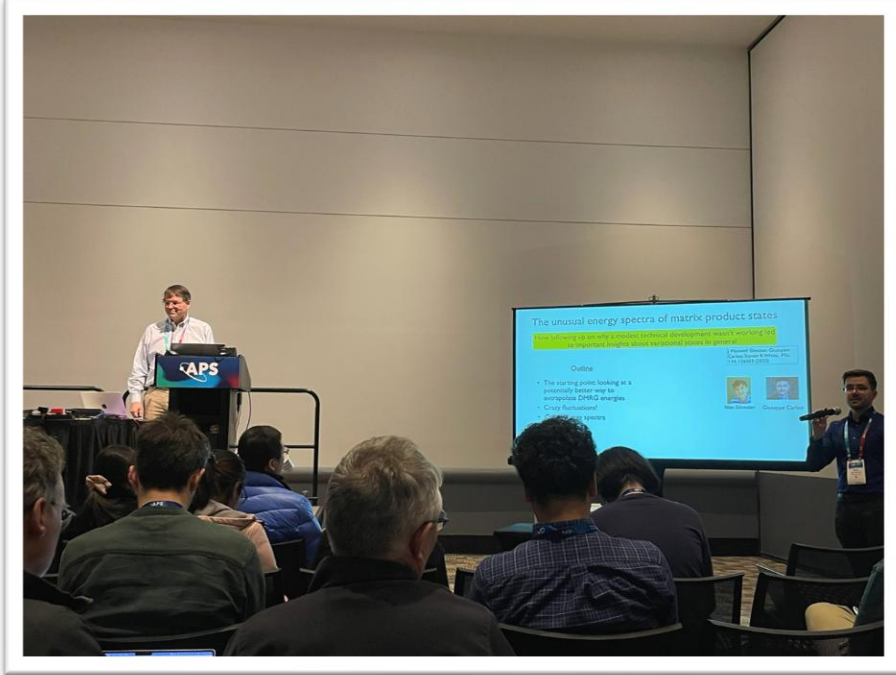
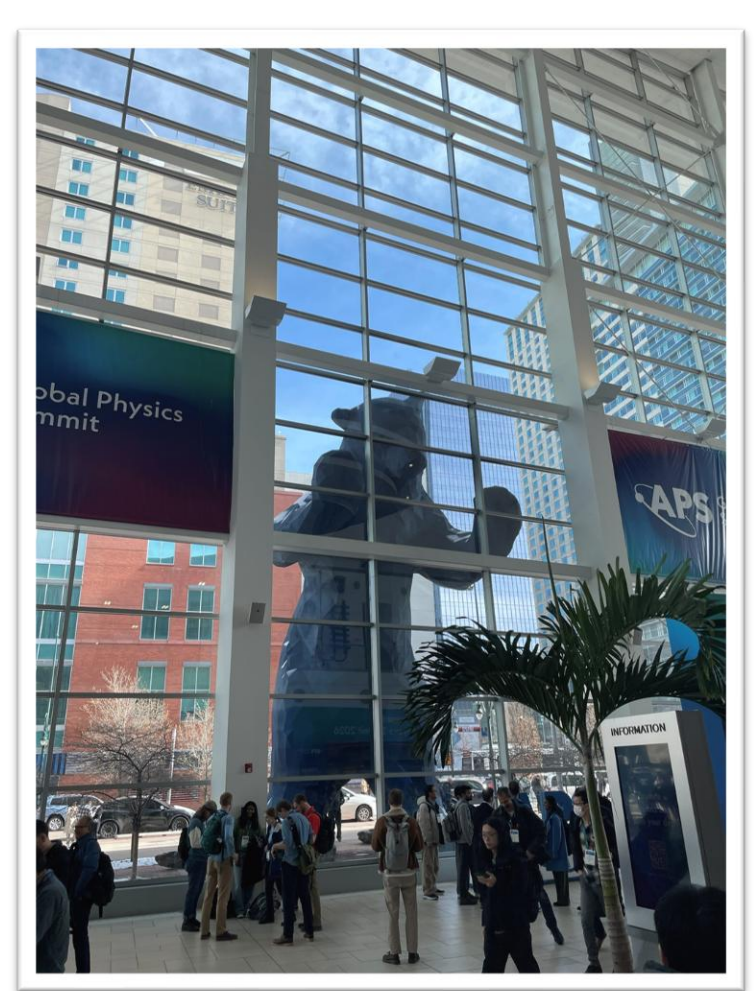




