Self Learning Monte Carlo Method

Jie Xiong
Chizhou Wang
2021 / 12 / 14

Contents

Part I: Problem Description

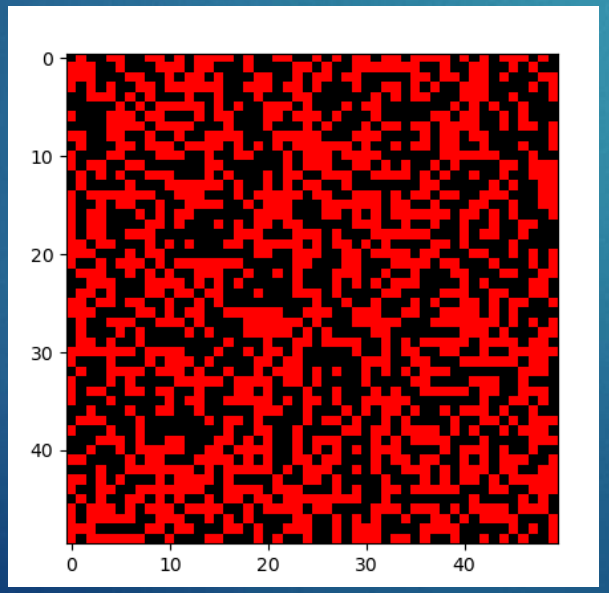
Problem Description

Monte Carlo Simulation (cf. lecture 2)

- Create a collection of samples which obeys Maxwell-Boltzmann distribution $\rho(\{\sigma_i\}) = \exp\{-\beta E(\{\sigma_i\})\}$;

- Use the samples to calculate the physical quantities by 'ensemble averaging'.

Question: How to create samples according to the distribution?



$$H = -J \sum_{\langle ij \rangle} S_i S_j - K \sum_{ijkl \in \square} S_i S_j S_k S_l$$

J = 1.0; K = 0.2

Problem Description

Metropolis-Hastings Method (cf. lecture 2)

Construct a stationary Markov-chain of spins configurations

$$P_{n+1}(x) = \sum_{y \neq x} W_{yx} P_n(y) + \left(1 - \sum_{y \neq x} W_{xy}\right) P_n(x) \rightarrow \sum_{y} W_{yx} P(y) = \sum_{y} W_{xy} P(x) \rightarrow W_{yx}\,\rho(y) = W_{xy}\,\rho(x)$$

Two steps:

- Propose a change x → y with probability $T_{x \to y}$;
- Accept the change with probability $A_{x \to y} = \min\left\{1, \frac{\rho(y) T_{yx}}{\rho(x) T_{xy}}\right\}$;
- $W_{x \to y} = T_{x \to y}\ A_{x \to y}$

Problem Description

Local Update Method

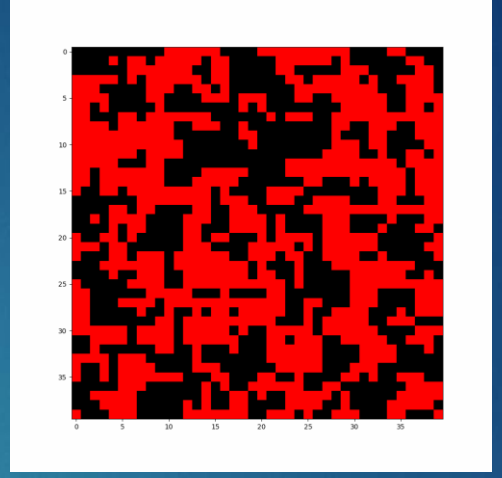
- $T_{x\to y} = T_{y\to x} = \frac{1}{N}$, configuration x and y only differs by only one spin.
- $A_{x\to y} = min\left\{1, \frac{\rho(y)}{\rho(x)}\right\}$;
- N local update steps are treated as one global update.

Drawback

- Slow exploration of the configuration space;
- Critical slowing down at phase transition temperature;
- Highly inefficient to describe long-range correlations.

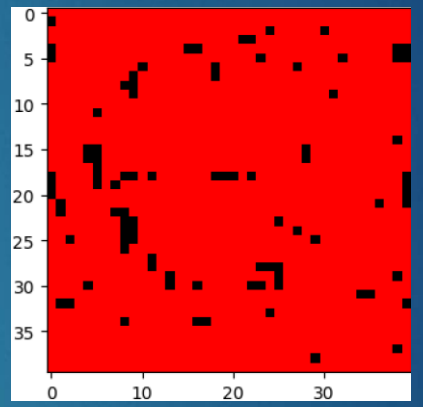
Advantage

- Fast calculation of energy;
- Comparatively high acceptance ratio

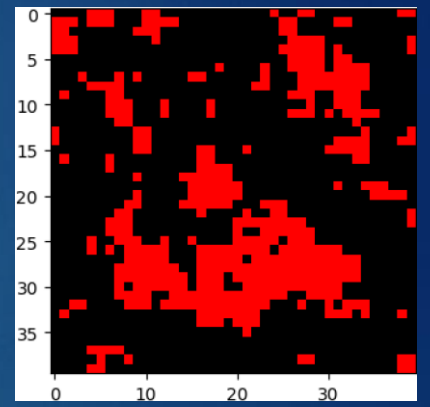


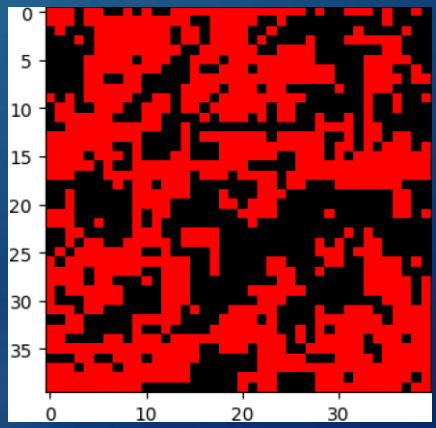
Tc = 2.5, Local Update Simulation

T = 2.2



Tc = 2.5





T = 3.0

Problem Description

Wolff Global Update

Proposed change:

