

Projector Quantum Monte Carlo Method for Nonlinear Wave Functions

非线性波函数的投影量子蒙特卡罗方法

原作者：Lauretta R. Schwarz, A. Alavi, George H. Booth

演讲人：安子鹏

目录



传统蒙特卡罗方法简介

非线性波函数的量子蒙特卡罗方法原理

结果讨论与对比

总结与思考

一、传统蒙特卡罗方法简介

核心问题: $i\hbar \frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle$

BO近似: $\hat{H}_{el} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_{\alpha} \frac{Z_{\alpha}}{r_{i\alpha}} + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{r_{ij}}$

HF方法



后HF方
法



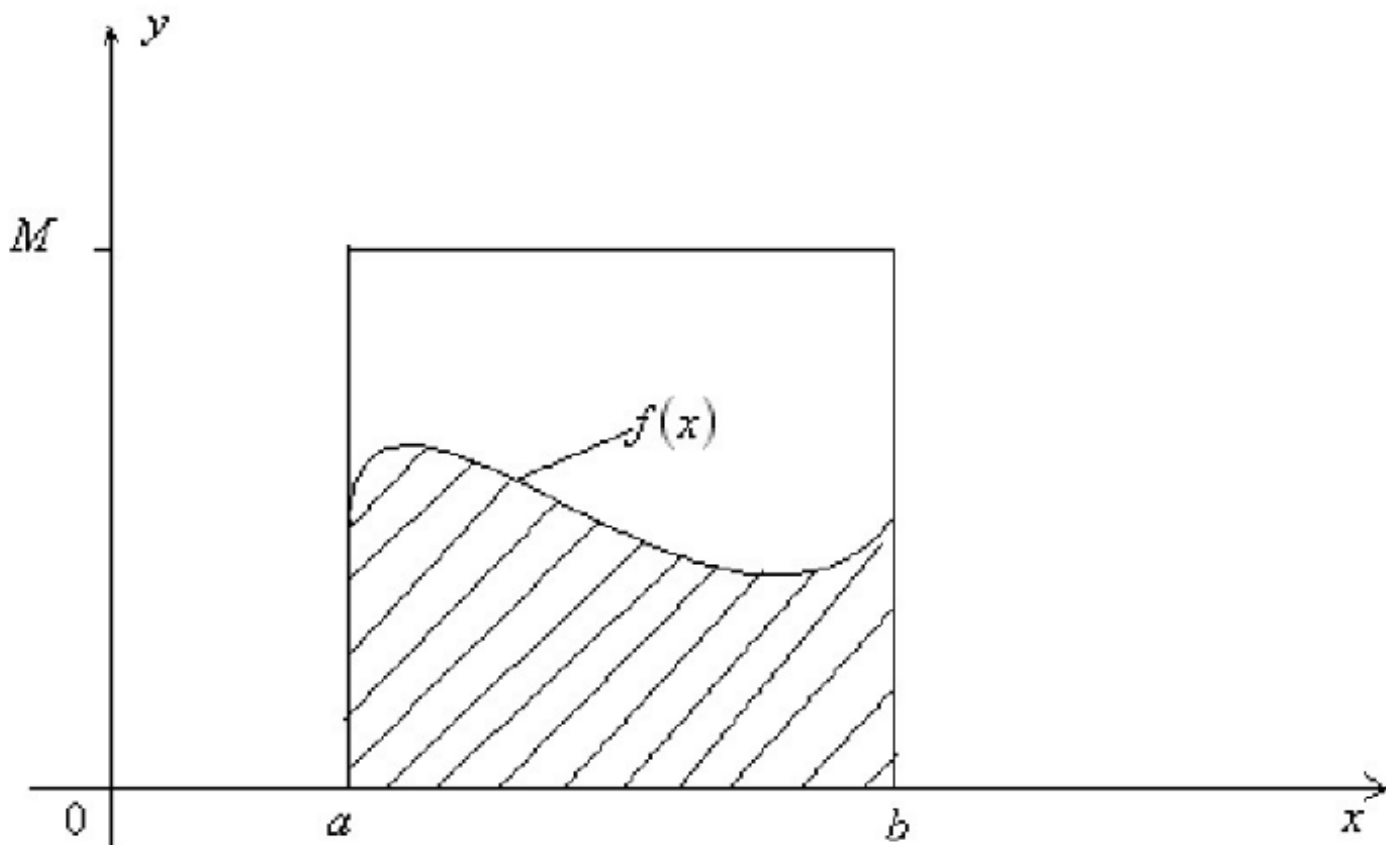
密度泛函
理论



量子蒙特
卡罗方法

蒙特卡罗方法（Monte Carlo Method, MC）基本思想

