

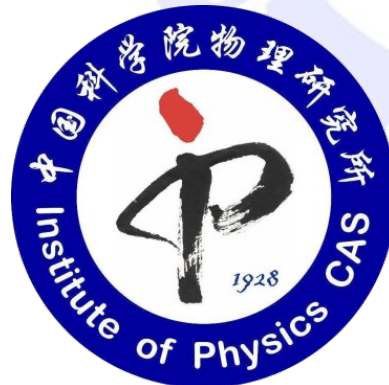
# Self-learning Monte Carlo Method

Zi Yang Meng

( 孟子杨 )

Institute of Physics, Chinese Academy of Sciences

<http://ziyangmeng.iphy.ac.cn/>



# Know thyself



**"Know thyself"** (Greek: γνῶθι σεαυτόν, gnothi seauton)

one of the Delphic maxims and was inscribed in the pronaos (forecourt) of the Temple of Apollo at Delphi



# Delphic Maxims



**"Know thyself"** (Greek: γνῶθι σεαυτόν, gnothi seauton). Thales of Miletus (c. 624 – c. 546 BC)

**"nothing in excess"** (Greek: μηδέν άγαν). Solon of Athens (c. 638 – 558 BC)

**"make a pledge and mischief is nigh"** (Greek: Έγγύα πάρα δ'άτη).

# Collaborators and References

---

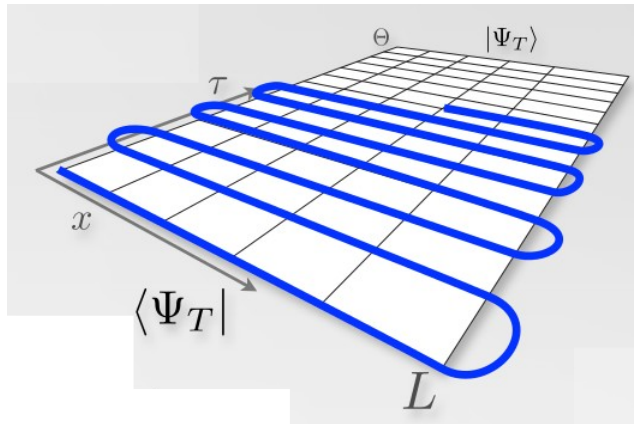
- Xiao Yan Xu, IOP, CAS
- Huitao Shen, Massachusetts Institute of Technology
- Jiuwei Liu, Massachusetts Institute of Technology
- Yang Qi, Massachusetts Institute of Technology & Fudan University, Shanghai
- Liang Fu, Massachusetts Institute of Technology

## Trilogy of SLMC

- Self-Learning Monte Carlo Method, arXiv:1610.08376
- Self-Learning Monte Carlo Method in Fermion Systems, arXiv:1611.09364
- Self-Learning Determinantal Quantum Monte Carlo Method, arXiv:1612.03804

# Quantum Monte Carlo simulation

## ■ Determinantal QMC for fermions



Hubbard-like models:

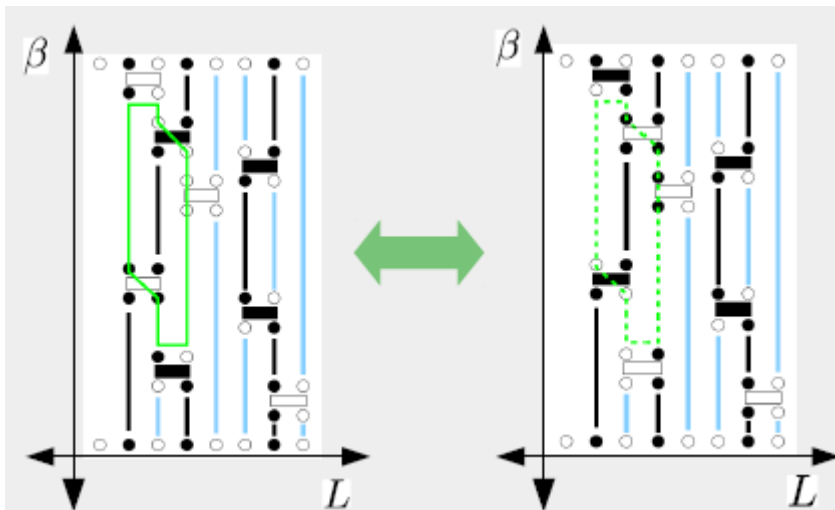
- Metal-Insulator transition
- Interaction effects on topological state of matter

Fermions coupled to critical bosonic mode:

- Itinerant quantum critical point
- Non-Fermi-liquid
- Gauge field couples to fermion

.....

## ■ World-line QMC for bosons



Heisenberg-like models:

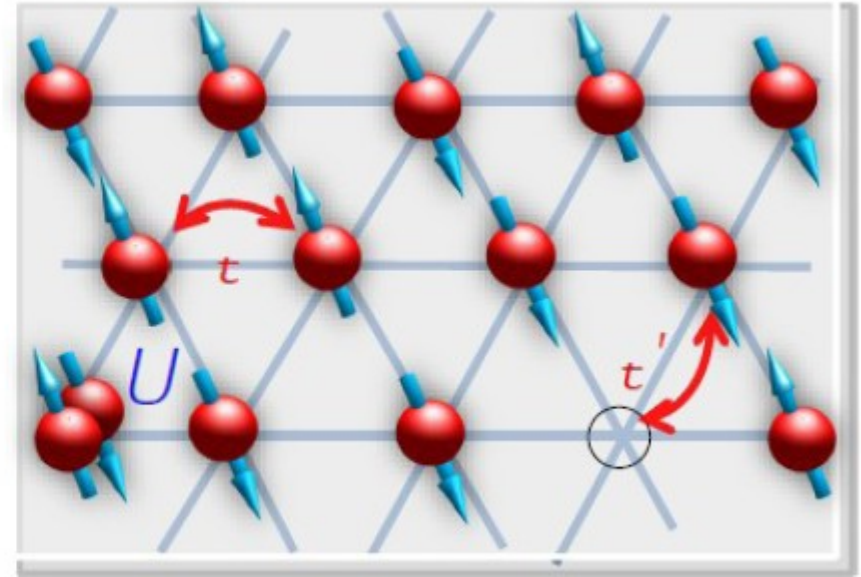
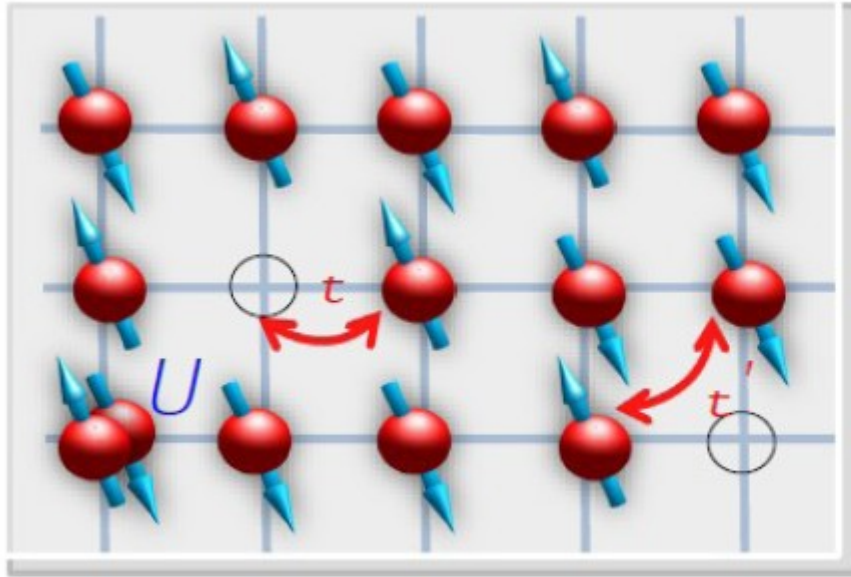
- Quantum magnetism
- Phase transition and critical phenomena
- Quantum spin liquids

Duality between DQCP and bosonic SPT:

- Deconfined quantum critical point
- Bosonic SPT and its critical point

.....

# Basic problem



Partition function: 
$$Z = \text{Tr} [e^{-\beta(\hat{H} - \mu\hat{N})}] = \sum_n \langle n | e^{-\beta(\hat{H} - \mu\hat{N})} | n \rangle$$

Observables: 
$$\langle \hat{A} \rangle = \frac{\text{Tr} [\hat{A} e^{-\beta(\hat{H} - \mu\hat{N})}]}{\text{Tr} [e^{-\beta(\hat{H} - \mu\hat{N})}]} = \frac{\sum_n \langle n | \hat{A} e^{-\beta(\hat{H} - \mu\hat{N})} | n \rangle}{\sum_n \langle n | e^{-\beta(\hat{H} - \mu\hat{N})} | n \rangle}$$

Fock space: 
$$\{|n\rangle\} \sim 2^{N_e} (e^{N_e \ln(2)}) \quad 4^{N_e} (e^{N_e \ln(4)})$$

# Monte Carlo simulation

---

- Widely used in statistical and quantum many-body physics
- Unbiased: statistical error  $1/\sqrt{N}$
- Universal: applies to any model without sign problem

$$Z = \sum_{\mathcal{C}} e^{-\beta H[\mathcal{C}]} = \sum_{\mathcal{C}} W(\mathcal{C})$$

- Markov chain Monte Carlo is a way to do important sampling

$$\cdots \rightarrow \mathcal{C}_{i-1} \rightarrow \mathcal{C}_i \rightarrow \mathcal{C}_{i+1} \rightarrow \cdots$$

- Distribution of  $\mathcal{C}$  converges to the Boltzmann distribution  $W(\mathcal{C})$
- Observable can be measured from a Markov chain

$$\langle O \rangle = \frac{\sum_{\mathcal{C}} O(\mathcal{C}) W(\mathcal{C})}{\sum_{\mathcal{C}} W(\mathcal{C})} = \frac{1}{\mathcal{N}} \sum_i O(\mathcal{C}_i)$$