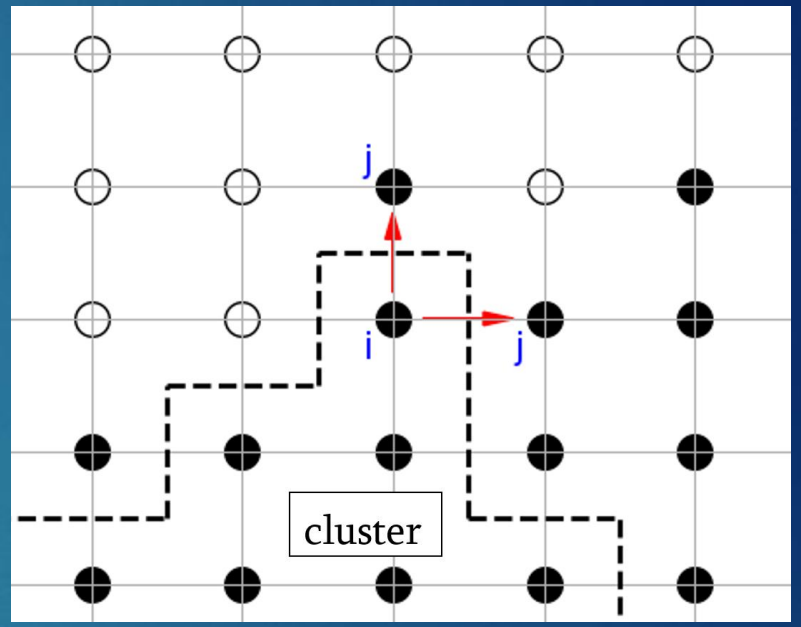
Construct a cluster to simulate long range correlations:

(1) Randomly choose a spin, add it to the cluster;

(2) For each site i in the cluster, consider the Nearest Neighbouring site j outside the cluster, activate the link with probability $p(i \to j) = \max\{0, 1 - e^{-2\beta J S_i S_j}\}$;

(3) Extend the cluster by adding all activated spins into it. Then go back to (2);

(4) Finish the construction if the extension stops or all the spins have been examined



$$\frac{T(x \to y)}{T(y \to x)} = \prod_{\langle i,j \rangle, i \in c, j \notin c} \frac{1 - p_{i \to j}(x)}{1 - p_{i \to j}(y)} = \prod_{\langle i,j \rangle, i \in c, j \notin c} e^{-2\beta J S_i^x S_j^x}$$

Problem Description

Wolff Global Update

Acceptance ratio: $A_{x\to y} = \min\{1, e^{\beta\Delta E}\}$

$$\Delta E = K\sum_{ijkl\in\square}(S_i^B S_j^B S_k^B S_l^B - S_i^A S_j^A S_k^A S_l^A)$$

Drawback

- Slower calculation of energy;
- Lower acceptance ratio;

Advantage

- Description of long-range correlation;
- Short correlation time;
- Extend to build some clusters, describe different parts of the correlation.

Tc = 2.5,
Wolff Update Simulation

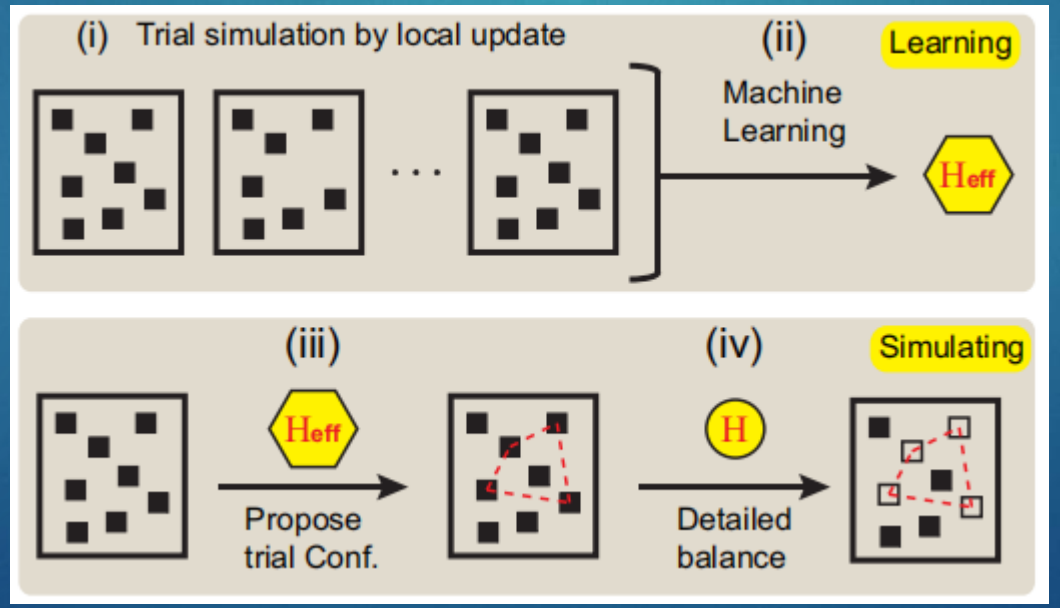
Part II
Self Learning Monte Carlo Method

Self Learning Monte Carlo Method

How to speed up the simulation & describe long range correlation?

Machine Learning:

- Reveal unobvious patterns in the model;
- Learn and self-optimized automatically.



Self Learning Monte Carlo Method

$$H_{eff} = E_0 - J_1 \sum_{\langle ij \rangle_1} S_i S_j - J_2 \sum_{\langle ij \rangle_2} S_i S_j - \cdots$$

- Propose an effective Hamiltonian;
- Construct the cluster “in the Wolff way” with this Hamiltonian;
- Acceptance ratio:

$$A_{x \to y} = \min\left\{1, \frac{\rho(y)}{\rho(x)} \frac{\rho_{eff}(x)}{\rho_{eff}(y)}\right\}$$

$$\rho_{eff} = e^{-\beta E_{eff}}$$

- Train H_{eff}, make it as closer to H as possible at given temperature.

