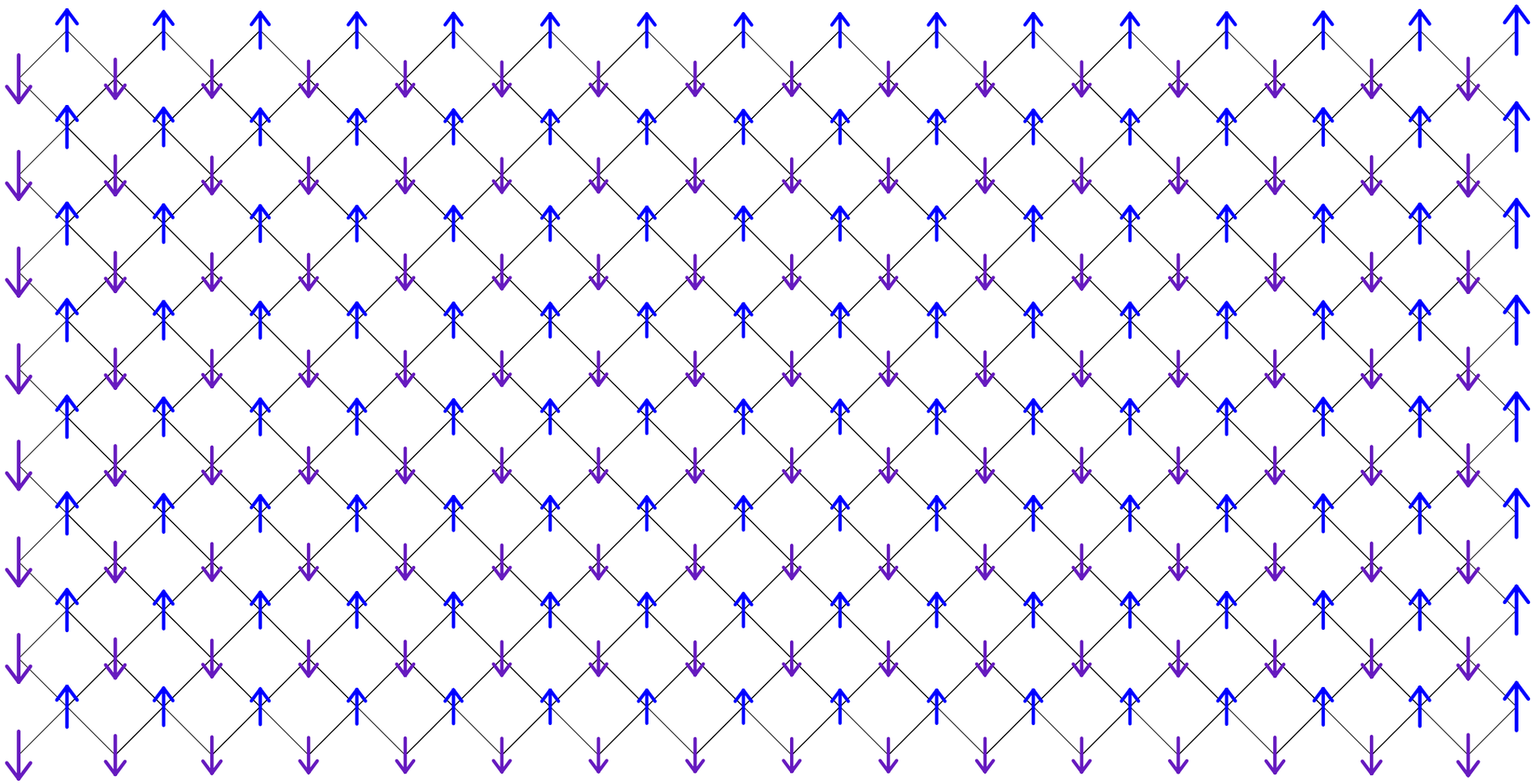
$\alpha = L_x/L_y$

Finite size spin wave theory



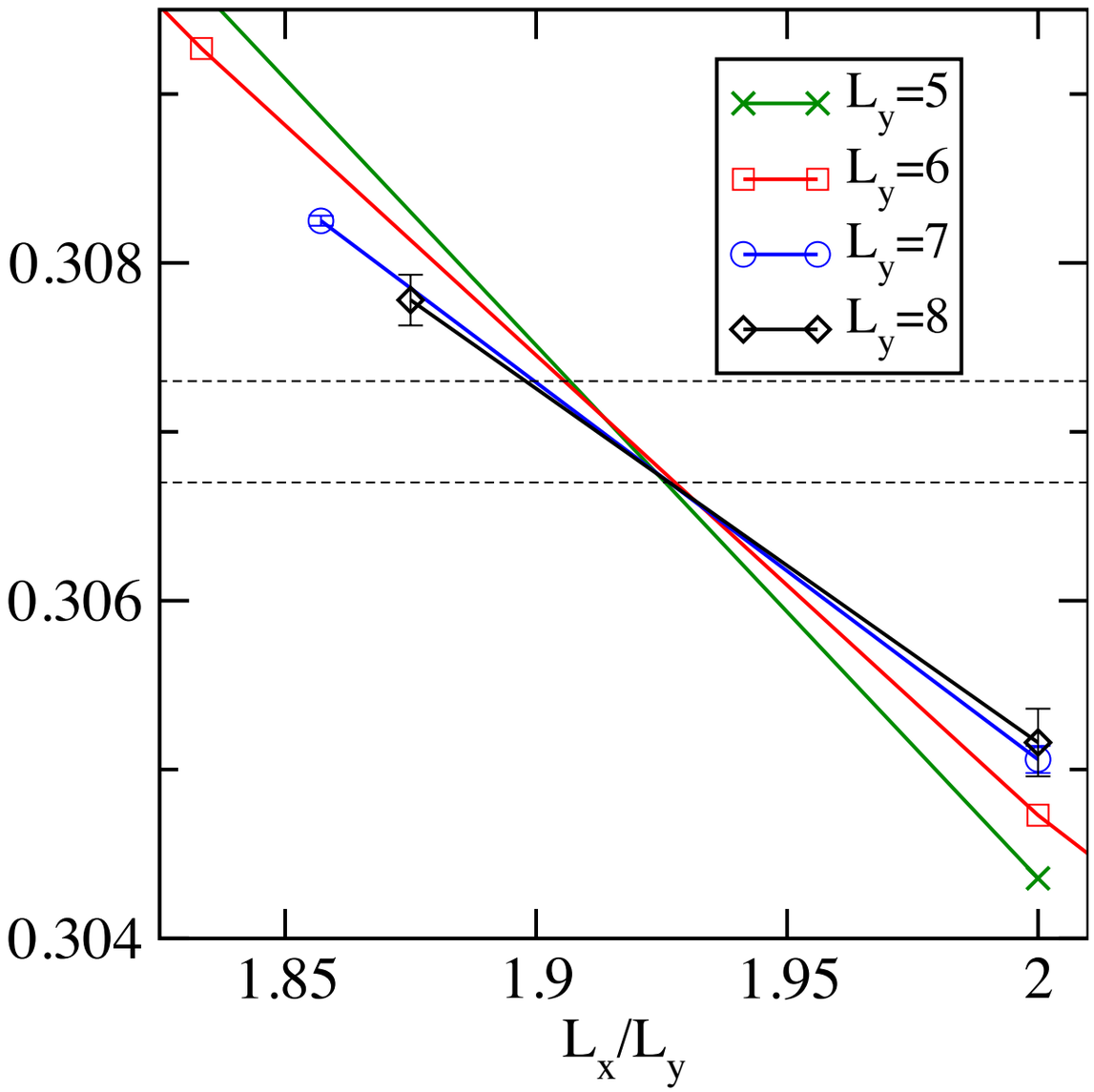
- Optimal choice $\alpha = 1.764$ eliminates linear term
- Even $\alpha = 1$ has much smaller finite size effects