# Autocorrelation time

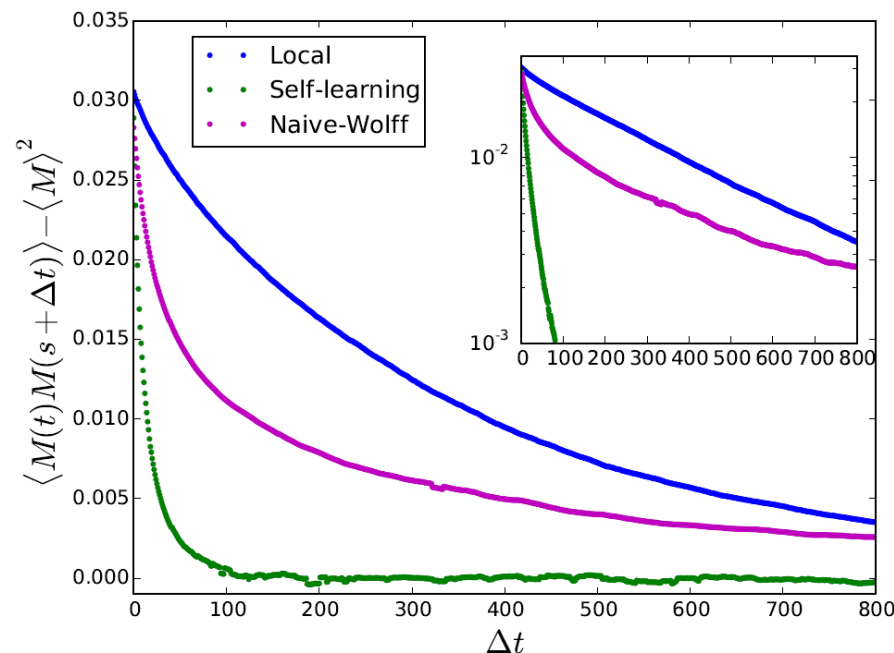
- Markov process, Monte Carlo time sequence

$$\dots \rightarrow O(t-1) \rightarrow O(t) \rightarrow O(t+1) \rightarrow \dots$$

$$O(t) = O[\mathcal{C}(t)]$$

- Autocorrelation function

$$A_O(\Delta t) = \langle O(t)O(t+\Delta t) \rangle - \langle O(t) \rangle^2 \propto e^{-\Delta t/\tau}$$





# Monte Carlo simulation

---

$$\cdots \rightarrow \mathcal{C}_{i-1} \rightarrow \mathcal{C}_i \rightarrow \mathcal{C}_{i+1} \rightarrow \cdots$$

- Detailed balance guarantees the Markov process converges to desired distribution

$$\frac{p(\mathcal{C} \rightarrow \mathcal{D})}{p(\mathcal{D} \rightarrow \mathcal{C})} = \frac{W(\mathcal{D})}{W(\mathcal{C})}$$

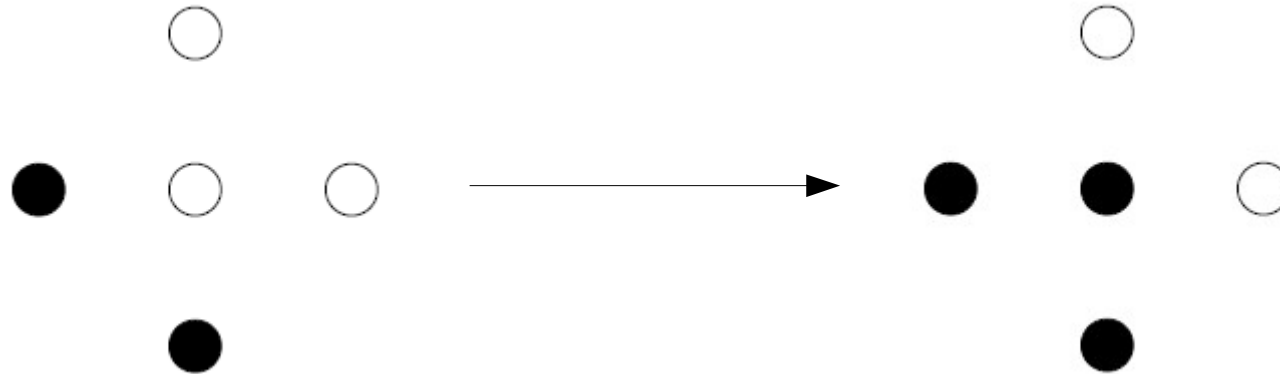
- Metropolis-Hastings algorithm: proposal – acceptance/rejection

$$p(\mathcal{C} \rightarrow \mathcal{D}) = q(\mathcal{C} \rightarrow \mathcal{D})\alpha(\mathcal{C} \rightarrow \mathcal{D})$$

$$\alpha(\mathcal{C} \rightarrow \mathcal{D}) = \min\left\{1, \frac{W(\mathcal{D})q(\mathcal{D} \rightarrow \mathcal{C})}{W(\mathcal{C})q(\mathcal{C} \rightarrow \mathcal{D})}\right\}$$

- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953)
- W. H. Hastings, Biometrika **57**, 97 (1970)

# Metropolis algorithm: local update



● Local update  $q(\mathcal{C} \rightarrow \mathcal{D}) = q(\mathcal{D} \rightarrow \mathcal{C}) = \frac{1}{N}$

● Acceptance ratio  $\alpha(\mathcal{C} \rightarrow \mathcal{D}) = \min\left\{1, \frac{W(\mathcal{D})}{W(\mathcal{C})}\right\}$

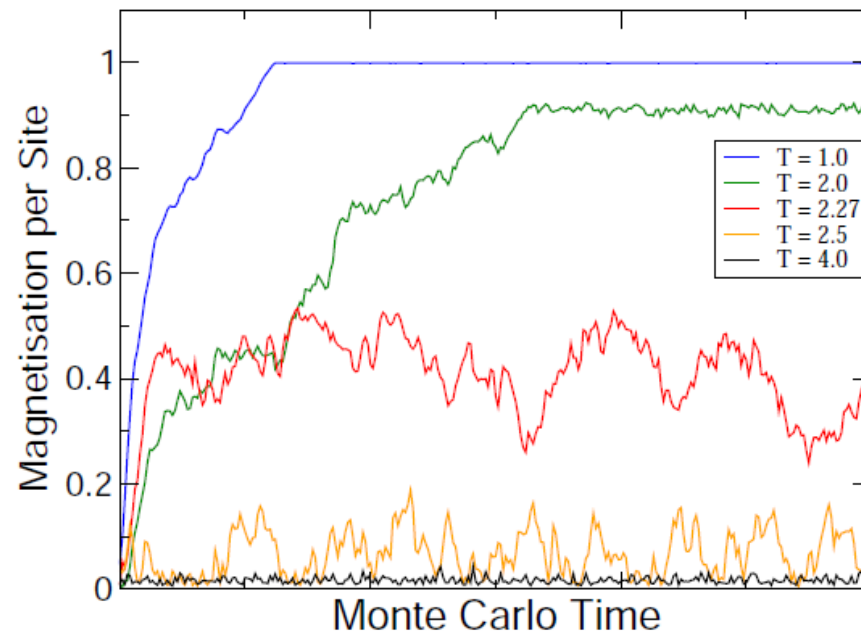
$$\frac{W(\mathcal{D})}{W(\mathcal{C})} = e^{-\beta(E(\mathcal{D}) - E(\mathcal{C}))} = e^{-\beta([-1-1+1+1] - [1+1-1-1])} = 1$$

➤ N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953)

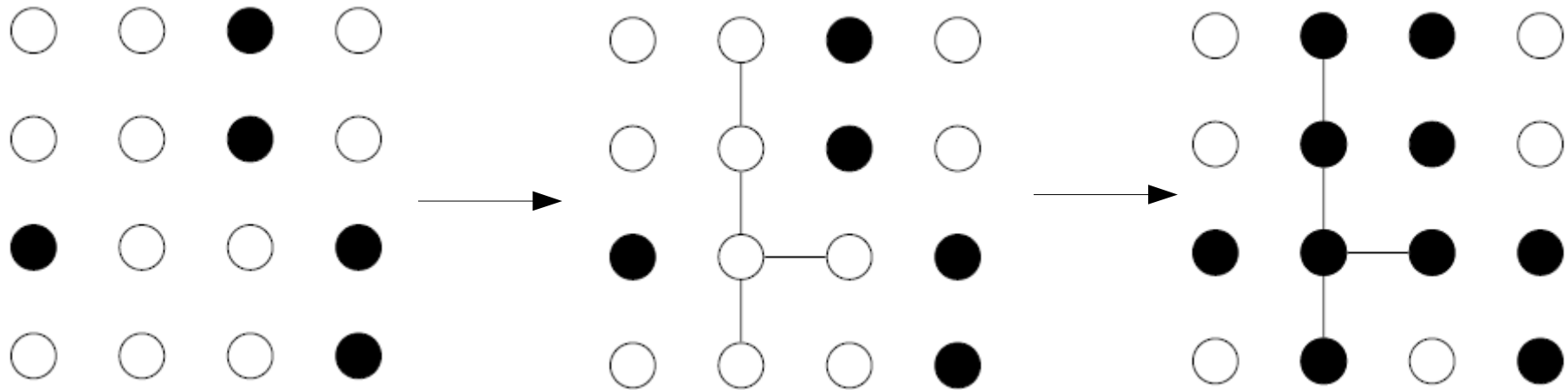
# Critical slowing down

- Dynamical relaxation time diverges at the critical point: critical system is slow to equilibrate.
- For 2D Ising model  $\tau \propto L^z, z = 2.125$