Self Learning Monte Carlo Method

Training Process (Self Learning)

(1) Temperature Descent

- Initial samples are obtained from Local Update Method at T = 5;
- Effective Hamiltonian is trained using these samples;
- Self learning update method is used to generate more samples at lower temperature;
- New effective Hamiltonian is trained;
- Repeat the process until T = Tc. Finally one will get H_{eff} at Tc.

(2) Ensemble average

- Initial samples are obtained from Local Update Method at T = 5;
- Effective Hamiltonian is trained using these samples;
- Self learning update method is used to generate more samples at T = Tc;
- Various effective Hamiltonians are trained at T = Tc, we take the average value.

Self Learning Monte Carlo Method

Advantage:

- High acceptance ratio;
- Low correlation time;
- Can describe long range correlation.

Drawback:

- Training process is time consuming;
- Calculation time increases much with increasing size.

Restricted Self Learning Monte Carlo Method

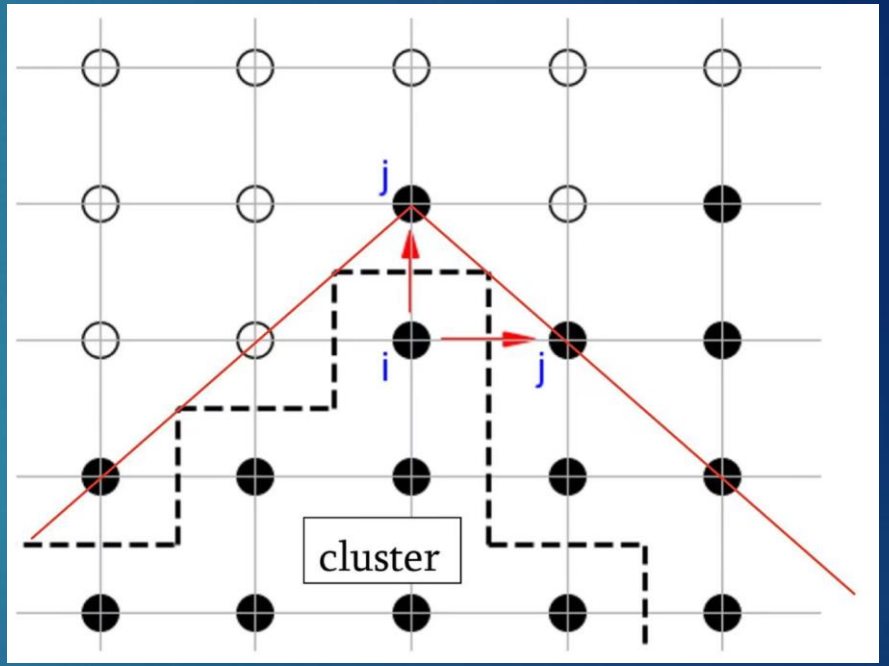
Restrict the cluster size

- A link is never activated if the site is out of the restriction;
- Build a cluster in the Wolff way.

Acceptance ratio:

$$A_{x\to y} = \min\left\{1, \frac{\rho(y)}{\rho(x)}\frac{\rho_{eff}(x)}{\rho_{eff}(y)}\prod\nolimits_{<ij>\in r} e^{-2\beta J S_i^x S_j^x}\right\}$$

Take N/A_r steps as one global update step when compared with SLMC method.



Restricted Self Learning Monte Carlo Method

Advantage:

- High acceptance ratio;
- Low correlation time;
- Can describe long range correlation;
- Comparatively short calculation time.

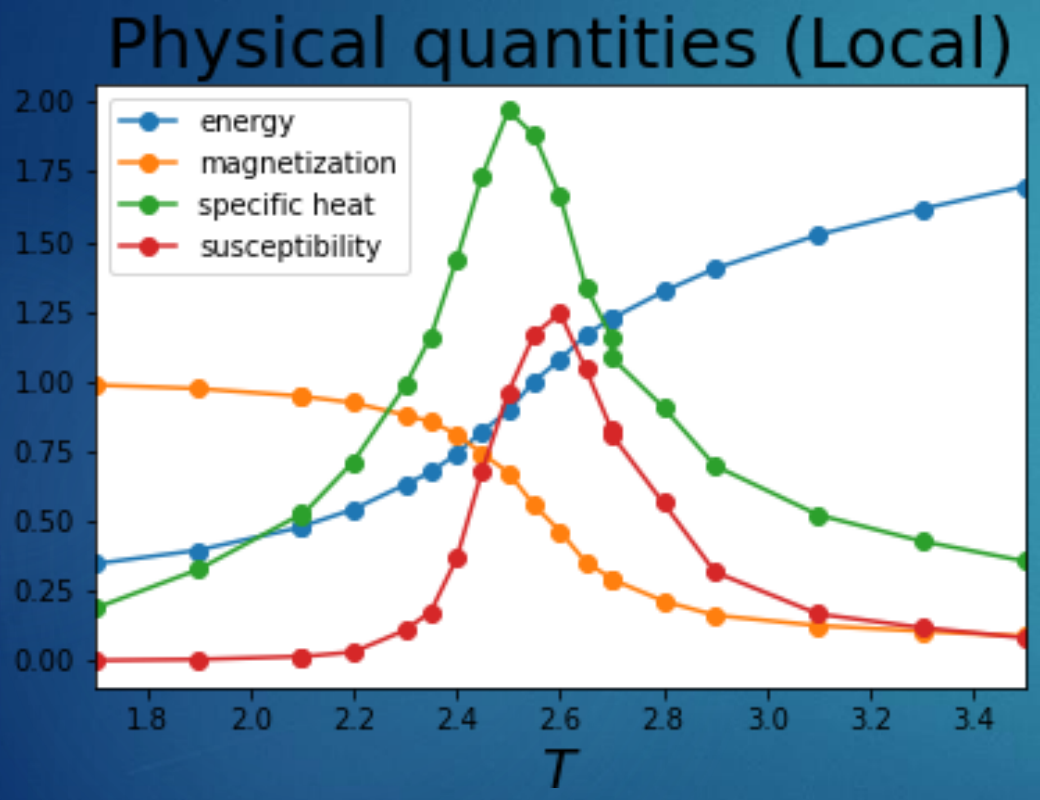
Drawback:

- Sacrifice a little bit of long range correlation outside the restriction area.