Tilted square lattice



- Tilted lattice has smaller DMRG errors for its width
- For this "16 $\sqrt{2}$ x 8 $\sqrt{2}$" obtain M = 0.3052(4)

Tilted square lattice



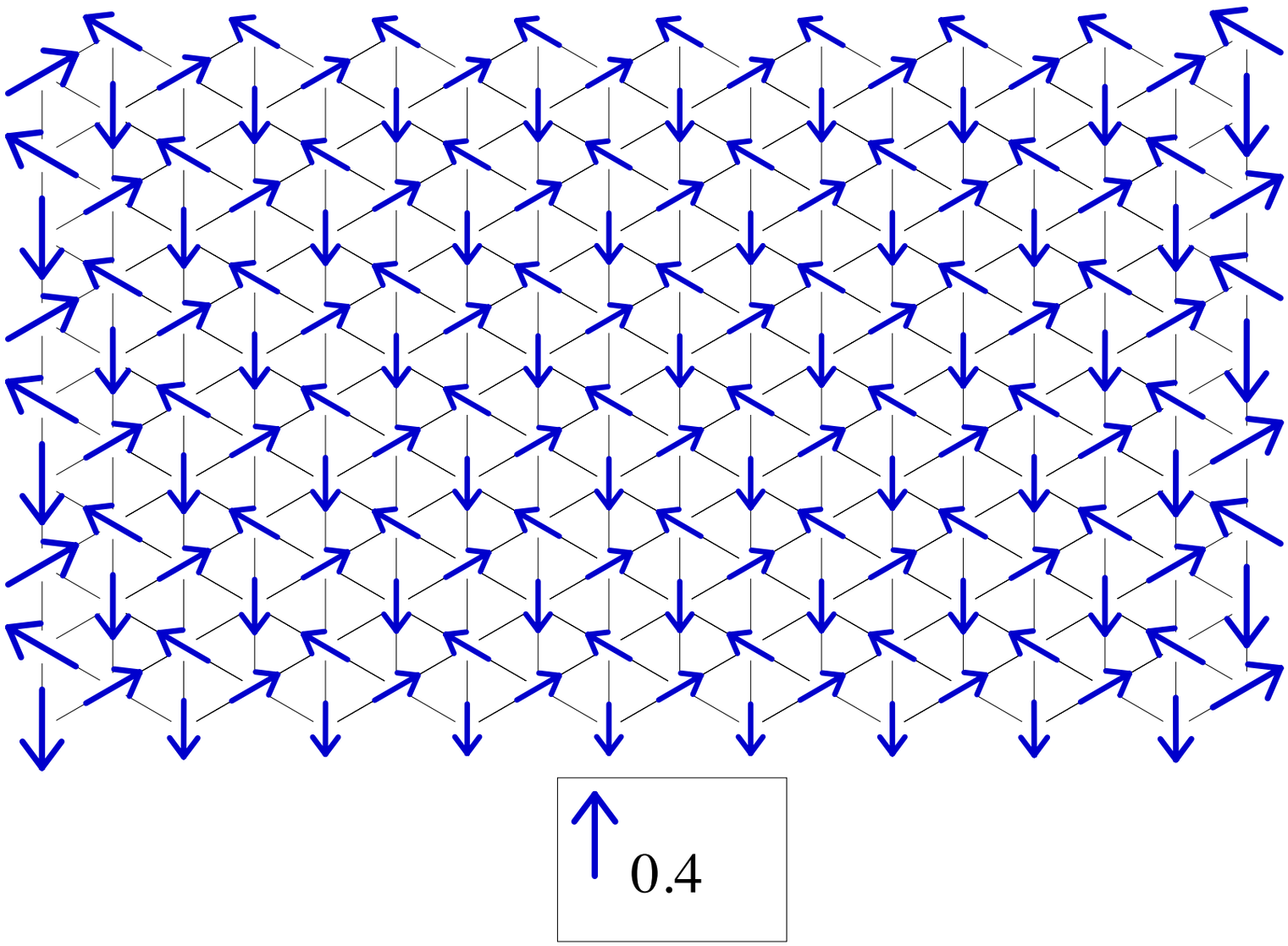
Sandvik, QMC

Energy, extrapolated to thermo limit: -0.669444(5)

Sandvik (1997): -0.669437(5)

- Results are consistent with and with comparable accuracy to QMC! (of 1997, at least)
- Latest QMC (Sandvik&Evertz) -0.30743(1) (No new E)