MPS-VMC



REF:

- [1] A.J. Ferris and G. Vidal, Phys. Rev. B 85, 165146 (2012).
- [2] A.W. Sandvik and G. Vidal, Phys. Rev. Lett. 99, 220602 (2007).
- [3] J. Maxwell Silvester ,Giuseppe Carleo and Steven R. White, Phys. Rev. Lett. 134, 126503 (2025)

DMRG compares with Imaginary Time Evolution (ITE) : *Energy Spectra*

$$H = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j, \quad \text{Model : 4*4 Heisenberg Square benchmarked by ED}$$

Both DMRG and ITE set Néel state originally , measuring energy spectra when they decay to the same energy.

$$|\psi\rangle = \sum_n c_n |n\rangle,$$

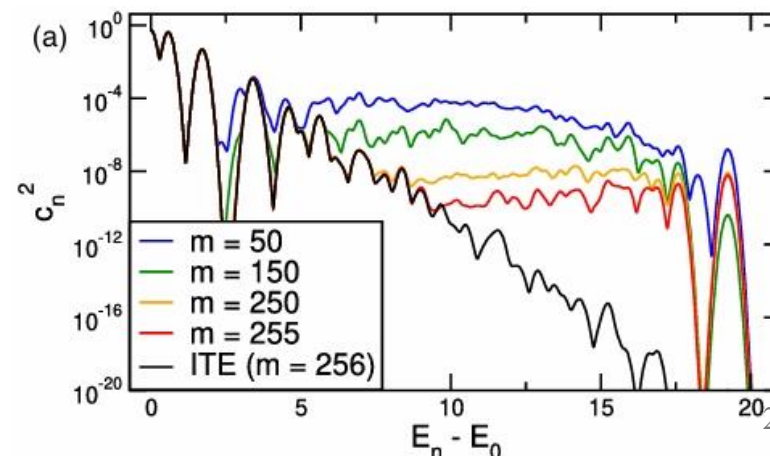
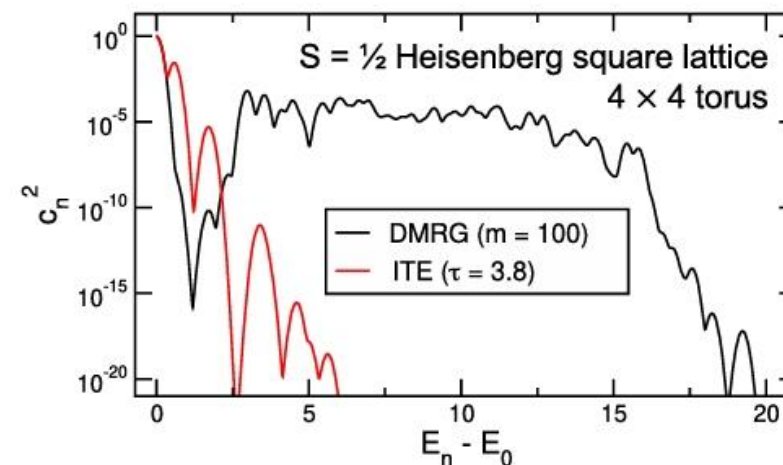
$$\frac{|c_n(\tau)|^2}{|c_0(\tau)|^2} = \frac{|c_n(0)|^2}{|c_0(0)|^2} e^{-2\tau(E_n - E_0)}$$

ITE methods should have exponentially decreased high-energy coefficients in the gapped system.

Coefficient c_n shows the energy spectra
For DMRG , high-energy coefficients does not decrease exponentially with the gap.

Max bond dimension:
 $d^{N/2} = 2^8 = 256$

The unusual energy spectra is due to the compression of the ITE state into an MPS ,
NOT the result of DMRG sweep procedure.