Metropolis Simulation on a 100x100 Grid



# Wolff algorithm: cluster update



- A cluster is built from bonds
- Probability of activating a bond is cleverly designed

$$q(i \rightarrow j) = 1 - e^{\min\{0, -2\beta S_i S_j\}}$$

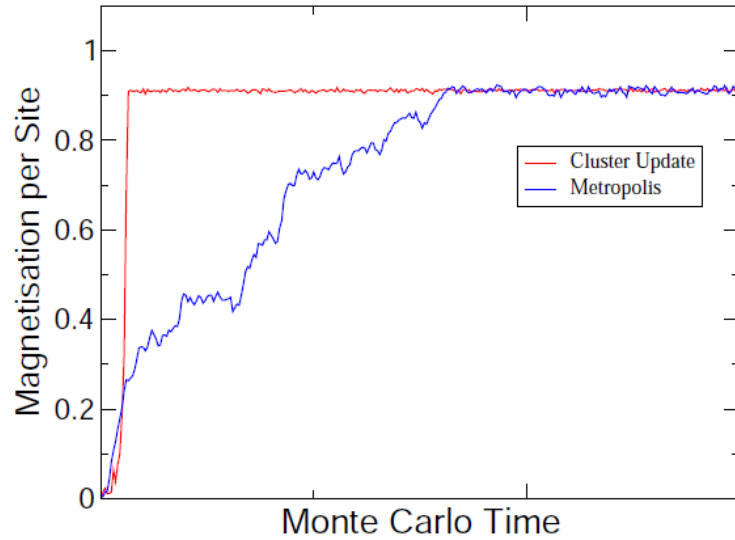
$$\frac{q(\mathcal{A} \rightarrow \mathcal{B})}{q(\mathcal{B} \rightarrow \mathcal{A})} = \prod_{\langle i, j \rangle, i \in c, j \notin c} \frac{1 - q(i \rightarrow j)_{\mathcal{A}}}{1 - q(i \rightarrow j)_{\mathcal{B}}} = \prod_{\langle i, j \rangle, i \in c, j \notin c} e^{-2\beta(S_i^{\mathcal{A}} S_j^{\mathcal{A}} - S_i^{\mathcal{B}} S_j^{\mathcal{B}})} = \frac{W(\mathcal{B})}{W(\mathcal{A})}$$

- an ideal acceptance ratio  $\alpha(\mathcal{A} \rightarrow \mathcal{B}) = 1$

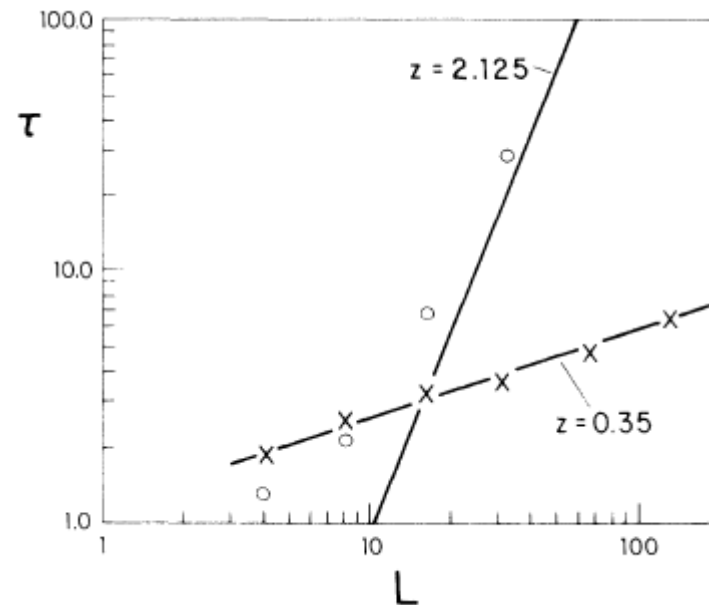
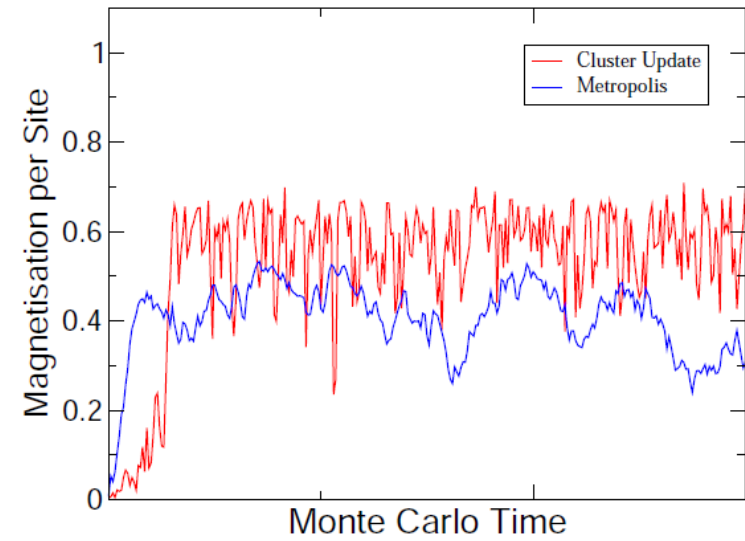
➤ U. Wolff, Phys. Rev. Lett. **62**, 361 (1989)

# Reduce critical slowing down

Simulations on a 100x100 Grid at  $T=2.0$



Simulations on a 100x100 Grid at  $T=2.27$

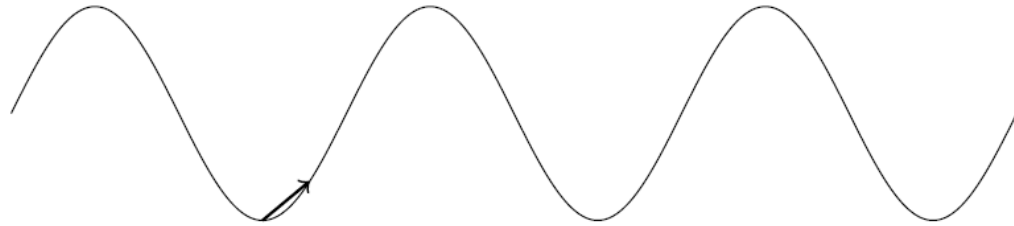


➤ Swendsen and Wang, Phys. Rev. Lett. **58**, 86 (1987)

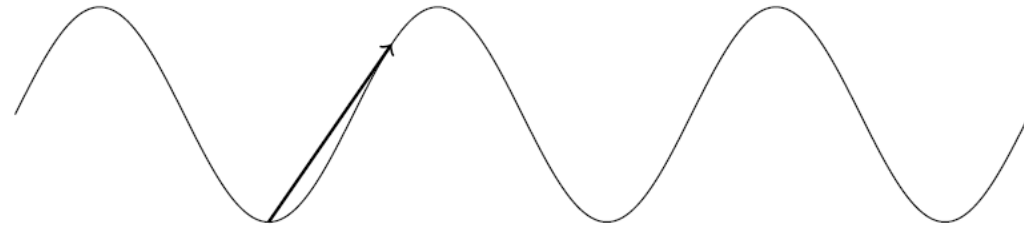
# Learn thyself

---

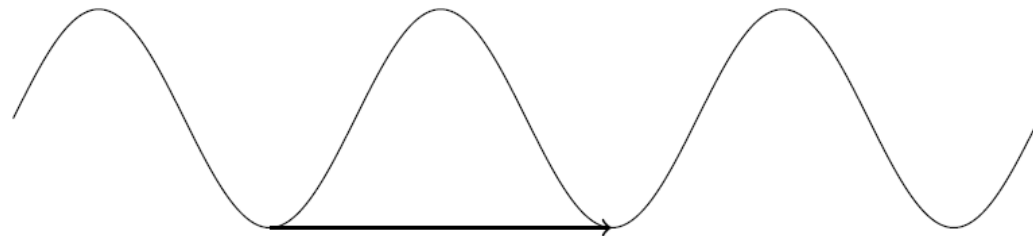
- Step too small: small difference, high acceptance



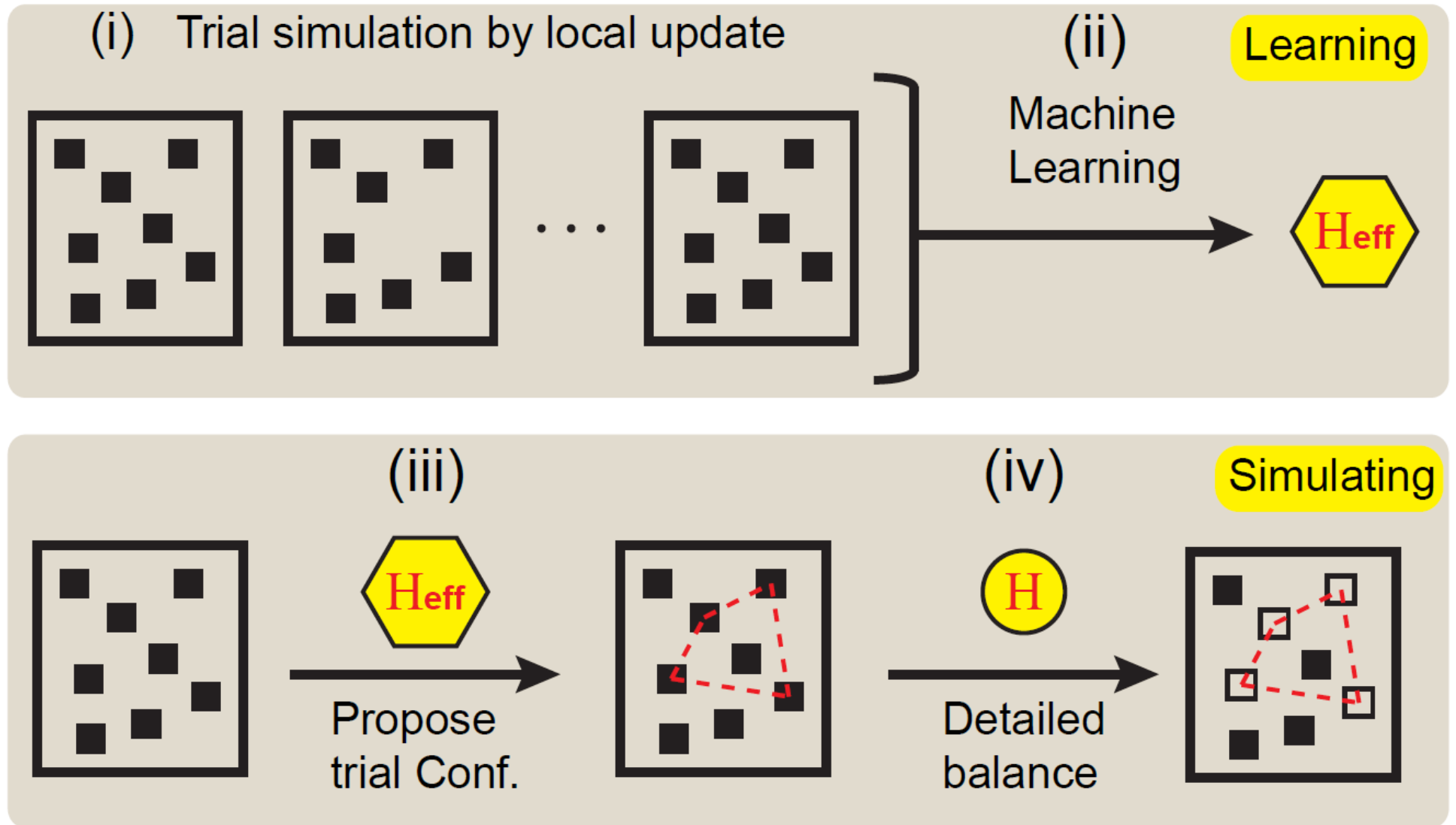
- Step too large: big difference, low acceptance