Traditional DMRG for triangular lattice Heisenberg model



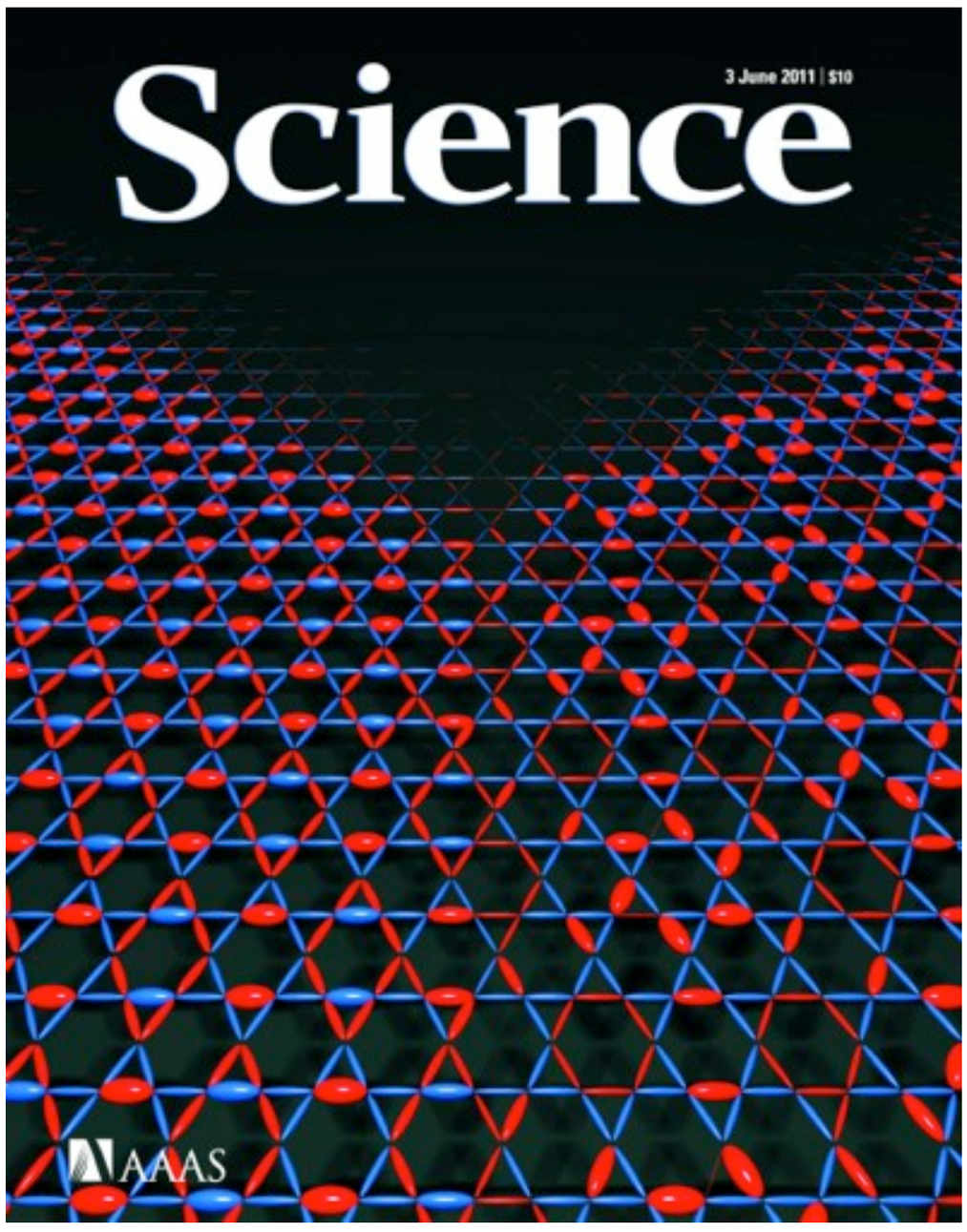
See White & Chernyshev, PRL 99, 127004 (2007)

$\Delta E \sim 0.3\%, \quad \Delta \langle S_z \rangle \sim 0.01$

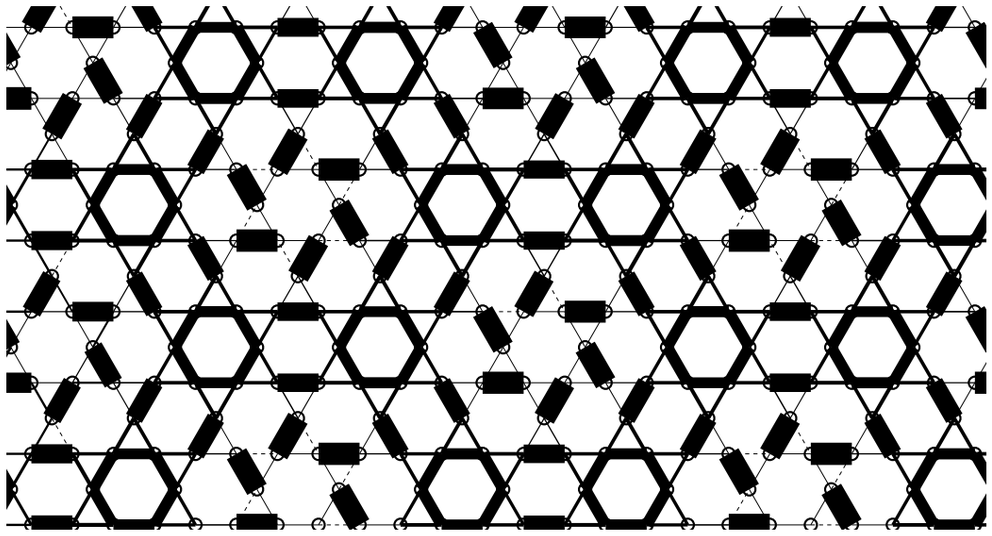
Extrap order param to thermodynamic limit: $M = 0.205(15)$