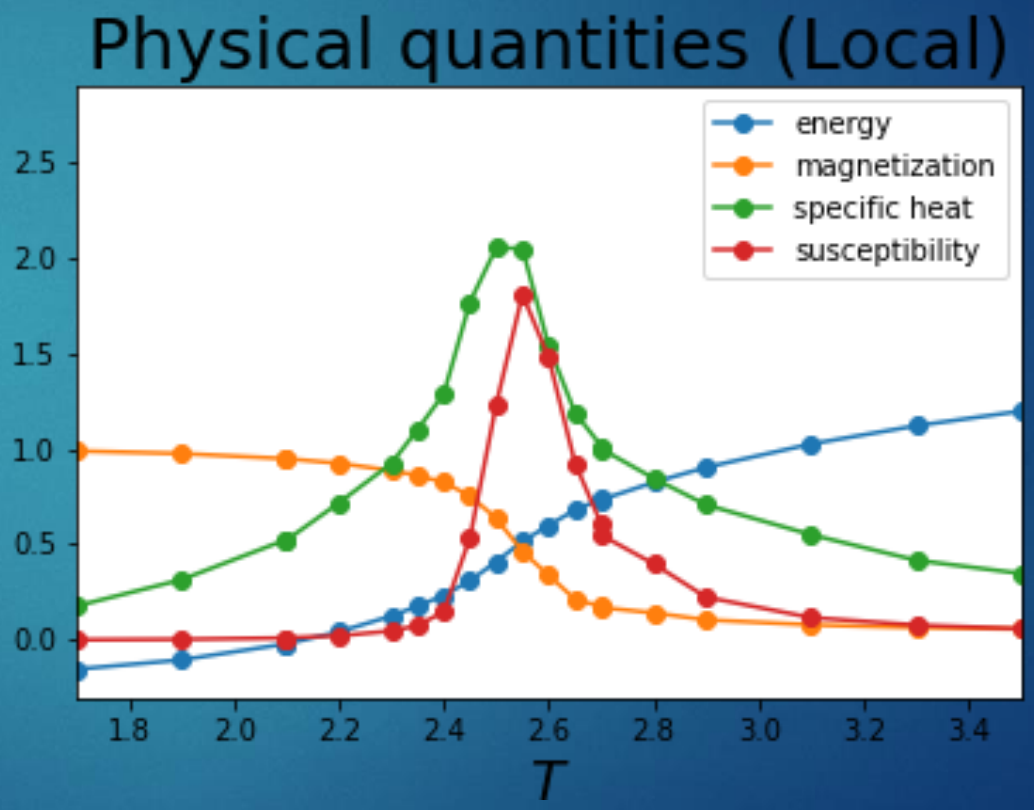
Part III: Results and Discussions

Overview of the system status vs. temperature

L = 25



L=40

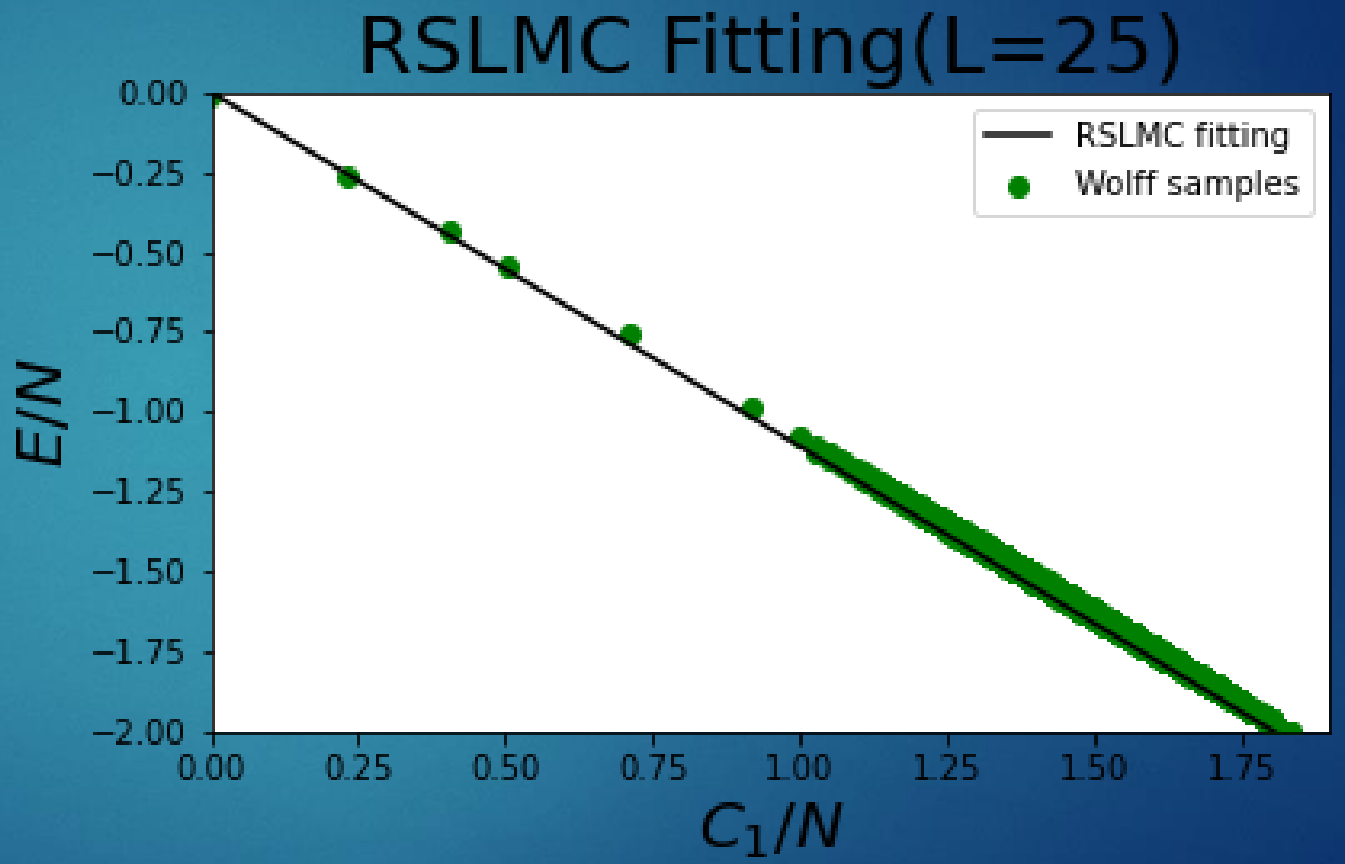


Training results of effective parameters

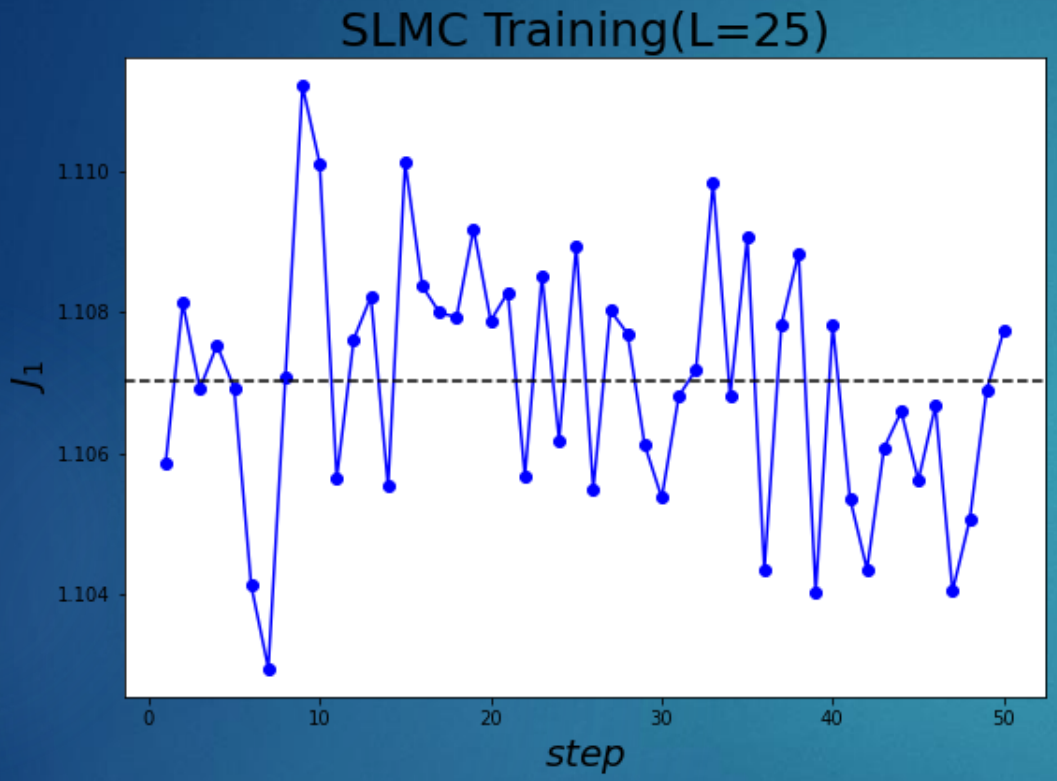
J_1 dominates in magnitude;