- Global update: explore the low-energy configurations



# SLMC: Learning+Simulating



# Trilogy I: SLMC for Bosons

$$H = -J \sum_{\langle ij \rangle} S_i S_j - K \sum_{ijkl \in \square} S_i S_j S_k S_l \quad K/J = 0.2$$

Ising transition with  $T_c = 2.493$

$$H_{\text{eff}} = E_0 - \tilde{J}_1 \sum_{\langle ij \rangle_1} S_i S_j - \tilde{J}_2 \sum_{\langle ij \rangle_2} S_i S_j - \dots$$

- The self-learning update: cluster is constructed using the effective model

$$\frac{q(\mathcal{C} \rightarrow \mathcal{D})}{q(\mathcal{D} \rightarrow \mathcal{C})} = \frac{W_{\text{eff}}(\mathcal{D})}{W_{\text{eff}}(\mathcal{C})}$$

- The acceptance ratio:

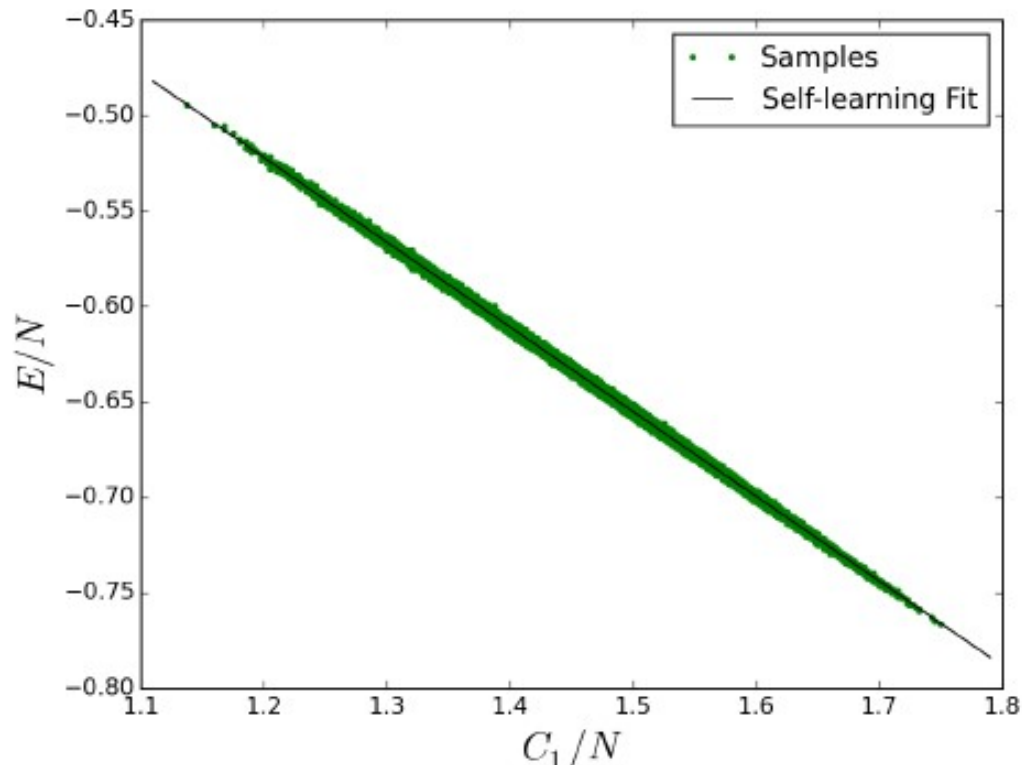
$$\alpha(\mathcal{C} \rightarrow \mathcal{D}) = \min\left\{1, \frac{W(\mathcal{D})W_{\text{eff}}(\mathcal{D})}{W(\mathcal{C})W_{\text{eff}}(\mathcal{C})}\right\} = \min\left\{1, e^{-\beta[(E(\mathcal{D}) - E_{\text{eff}}(\mathcal{D})) - (E(\mathcal{C}) - E_{\text{eff}}(\mathcal{C}))]}\right\}$$

- The acceptance ratio can be very high, autocorrelation time can be very short
- effective model capture the low-energy physics



# Trilogy I: SLMC for Bosons

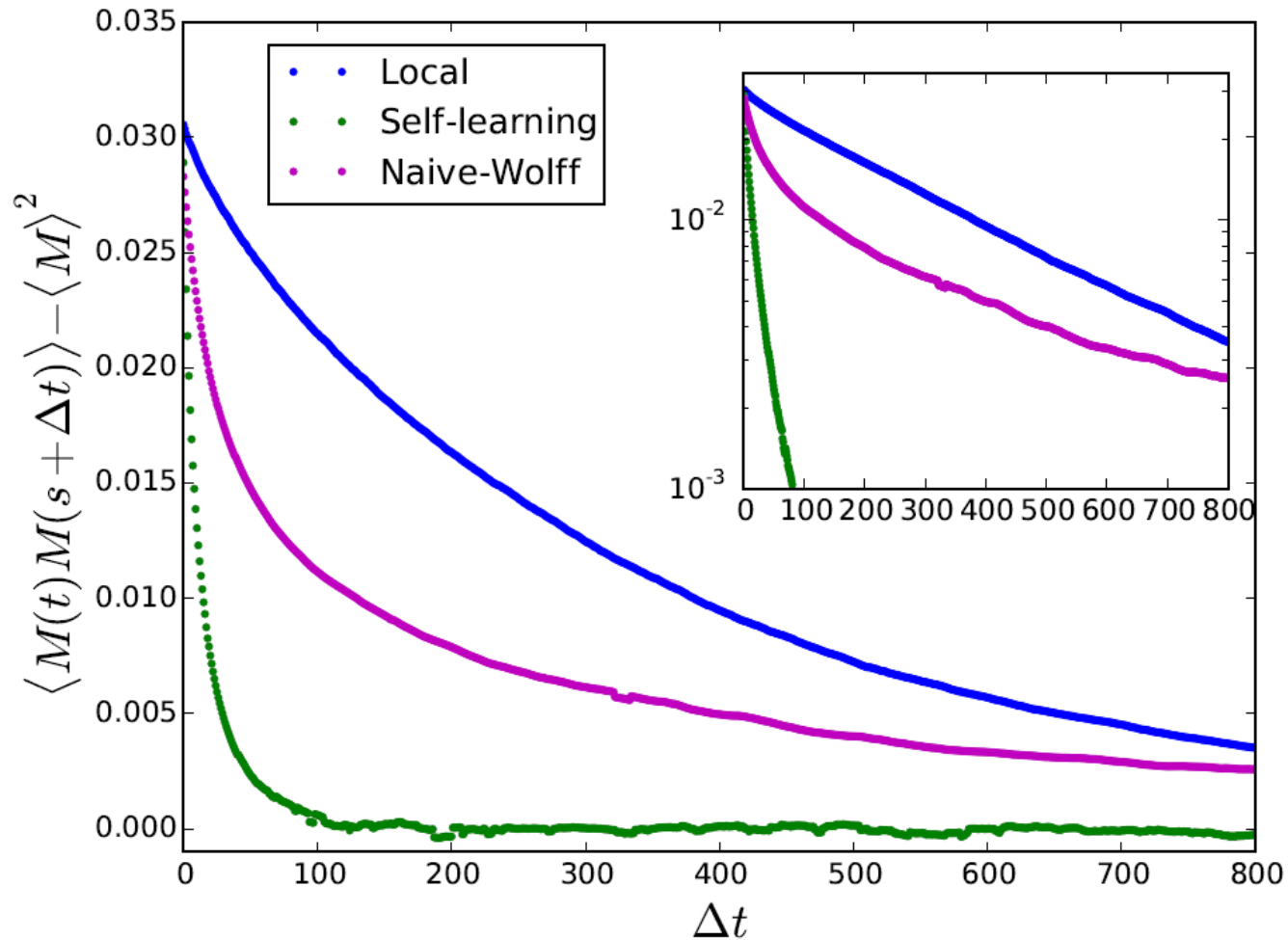
- Generate configurations with local update, at  $T=5 > T_c$ .
- Perform linear regression
- Generate configurations with reinforced learning at  $T_c$