Spin Liquid Ground state of the S=1/2 Heisenberg model on the Kagome lattice

Collaborators: Simeng Yan and David Huse



A little history



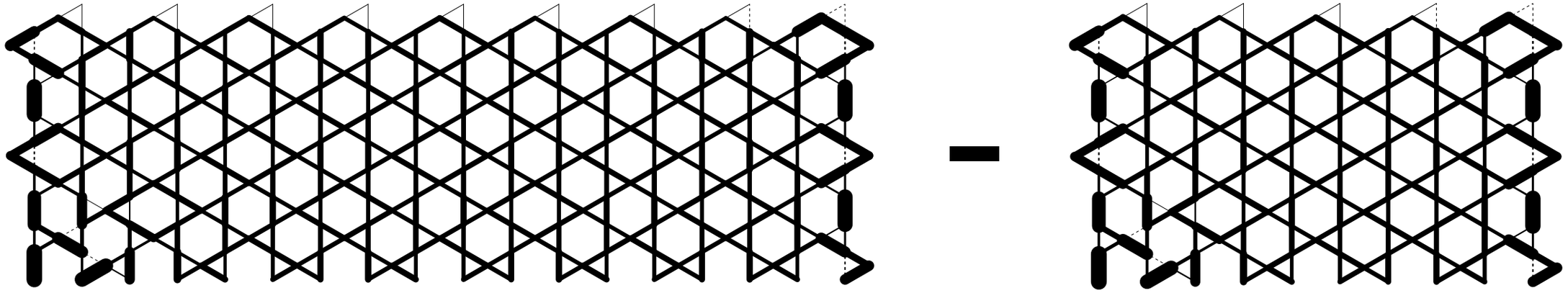
The S=1/2 Heisenberg Kagome systems has long been thought to be an ideal candidate for a spin liquid because of its high frustration. General agreement there is no magnetic order.

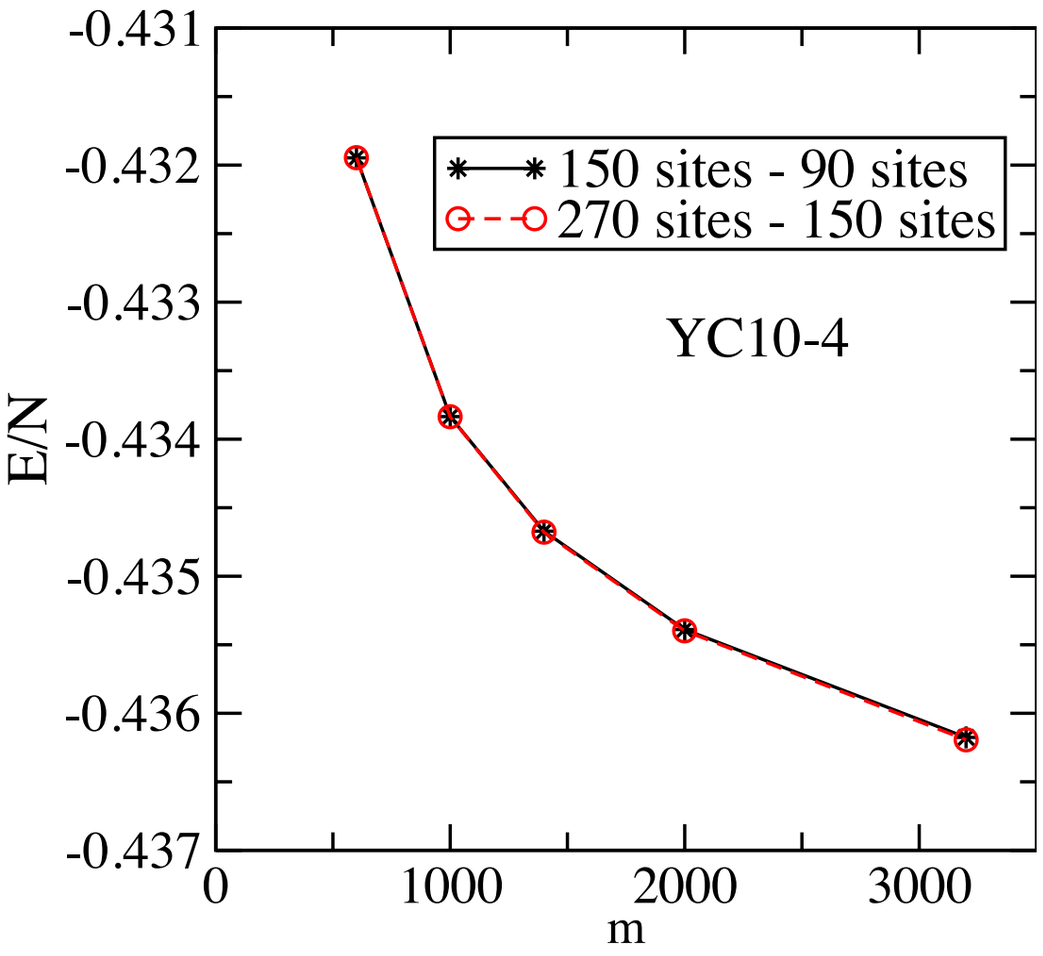
- Key question: is it a valence bond crystal or a spin liquid? What kind of VBC or SL?
- Until our work most of the work favored this 36 site "honeycomb" VBC