J_2 and J_3 contribute little to the accuracy of H_{eff}

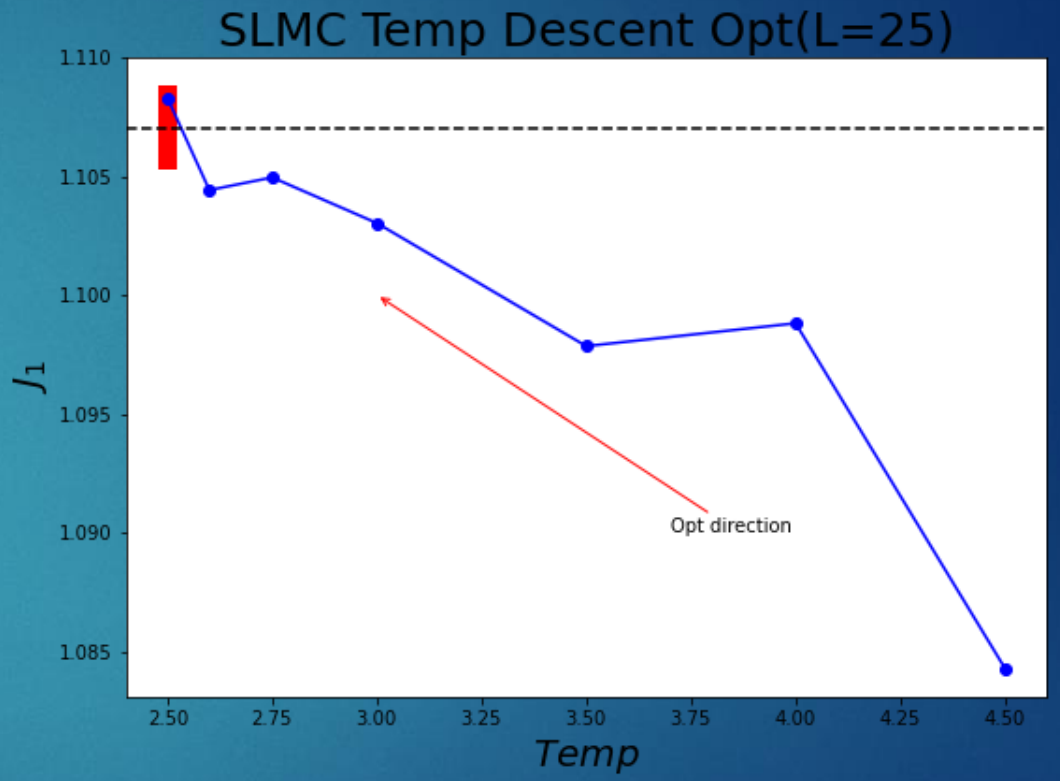
Size	Method	J_1	J_2	J_3	Mean error
L=25	SLMC	1.2454	-0.0912	-0.0143	0.0026
		1.1070			0.0025
	RSLMC	1.2558	-0.0850	-0.0850	0.0026
		1.1073			0.0026
L=40	SLMC	1.2448	-0.0862	-0.0134	
		1.1071			
	RSLMC	1.1068			