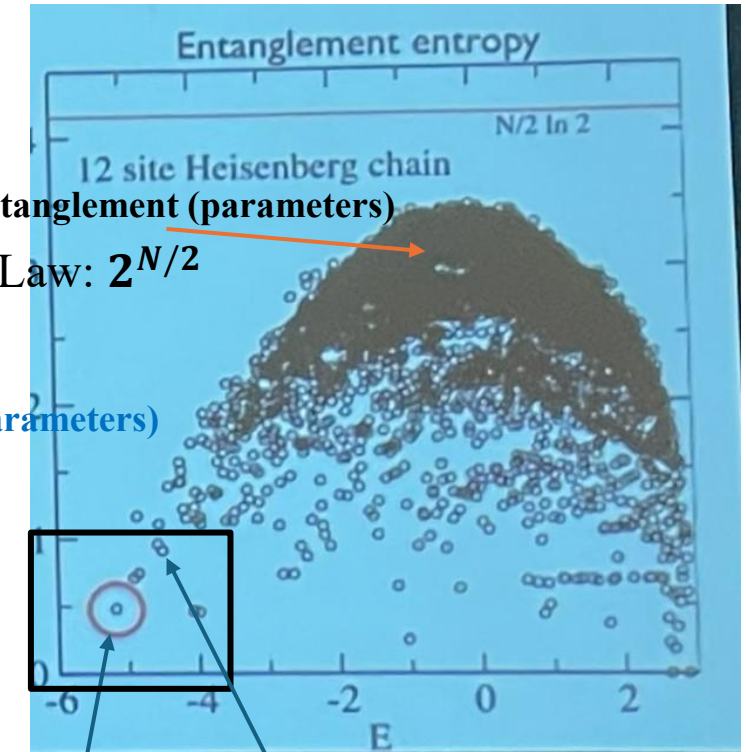
Why it occurs ? The setting bond dimension m is a compression of the real number of parameters . It borrows high energy state parameters to cancel extra parameters.

When we carry out a fixed bond dimension $m = 100$ DMRG to certain MPS :

- MPS state , The approximate ground state is the low lying states
- It consists $m = 1000$ and $m = 100$ states . Hard to set $m = 1000$, lack computational power $O(Nm^3)$
- 这里只是举例，实际就设 $m=1000$ 一个sweep也只要不超过半小时。
- $|\psi\rangle = \sum_n c_n |n\rangle$ Now $|\psi\rangle$ is the set MPS with $Nm^2 = 10^4 N$ parameters.
- Now $|n\rangle$ is the approximate ground state , containing $Nm^2 = 10^6 N$ parameters.
- After SVD ,we want $|\psi\rangle$ goes to $Nm^2 = 10^4 N$ parameters , as set.
- But the real state contains $Nm^2 = 10^6 N$ parameters.
- $|\psi(10000)\rangle = \sum_n c_n |n(1000000)\rangle$ we want to cancel extra parameters.
- **The number of low lying states is not enough to cancel extra parameters ! (functions way fewer than parameters)**
- We have many high energy states.
- $|n\rangle$ has to borrow high energy states so that DMRG can continue
- High energy states has exponent entanglement (parameters)
- We should borrow more to cancel the extra parameters from high energy
- We borrow almost everything, resulting long tail energy spectra.

High energy state
Many states , high entanglement (parameters)

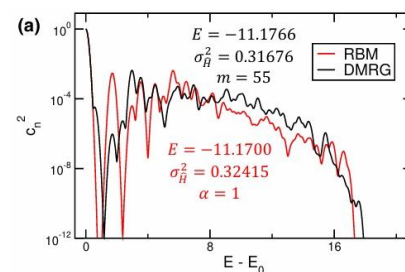
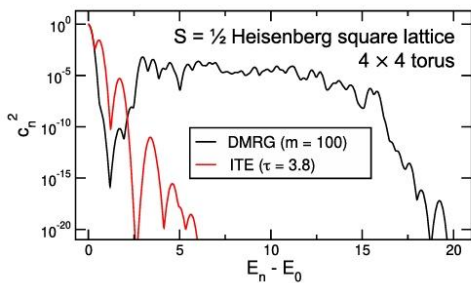
Volume Law: $2^{N/2}$



Area Law Nm^2 , eg: $m = 1000$

Area Law Nm^2 , eg: $m = 100$

MPS state , The approximate ground state
Few states , low entanglement (parameters)



Parameters : singular value matrix cutoff to $m \times m$

- Similar results in neural quantum states (RBM) , since the hidden unit parameters are always not enough to represent the full wave function.