Practical Issues for Kagome

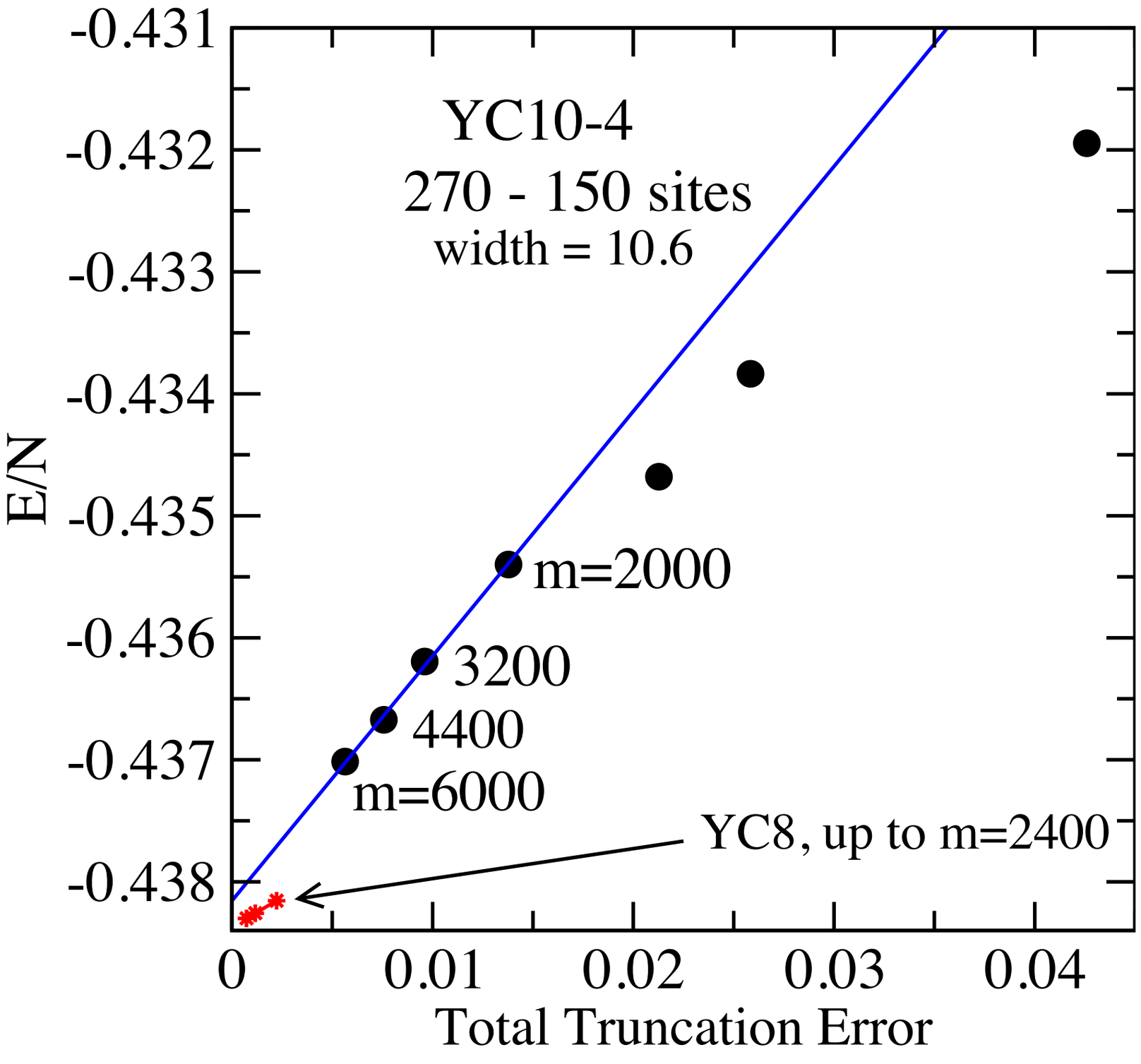
1. Metastability: getting stuck in a higher energy state (usually an issue only on wider cylinders)
   - Need to understand system and find a simple state close to the ground state to initialize DMRG
2. Strong dependence on width (and shift) of cylinders
   - Need to do many cylinders and understand patterns of behavior
3. Open edges--obtaining bulk cylinder behavior
   - This is a minor problem for this system
   - Open ends useful for pinning, selecting different topological sectors...