We choose to keep the NN interaction. Then the coefficient J1 is the target of traning. The figure of fitting shows that the relation between the energies of samples and the sum of NN spin produtes are very close to a linear one.



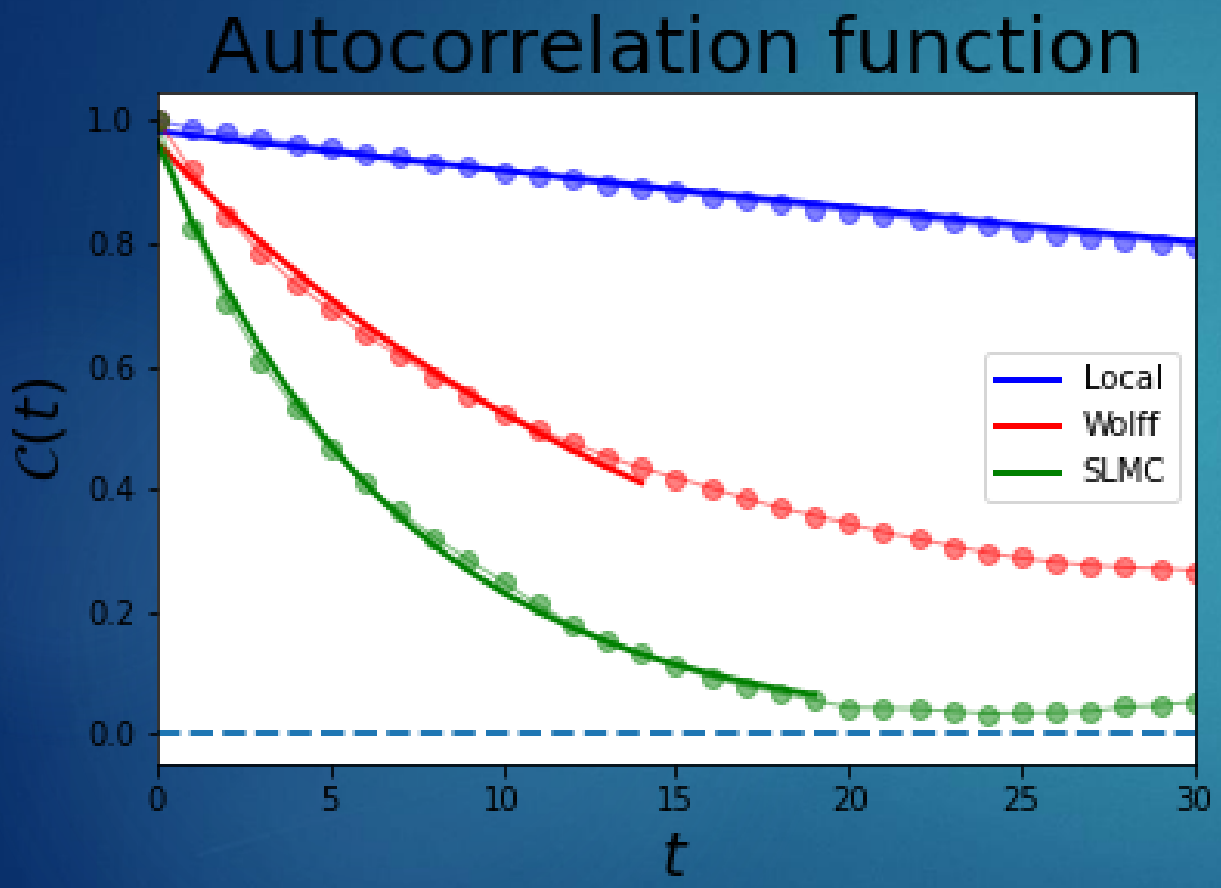
One step from T=5 to T=2.5: is that OK?





The differernce between J1 at T=5 and J1 at T =2.5 is quite small. Heff from T=5 converges quickly during the iteration at T=2.5. The final results are also close. (1.0701 vs 1.0828)