$$C_1 = \sum_{\langle i,j \rangle} S_i S_j$$

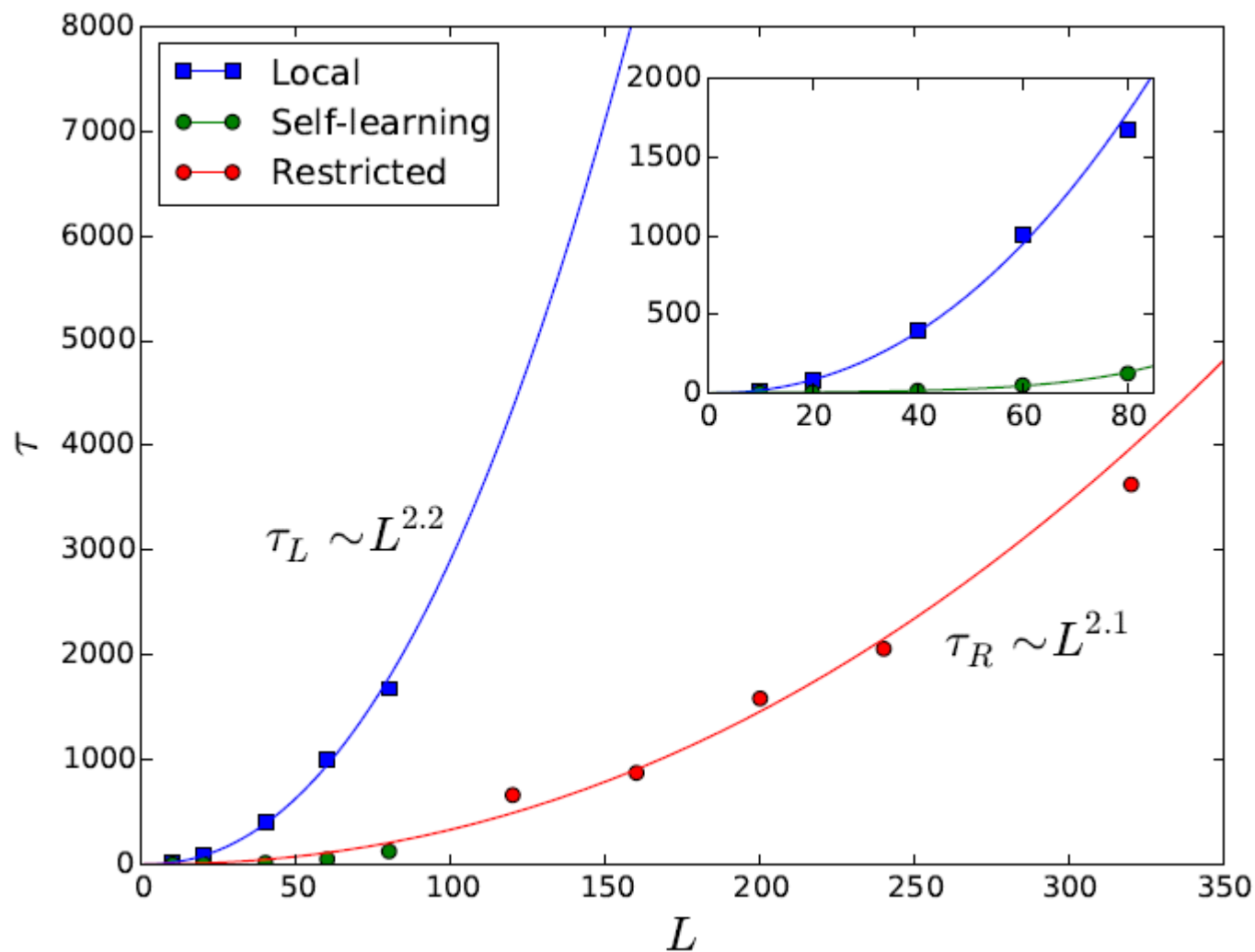
	$J_1$	$J_2$	$J_3$
Train 1	1.2444	-0.0873	-0.0120
Train 2	1.1064	-	-

# Trilogy I: SLMC for Bosons



System size  $40 \times 40$  at  $T_c$

# Trilogy I: SLMC for Bosons



● Speedup of 10~20 times

# Trilogy II: SLMC for Fermions

- Double exchange model

$$\hat{H} = -t \sum_{\langle ij \rangle, \alpha} (\hat{c}_{i\alpha}^\dagger \hat{c}_{j\alpha} + \text{h.c.}) - \frac{J}{2} \sum_{i, \alpha, \beta} \vec{S}_i \cdot \hat{c}_{i\alpha}^\dagger \vec{\sigma}_{\alpha\beta} \hat{c}_{i\beta}$$

$$Z = \sum_{\phi} \det \left[ \mathbf{I} + e^{-\beta H_f[\phi]} \right] \equiv \sum_{\phi} W[\phi]$$

- Computational complexity

$$O(\tau_0 \times L^{3d} \times L^d) = O(\tau_0 \times L^{4d})$$

- Fit effective model

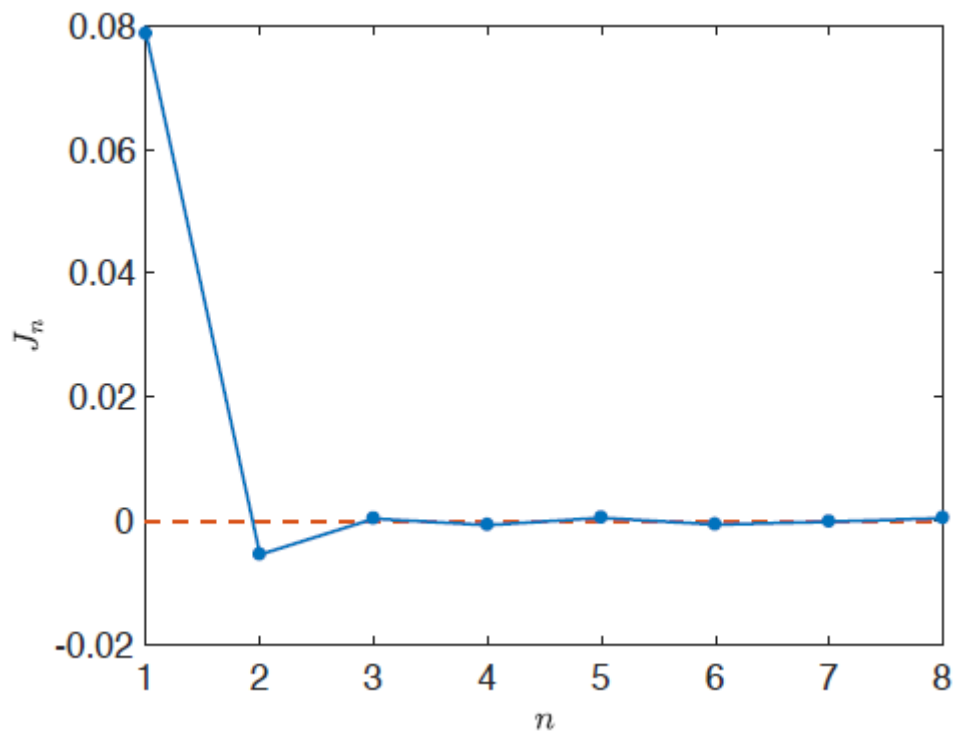
$$W[\phi] \simeq e^{-\beta H_{\text{eff}}[\phi]}$$

$$H_{\text{eff}} = E_0 - J_1 \sum_{\langle ij \rangle_1} \vec{S}_i \cdot \vec{S}_j - J_2 \sum_{\langle ij \rangle_2} \vec{S}_i \cdot \vec{S}_j - \dots$$

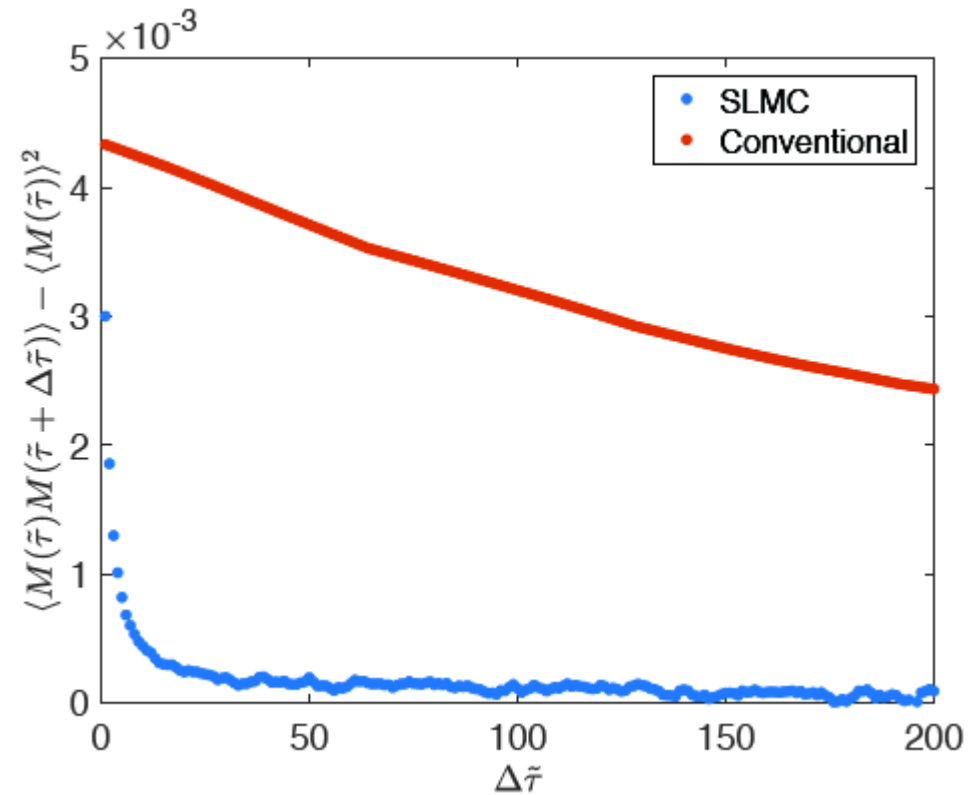
# Trilogy II: SLMC for Fermions

- effective model captures the low-energy physics, RKKY interaction.
- only need to learn from small system sizes