Ground state energies per site

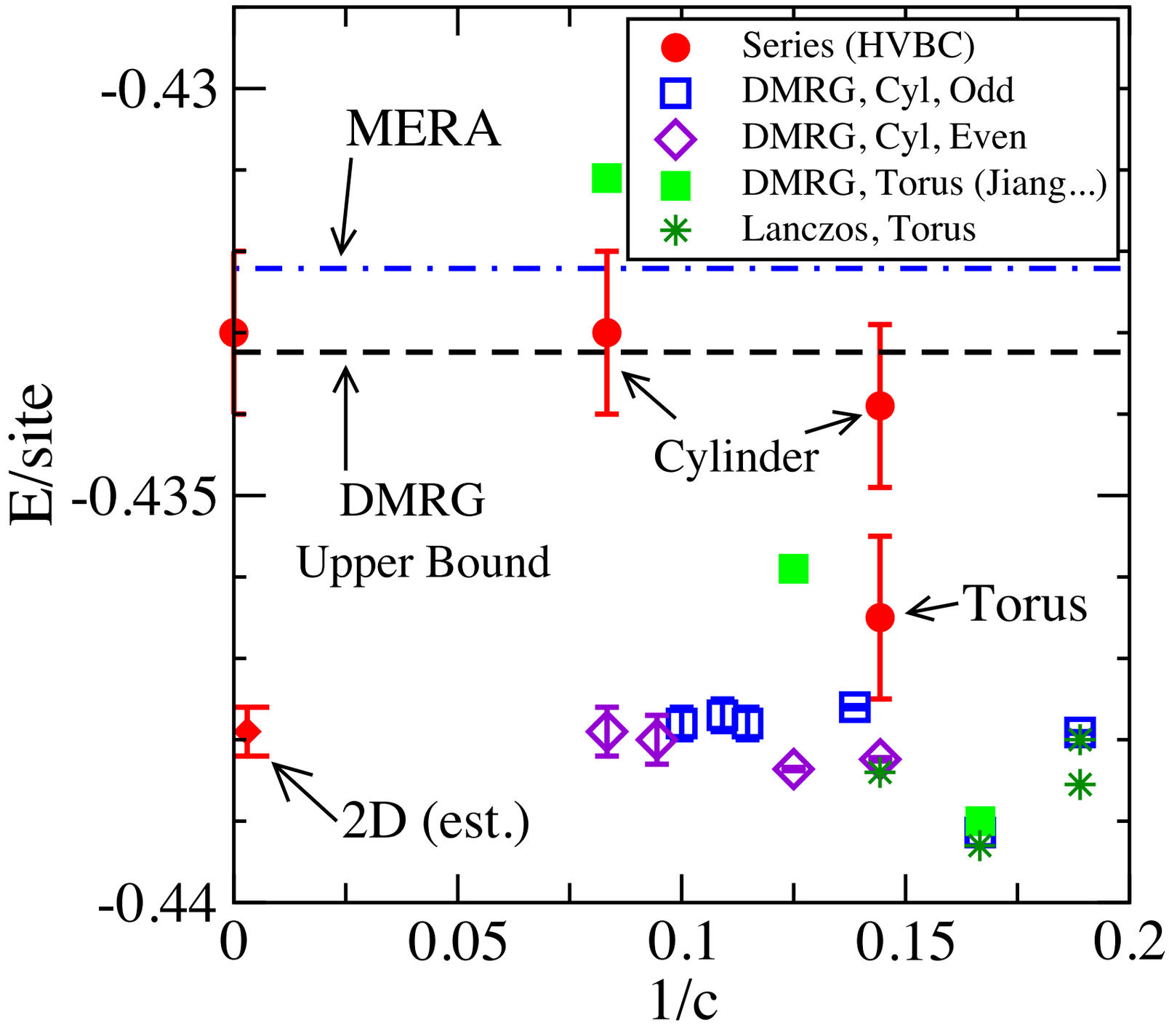




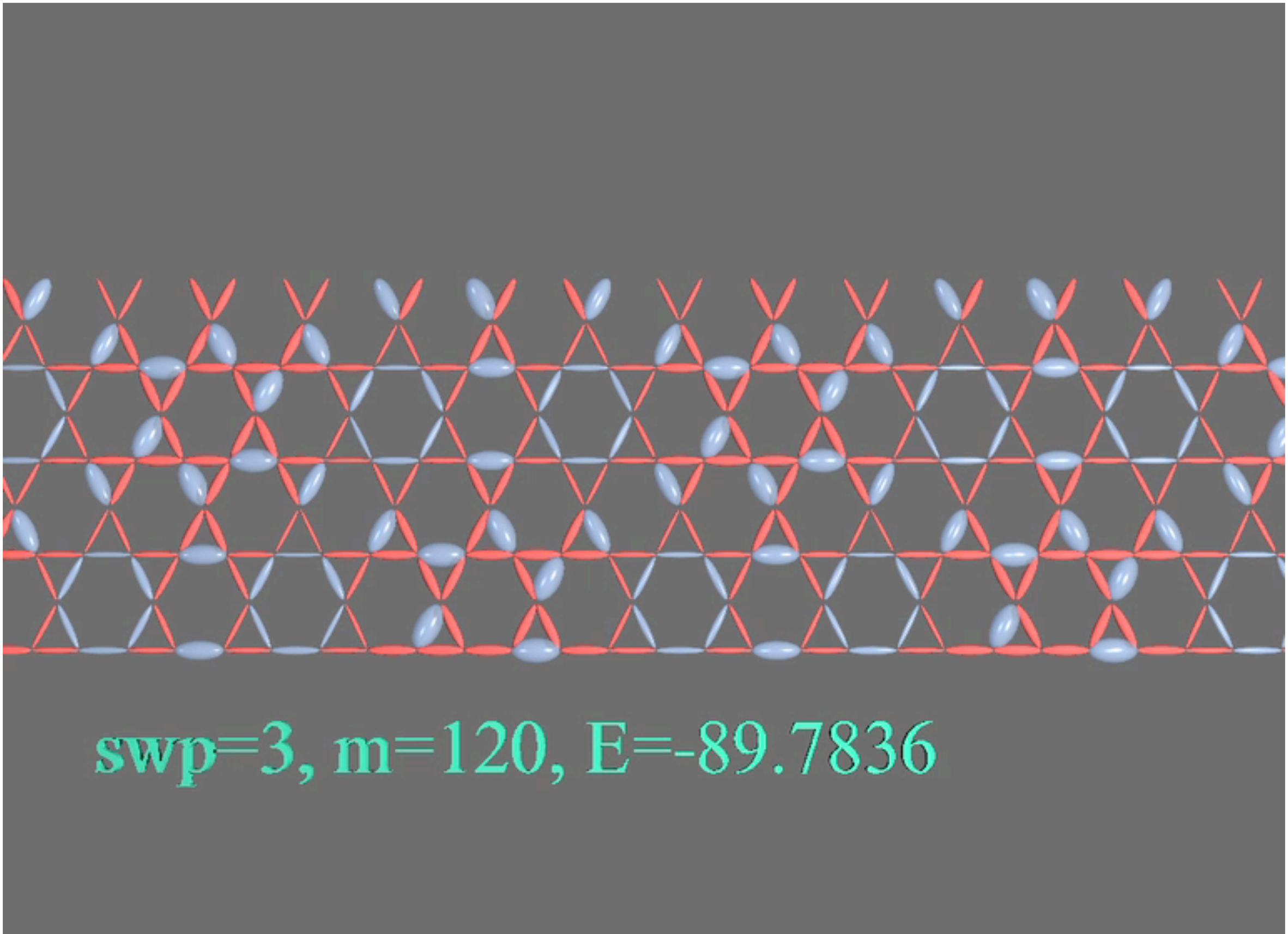