Application : Sampling MPS ----- Creating Samples

Instead of doing VMC through a Markov chain , now we do a sweep , sampling MPS states s

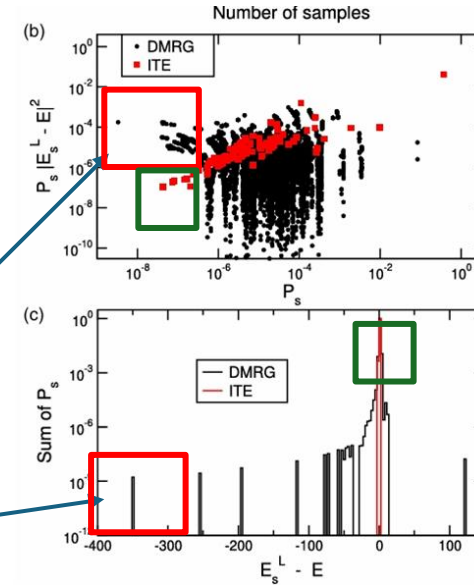
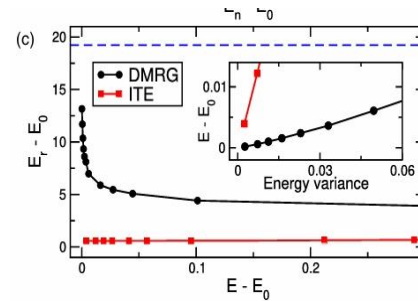
$$P_s = |\langle \psi | s \rangle|^2 \quad E_s^L = \langle s | \hat{H} | \psi \rangle / \langle s | \psi \rangle$$

- We have stored the local tensor for each site, we know the right is orthogonalized, don't need to do the sweep again.
- Suppose the left index ($s_1 \dots s_{i-1}$) has been contracted.
- Now we are working on site index i , **with** s_i . calculating P_s . Eg; spin system $s_i \in \{\uparrow, \downarrow\}$, right equals to 1.
- Follow DMRG procedure, add a block : Left part $V_{\alpha_i}^{(s_i)} = \sum_{\alpha_{i-1}=1}^m L_{\alpha_{i-1}} A_{\alpha_{i-1}, \alpha_i}^{[i] s_i}$ A is the local tensor of i ($m \times m$), L is the orthogonalized left block ($1 \times m$), this process is a Tensor contraction, costing $O(dm^2 + md^2)$
- $P(s_1, s_2, \dots, s_N) = |\langle s_1, s_2, \dots, s_N | \Psi \rangle|^2 = P(s_1)P(s_2 | s_1)P(s_3 | s_1, s_2) \dots P(s_N | s_1, \dots, s_{N-1})$
 $P(s_i | s_1^*, \dots, s_{i-1}^*) = \sum_{\alpha_i=1}^m |V_{\alpha_i}^{(s_i)}|^2$ Roll a $P(s_i | s_1^*, \dots, s_{i-1}^*)$ to decide s_i , multiply $P(s_i | s_1^*, \dots, s_{i-1}^*)$ to get P_s
- Calculating E_s^L following the right contracting procedure, costing $O(wdm^2 + d^2w^2m)$ (w : *MPO bond dimension is usually small*)
- Go on to site index $i+1$
- **Creating a sample cost $O(wdm^2 + d^2w^2m)$, a DMRG sweep $O(wdm^3 + d^2w^2m^2)$ equals to thousands of samples in complexity**
- "Perfect (Direct) Sampling":
 - 1、 Directly sampled according to their probabilities in the wave-function, without resorting to a Markov chain process (no Equilibration Time)
 - 2、 Don't need to roll a random number according to $E(s_i)$, acceptance is 100% .

The diagram illustrates the contraction of tensors for sampling. It shows the contraction of a state $|\psi\rangle$ with a local operator $\hat{O}$, resulting in a trace of a product of tensors. The diagram uses various indices: m for MPS bond dimension, d for local Hilbert space dimension, w for MPO bond dimension, and s for physical index. It shows how the trace is calculated by contracting the tensors in a specific order, leading to a final result of 1.