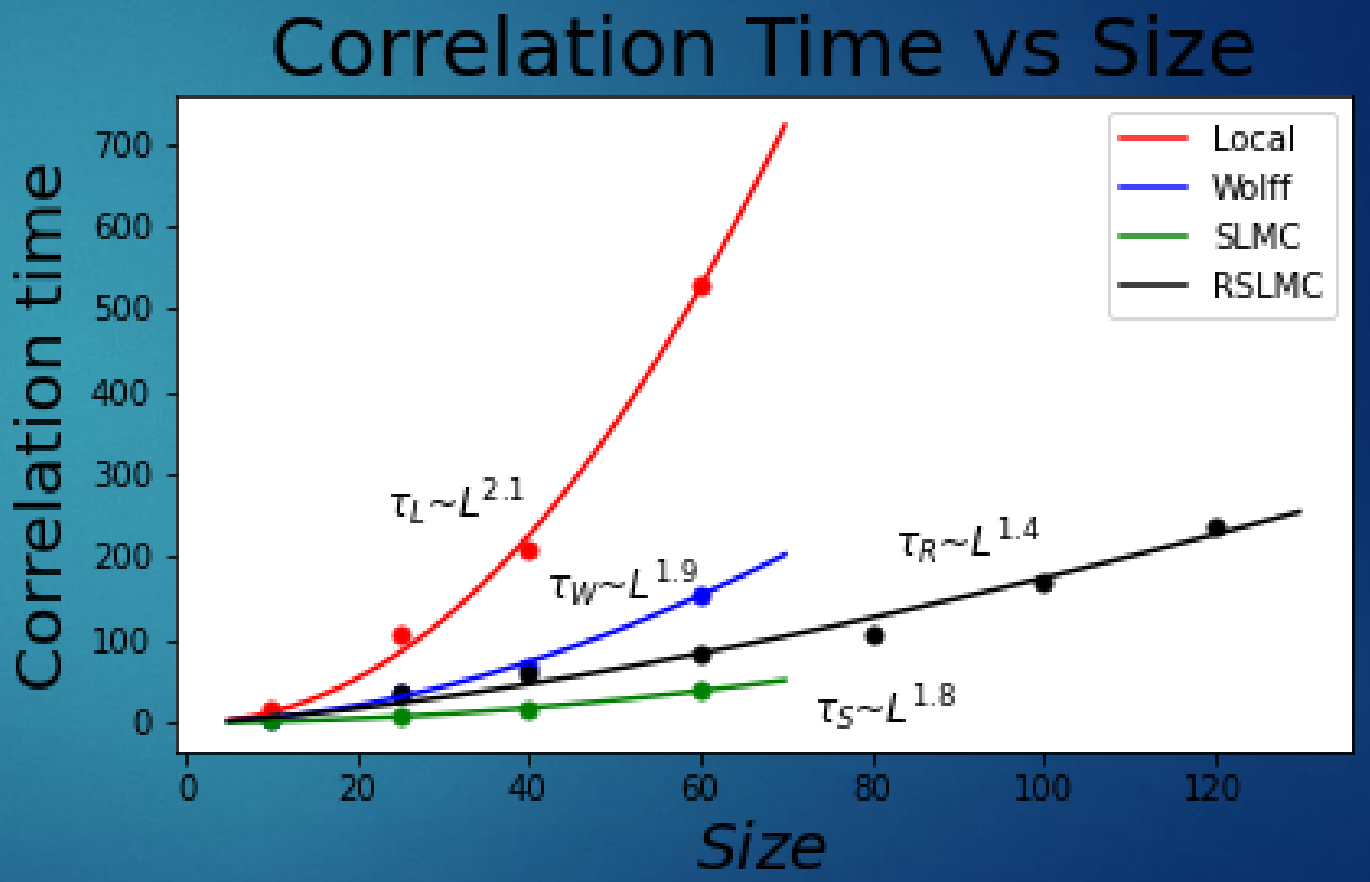
Autocorrelation function of different methods



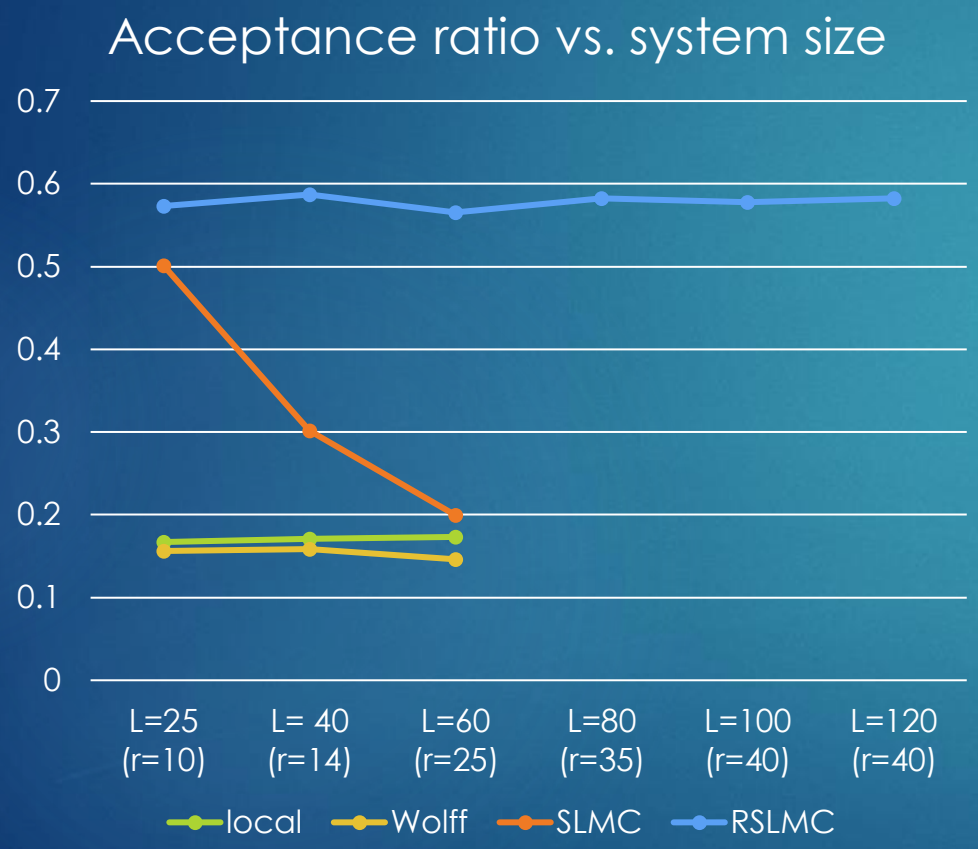
- Fit the autocorrelation functions to exponential ones. The three correlation times are
- Global updates have great advantage over local updates at Tc.
- SL made much improvements to the global update method.