Smaller widths are nearly exact



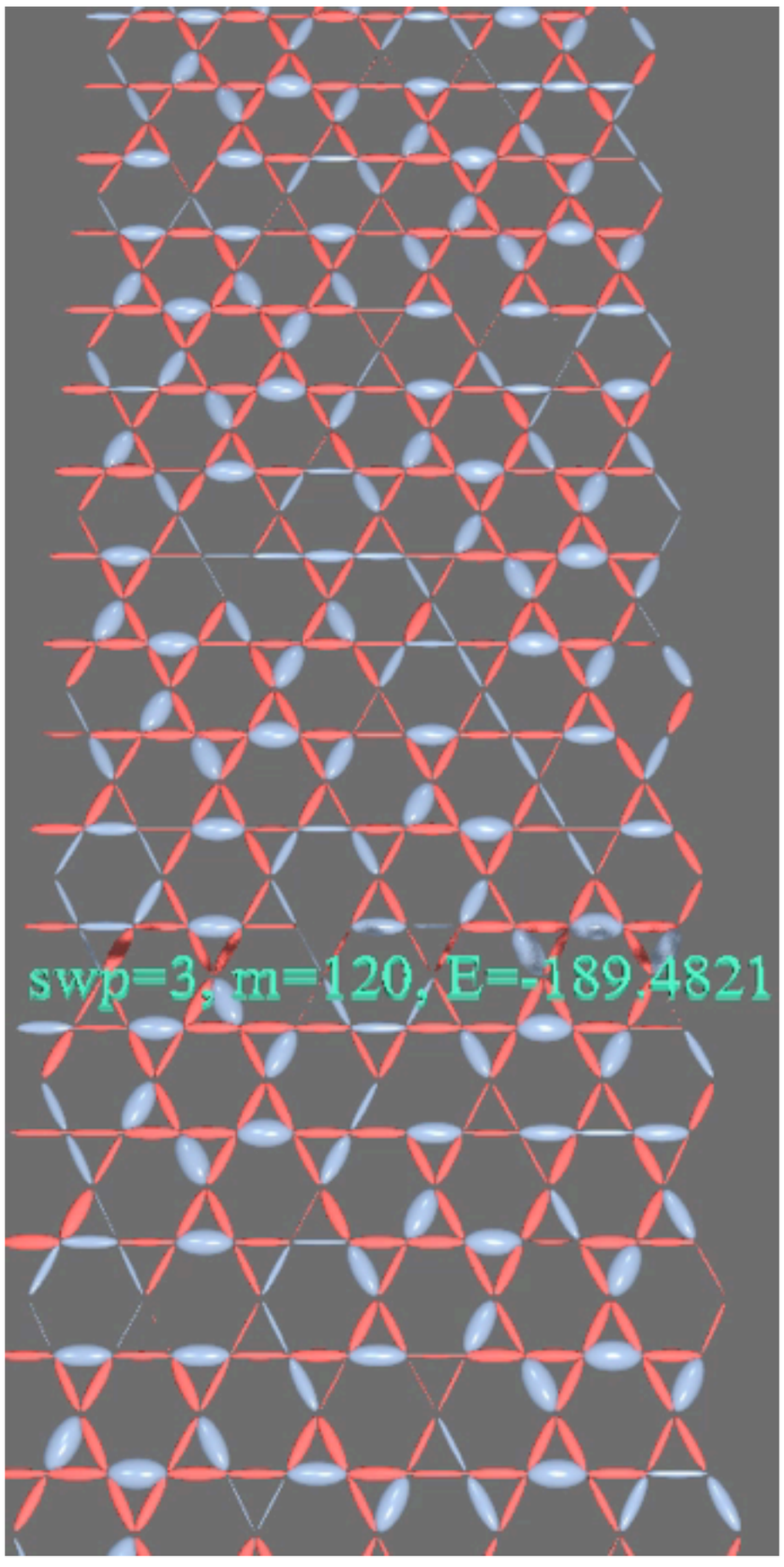
Energies of various cylinders and methods