Application : Sampling MPS -----Finding G.S. Observables

- Any DMRG MPS $|\psi\rangle = \cos\theta |0\rangle + \sin\theta |r\rangle$ r is the residue, $\sin\theta$ is small, if we stick to **ONE** DMRG sweep with $|\psi\rangle$, we have constant $E_r - E_0$, we don't care it's specific figure.
- In ITE, $E_r - E_0$ is almost zero, in DMRG **we have non-zero number due to the long-tail of high energy part ie: non-zero energy variance** We can utilize that !
- We have $|\psi\rangle$ and $E_\psi = \langle\psi|H|\psi\rangle$ given by DMRG sweep.
- Define energy variance: $\langle\psi|(\hat{H} - E_\psi)^2|\psi\rangle$ and energy $\Delta E = E_\psi - E_0$
- Proof: Energy Variance is Linear to Energy (Same with other local operators)*
- Thanks to the non-zero $E_r - E_0$, we can linear fit the function.
- In the algorithm, we carry out several sweeps with different bond dimensions, as in the figure, and gain several $|\psi\rangle$ and E_ψ .
- We already have $P_s = |\langle\psi|s\rangle|^2$ $E_s^L = \langle s|\hat{H}|\psi\rangle / \langle s|\psi\rangle$ in the samples.
- $\sigma_H^2 = \sum_s \langle\psi|\hat{H} - E|s\rangle \langle s|\hat{H} - E|\psi\rangle = \Delta_s + \sum_{|s\rangle \in S} P_s |E_s^L - E_\psi|^2$
- We can do that linear fit, and the **intercept is G.S. energy**.



$$|\psi\rangle = \cos\theta |0\rangle + \sin\theta |r\rangle \approx (1 - \frac{1}{2}\delta^2) |0\rangle + \delta |r\rangle$$

$$E_\psi = \langle\psi|H|\psi\rangle \approx (1 - \delta^2)E_0 + \delta^2 E_r$$

$$\Delta E = E_\psi - E_0 = \delta^2(E_r - E_0)$$

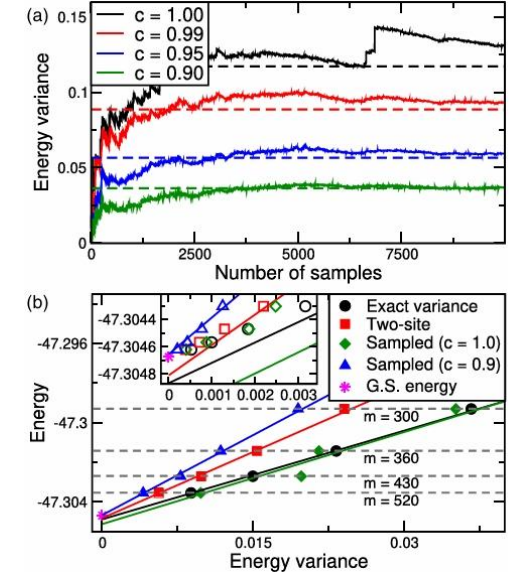
$$\delta_H^2 = \langle\psi|(\hat{H} - E_\psi)(\hat{H} - E_\psi)|\psi\rangle. \text{ Let } |R\rangle = (\hat{H} - E_\psi)|\psi\rangle$$

$$|R\rangle = [\hat{H} - E_0 - \delta^2(E_r - E_0)][(1 - \frac{1}{2}\delta^2)|0\rangle + \delta|r\rangle]$$

$$= \delta^2(E_r - E_0)|0\rangle + \delta(E_r - E_0)|r\rangle + \text{higher order.}$$

$$\delta_H^2 = \langle R|R\rangle \approx \delta^2(E_r - E_0)^2 + \text{higher order.}$$

$$\text{Therefore: } \delta_H^2 = (E_r - E_0) \cdot \Delta E \quad \boxed{E_\psi = c_0 \cdot \delta_H^2 + E_0}$$



- Multiply $P(s_i | s_1^*, \dots, s_{i-1}^*)$ to get P_s , it is usually small.
- Right figure suggests the DMRG distribution has samples with very small probability but substantial contributions to the variance.
- Δ_c is the cutoff according to energy, it is $\sum_{|s\rangle \in C} P_s |\varepsilon_c - \varepsilon|^2$ by replacing $E_s^L - E_\psi$ with $\varepsilon_{\max}$ if it is larger than $\varepsilon_{\max}$
- Define the cutoff ratio, c , as the percentage of the small $E_s^L - E_\psi$, replace other energies.
- $c = 0.9$ is the most optimized cutoff.