Increase of correlation times with L

- Index for the correlation time of RSLMC ($\alpha_R = 1.4$) is significantly smaller than three others ($\alpha_L = 2.1$, $\alpha_W = 1.9$, $\alpha_S = 1.8$).
- Generally, global updates perform better than local updates when dealing with long-length components.
- Even normalized correlation time cannot accurately show the computing speed. RSLMC runs faster than SLMC under small sizes, even though its correlation time is larger.



Acceptance ratios of the clusters in different systems



- Correction of Hamiltonian in the ratio comes from the boundary of the cluster. Bigger L, larger cluster size, more errors for Hamiltonian, then lower acceptance ratio for SLMC.
- RSLMC makes its acceptance ratios nearly unchanged with L by restricting the cluster size.
- Puzzle1: Why the same mechanism doesn't work for Wolff updates?
- Puzzle2:Why the correlation time of Wolff updates grows faster with L than SLMC does, even though its acceptance ratios are less sensitive to L?

Part IV: Conclusions and Outlook

Conclusions

- We increased the accuracy of effective Hamiltonian by machine-learning. The mean error is now below 0.3%, which means higher acceptance ratios of cluster and a better match between the clusters and system components. These two improvements both contribute to a shorter correlation time.

- Most amazingly, we still only have to consider NN interactions. The same computing amount compared to naive Wolff updates!

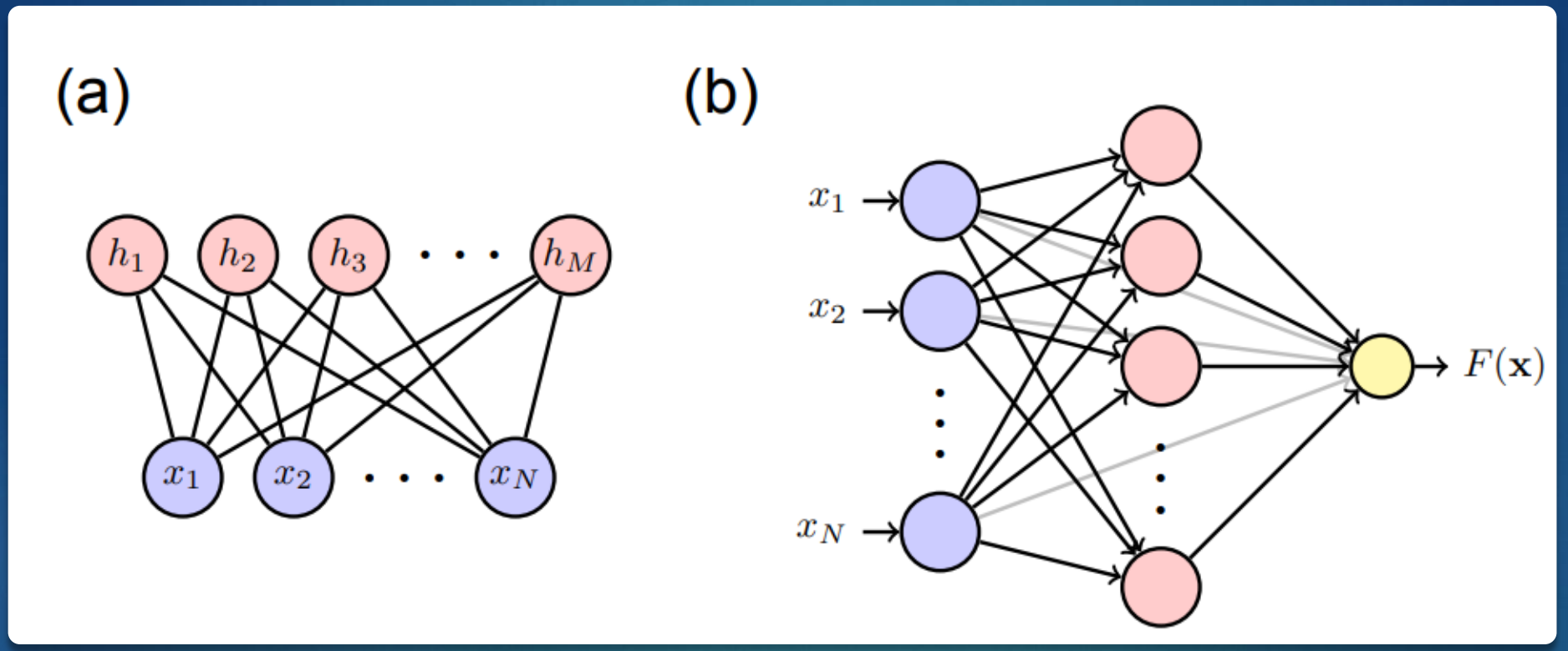
- By applying the restriction of cluster sizes to SLMC, RSLMC can adapt to large systems. We reduce the increase of the correlation time from $L^{1.8}$ to $L^{1.4}$.

Outlook

- More than Ising model: develop global update methods for strongly correlated systems with many-body interactions, like Fermion systems

- Mutual benefits: original Hamiltonian is not always exactly known. By studying the system with SLMC, we can improve the theory.

Outlook



a) Input variables → items of the result

b)Input variables → hidden variable + input variables → result

The method we used to optimize H_{eff} is still naive. The neural network may fit better for this problem.

References

- Junwei Liu, Yang Qi, Zi Yang Meng, and Liang Fu. Self-learning monte carlo method. *Phys. Rev. B*, 95:041101, Jan 2017.

- Li Huang and Lei Wang. Accelerated monte carlo simulations with restricted boltzmann machines. *Phys. Rev. B*, 95:035105, Jan 2017.

Thank you for watching!

Welcome questions!

JaySchon/PHY571-Project: PHY571 Project: Self Learning Monte Carlo Method (github.com)