$$L = 4$$



$$L = 4$$



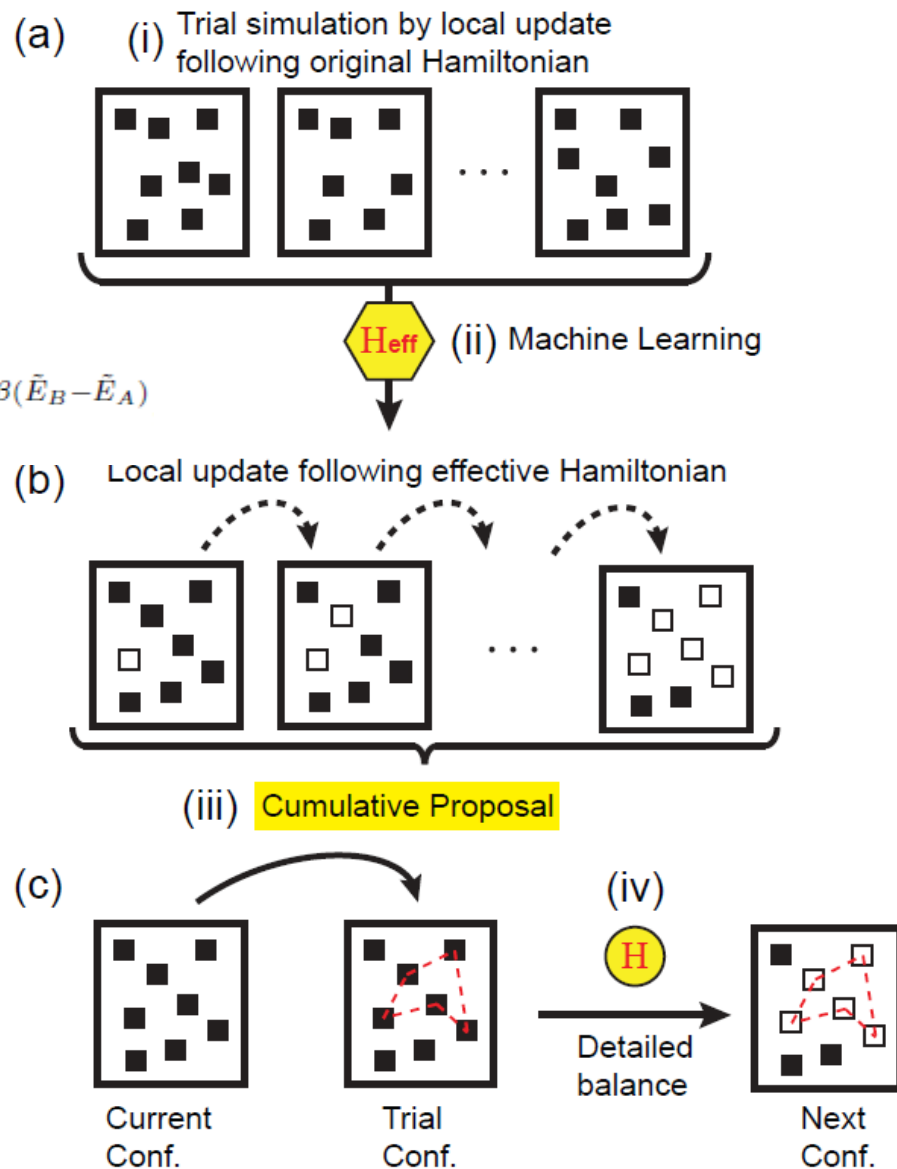
# Trilogy II: SLMC for Fermions

- Cumulative update

$$\frac{q(\mathcal{C} \rightarrow \mathcal{D})}{q(\mathcal{D} \rightarrow \mathcal{C})} = \frac{W_{\text{eff}}(\mathcal{D})}{W_{\text{eff}}(\mathcal{C})}$$

$$\frac{S(A \rightarrow B)}{S(B \rightarrow A)} = \prod_{i=0}^{n_c-1} \frac{\tilde{P}(C_i \rightarrow C_{i+1})}{\tilde{P}(C_{i+1} \rightarrow C_i)} = \prod_{i=0}^{n_c-1} e^{-\beta(\tilde{E}_{i+1} - \tilde{E}_i)} = e^{-\beta(\tilde{E}_B - \tilde{E}_A)}$$

$$p(A \rightarrow B) = \min\{1, e^{-\beta(E_B - \tilde{E}_B) - (E_A - \tilde{E}_A)}\}$$

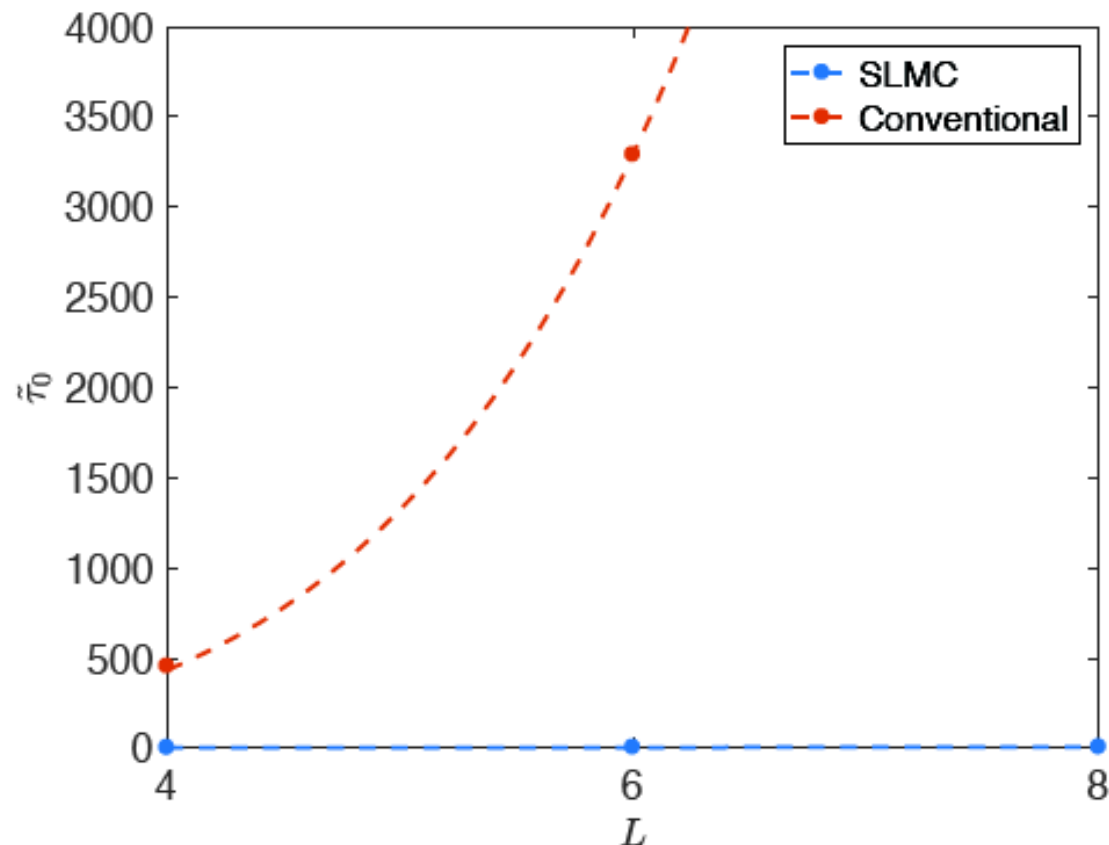


# Trilogy II: SLMC for Fermions

- Computation complexity at most

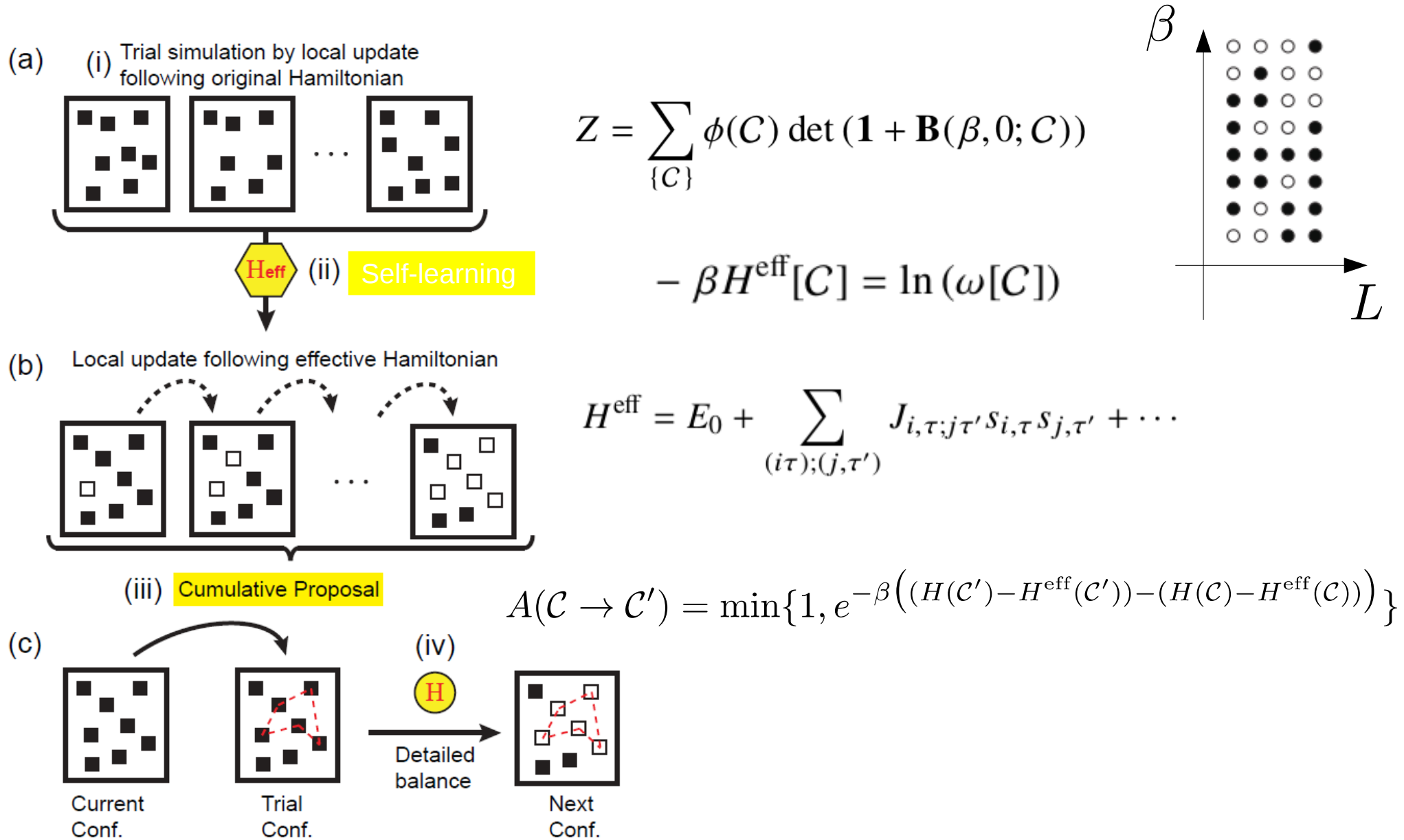
$$O(l_c) + O(\tau_0 \times L^{3d}) = O(\tau_0 \times L^{3d})$$

- Speedup of  $O(L^z L^d) = O(L^{d+z})$



- L=4,6,8, at L=8, 10<sup>3</sup> times faster.