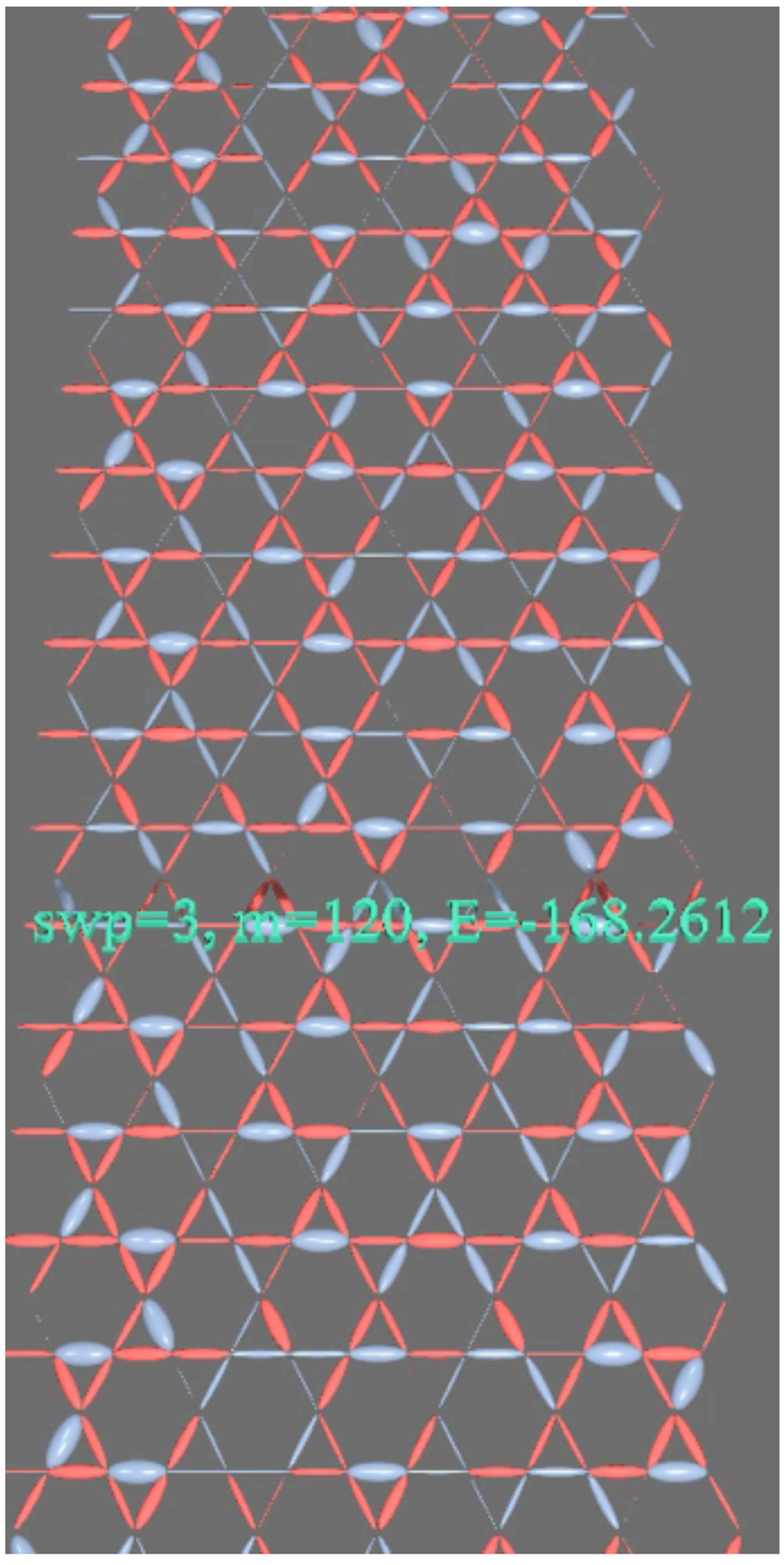


XC8 cylinder, biased to HVBC

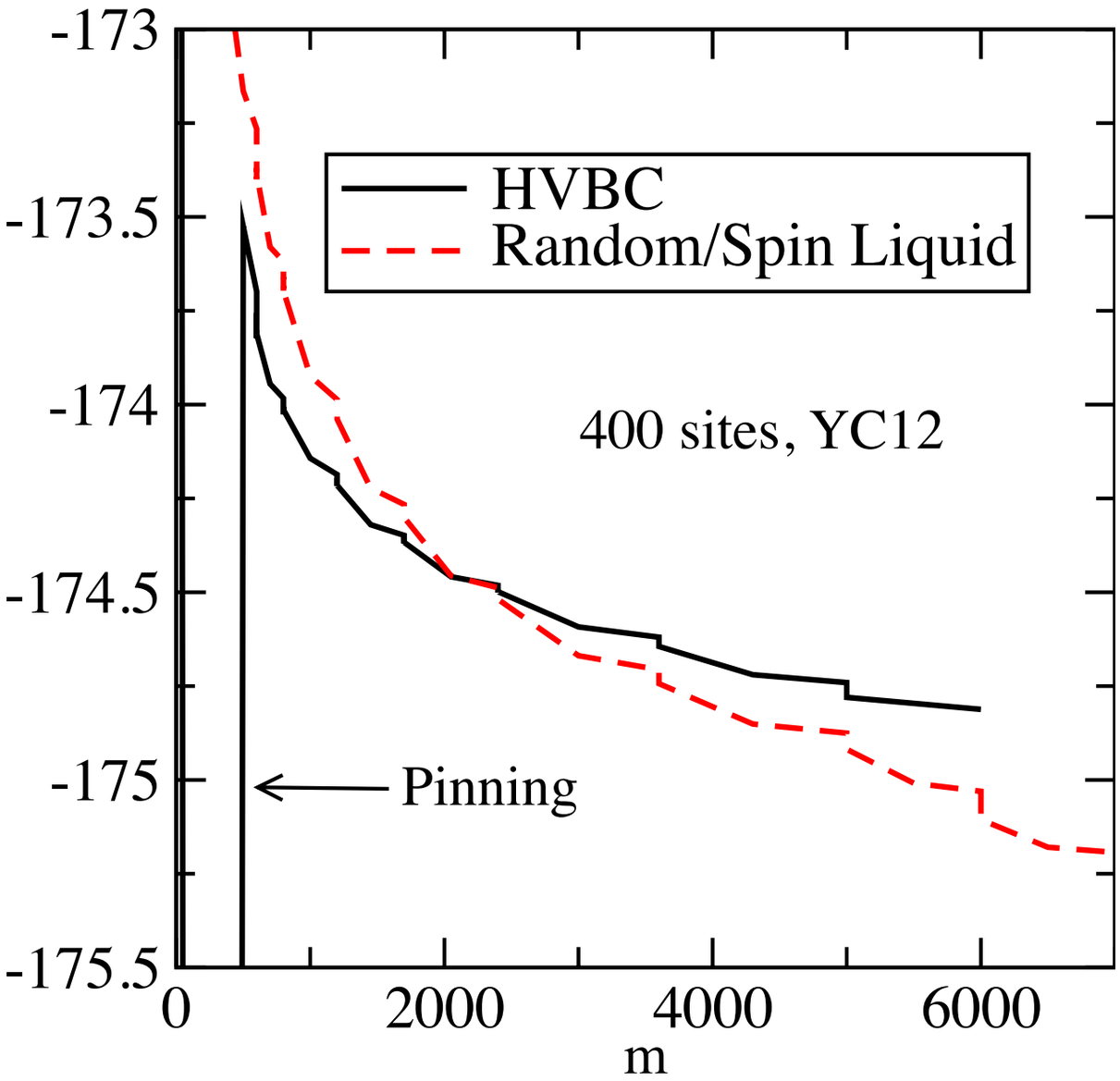


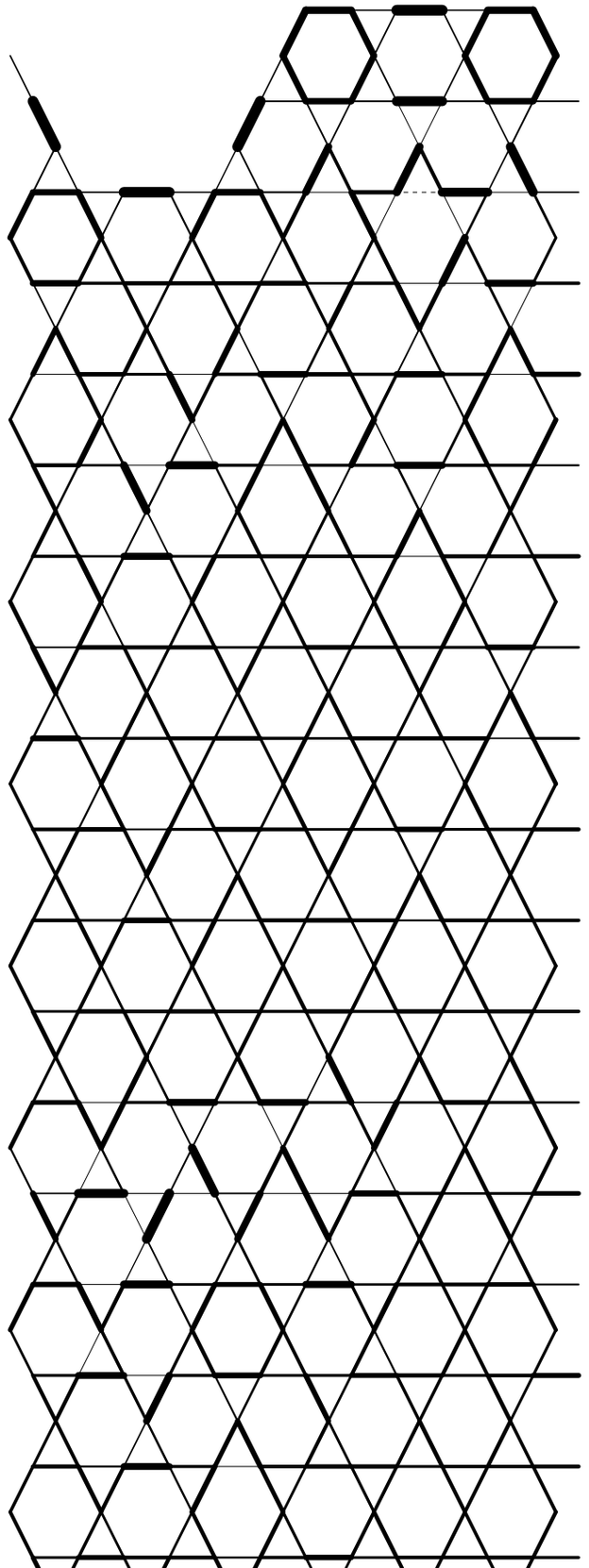
YC12 cylinder, one started in HVBC





Ruling out an HVBC on a width 12 cylinder





Summary

- DMRG and related low entanglement approximations have been the most powerful and diverse techniques for 1D systems known.
- Recently, many 2D models with either frustration or fermions can be treated on cylinders large enough to extrapolate to 2D
- Lots of fascinating connections to quantum information and entanglement