# Trilogy III: SLMC for DQMC





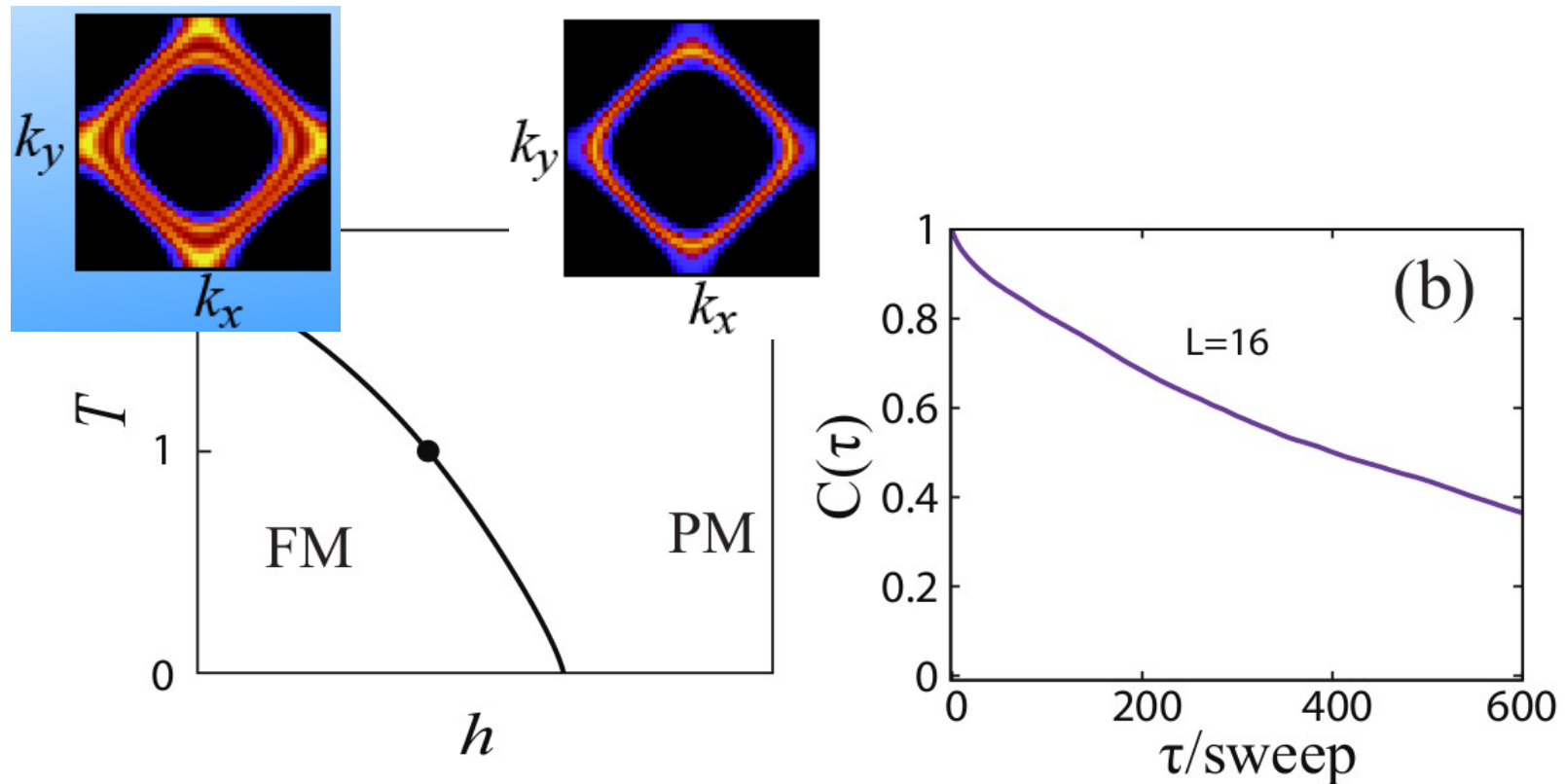
# Trilogy III: SLMC for DQMC

Fermions coupled to critical bosonic mode

- Itinerant quantum critical point
- Non-Fermi-liquid

➤ arXiv:1602.07150

➤ arXiv:1612.06075



Complexity for getting an independent configuration:  $\beta N^3 \tau_L$

# Trilogy III: SLMC for DQMC

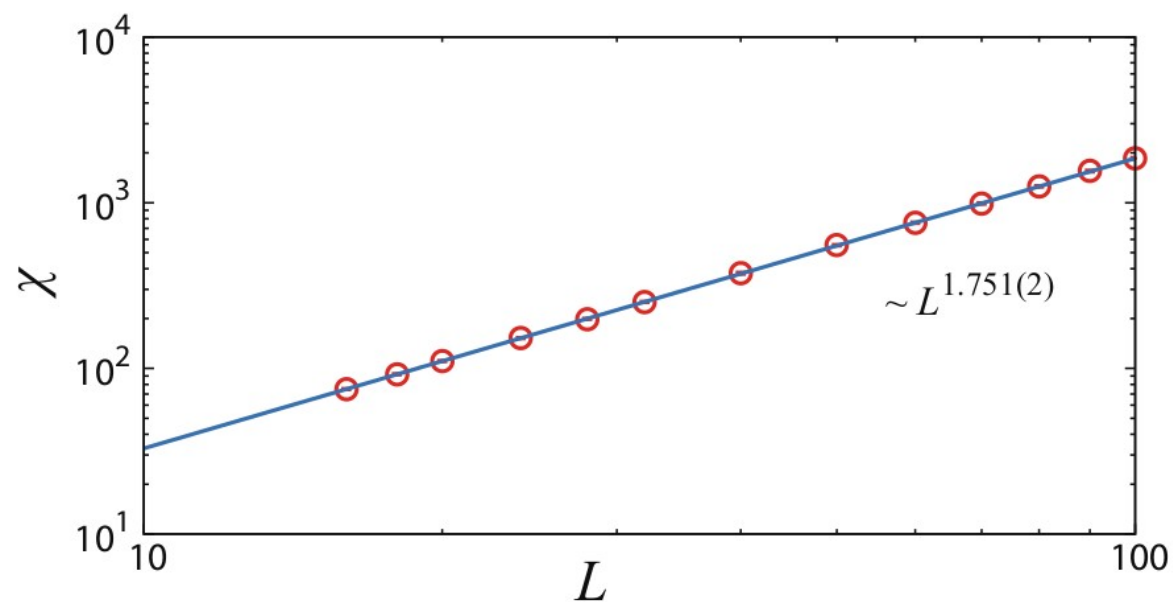
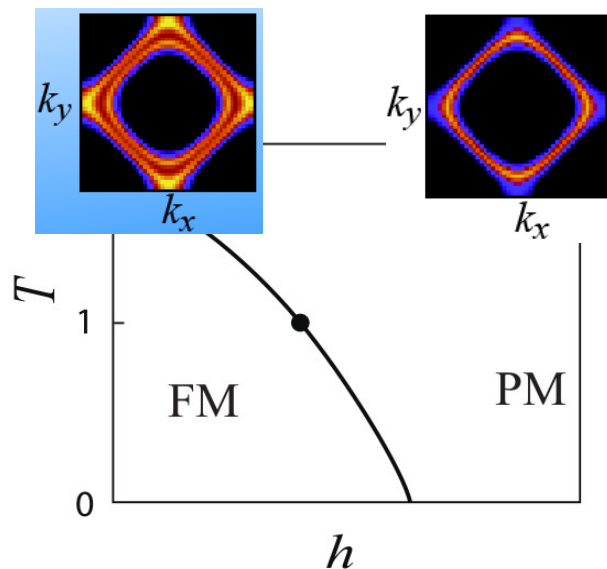
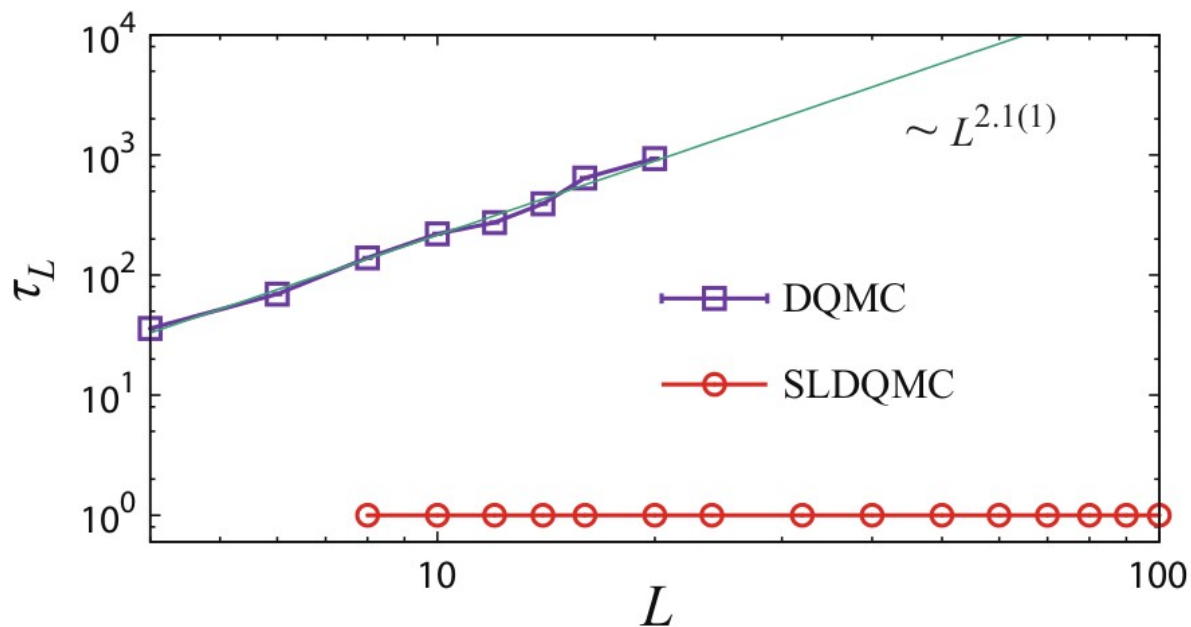
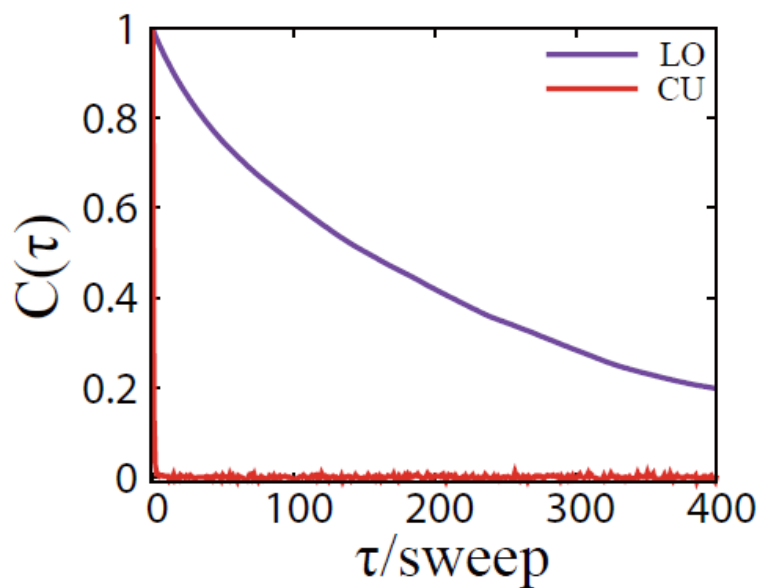
Complexity for obtaining an independent configuration:  $\beta N^3 \tau_L$

Complexity for SLMC

- Cumulative update:  $\gamma \beta N \tau_L$
- Detail balance: 
$$\begin{aligned} N^3 \omega_C &= \phi(\mathcal{C}) \det(\mathbf{1} + \mathbf{B}(\beta, \tau) \mathbf{B}(\tau, 0)) \\ &= \phi(\mathcal{C}) \det(\mathbf{G}(0, 0))^{-1} \end{aligned}$$
- Sweep Green's function:  $\beta N^2$

$$\text{Complexity speed up } \mathcal{S} = \min\left(\frac{N^2}{\gamma}, \beta \tau_L, N \tau_L\right)$$

# Trilogy III: SLMC for DQMC



# Advertisement

The screenshot shows the website for the Kavli Institute for Theoretical Sciences (KITS) at the University of Chinese Academy of Sciences (UCAS). The page is titled "Machine Learning and Many-Body Physics (Jun. 28th - Jul. 7th, 2017)". It features a navigation menu with links for home, About, Facilities, Research, News, Events, People, and Links. A search bar is located in the top right corner. The main content area lists the local coordinators and the international steering committee members.

KITS Kavli Institute for Theoretical Sciences  
University of Chinese Academy of Sciences (UCAS)

Site Map Login

Search ...

home About Facilities Research News Events People Links

Home / Events / Workshops / Machine Learning and Many-Body Physics (Jun. 28th - Jul. 7th, 2017)

Calendar  
Conferences  
Programs  
Schools  
**Workshops**  
Talks

## Machine Learning and Many-Body Physics (Jun. 28th - Jul. 7th, 2017)

© Published: 2017-02-20

### Local Coordinators

Lei Wang (Institute of Physics, CAS), [wanglei@iphy.ac.cn](mailto:wanglei@iphy.ac.cn)  
Zi Yang Meng (Institute of Physics, CAS), [zymeng@iphy.ac.cn](mailto:zymeng@iphy.ac.cn)  
Zhi-Yuan Xie (Renmin University of China), [qingtaoxie@ruc.edu.cn](mailto:qingtaoxie@ruc.edu.cn)

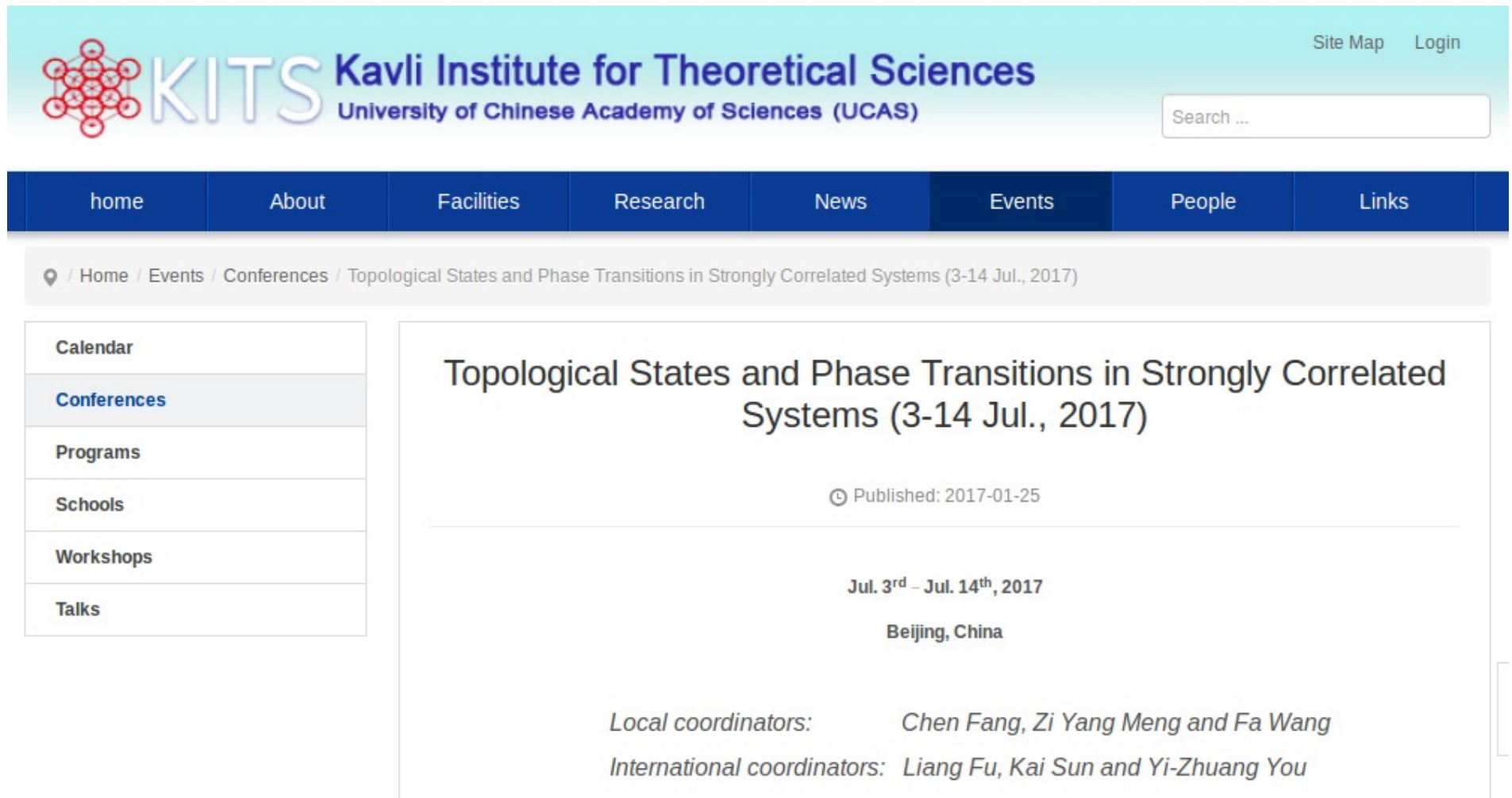
### International Steering Committee

Matthias Troyer (ETH Zurich and Microsoft Research)  
Roger Melko (Perimeter Institute)  
Hong Guo (McGill)  
Xi Dai (Institute of Physics, CAS)  
Tao Xiang (Institute of Physics, CAS),

## Scope of the Workshop

- Conceptual connections of machine learning and many-body physics
- Machine learning techniques for solving many-body physics/chemistry problems
- Quantum algorithms and quantum hardwares for machine learning

# Advertisement



The screenshot shows the website for Kavli Institute for Theoretical Sciences (KITS) at the University of Chinese Academy of Sciences (UCAS). The header includes the KITS logo, the name of the institute, and a search bar. A navigation menu contains links for home, About, Facilities, Research, News, Events, People, and Links. The breadcrumb trail indicates the current page is for a workshop titled "Topological States and Phase Transitions in Strongly Correlated Systems (3-14 Jul., 2017)". A sidebar on the left lists various categories: Calendar, Conferences (highlighted), Programs, Schools, Workshops, and Talks. The main content area features the workshop title, a publication date of 2017-01-25, the dates "Jul. 3<sup>rd</sup> - Jul. 14<sup>th</sup>, 2017", and the location "Beijing, China". It also lists local coordinators (Chen Fang, Zi Yang Meng, and Fa Wang) and international coordinators (Liang Fu, Kai Sun, and Yi-Zhuang You).

KITS Kavli Institute for Theoretical Sciences  
University of Chinese Academy of Sciences (UCAS)

Site Map Login

Search ...

home About Facilities Research News Events People Links

/ Home / Events / Conferences / Topological States and Phase Transitions in Strongly Correlated Systems (3-14 Jul., 2017)

Calendar  
Conferences  
Programs  
Schools  
Workshops  
Talks

## Topological States and Phase Transitions in Strongly Correlated Systems (3-14 Jul., 2017)

© Published: 2017-01-25

Jul. 3<sup>rd</sup> - Jul. 14<sup>th</sup>, 2017  
Beijing, China

*Local coordinators:* Chen Fang, Zi Yang Meng and Fa Wang  
*International coordinators:* Liang Fu, Kai Sun and Yi-Zhuang You

## Scope of the Workshop

- Topological classification of strongly correlated systems
- Topological phase transitions
- Realizations of topological orders

# 爱上物理 改变世界

中科院物理所 | 微信ID: cas-iop



长按二维码→自动识别→快速关注！

推荐文章或反馈意见可直接在公众号内留言  
投稿或法律相关事宜请通过邮件联系我们  
邮箱：[zhc@iphy.ac.cn](mailto:zhc@iphy.ac.cn)

## “认识你自己”——自学习蒙特卡洛三部曲

在这篇文章中，笔者只希望讲述我们最近发展的自学习蒙特卡洛方法三部曲，讲述我们如何通过自我观照、自我学习蒙特卡洛构型，设计出自学习蒙特卡洛方法，解决了凝聚态量子多体系统蒙特卡洛模拟中一些诸如临界慢化和接收概率低等瓶颈性的问题，推动领域的发展。

2016-12-22