假设点 \mathbf{R} 的表达式为： $\mathbf{R} = (r_1, r_2, \dots, r_N)$



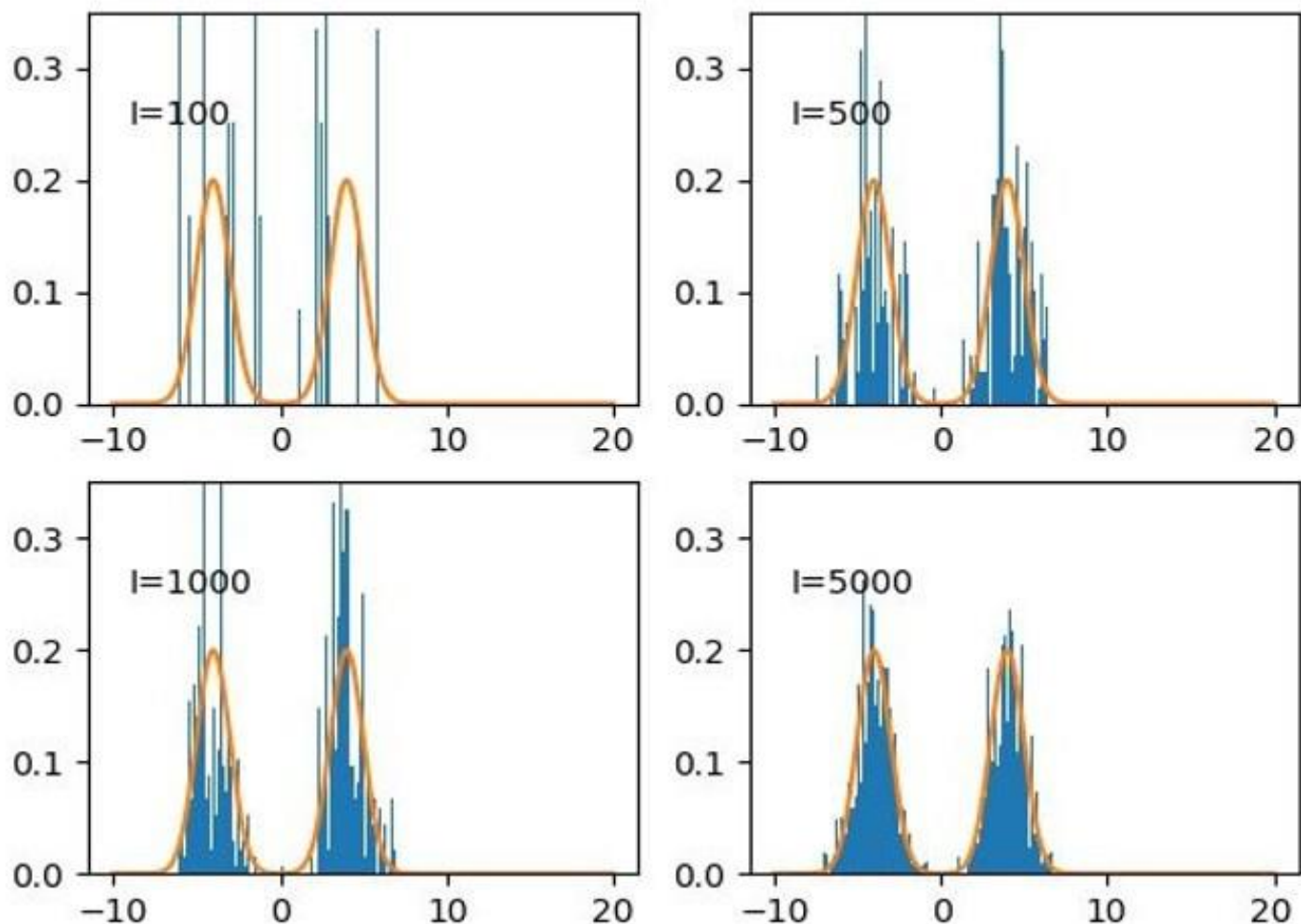
$$I = \int_a^b f(\mathbf{R}) p(\mathbf{R}) d\mathbf{R}$$

Monte Carlo Method

$$I = \frac{1}{M} \sum_{m=1}^M f(\mathbf{R}_m)$$

Metropolis算法

Metropolis_Hastings for MCMC(Normal)



1. 给定随机起始态 \mathbf{R} ;
2. 选择一个新态 \mathbf{R}' , 根据规则确定是否步行至新态;

$$A(\mathbf{R}' \leftarrow \mathbf{R}) = \text{Min} \left(1, \frac{T(\mathbf{R} \leftarrow \mathbf{R}') \mathcal{P}(\mathbf{R}')}{T(\mathbf{R}' \leftarrow \mathbf{R}) \mathcal{P}(\mathbf{R})} \right)$$

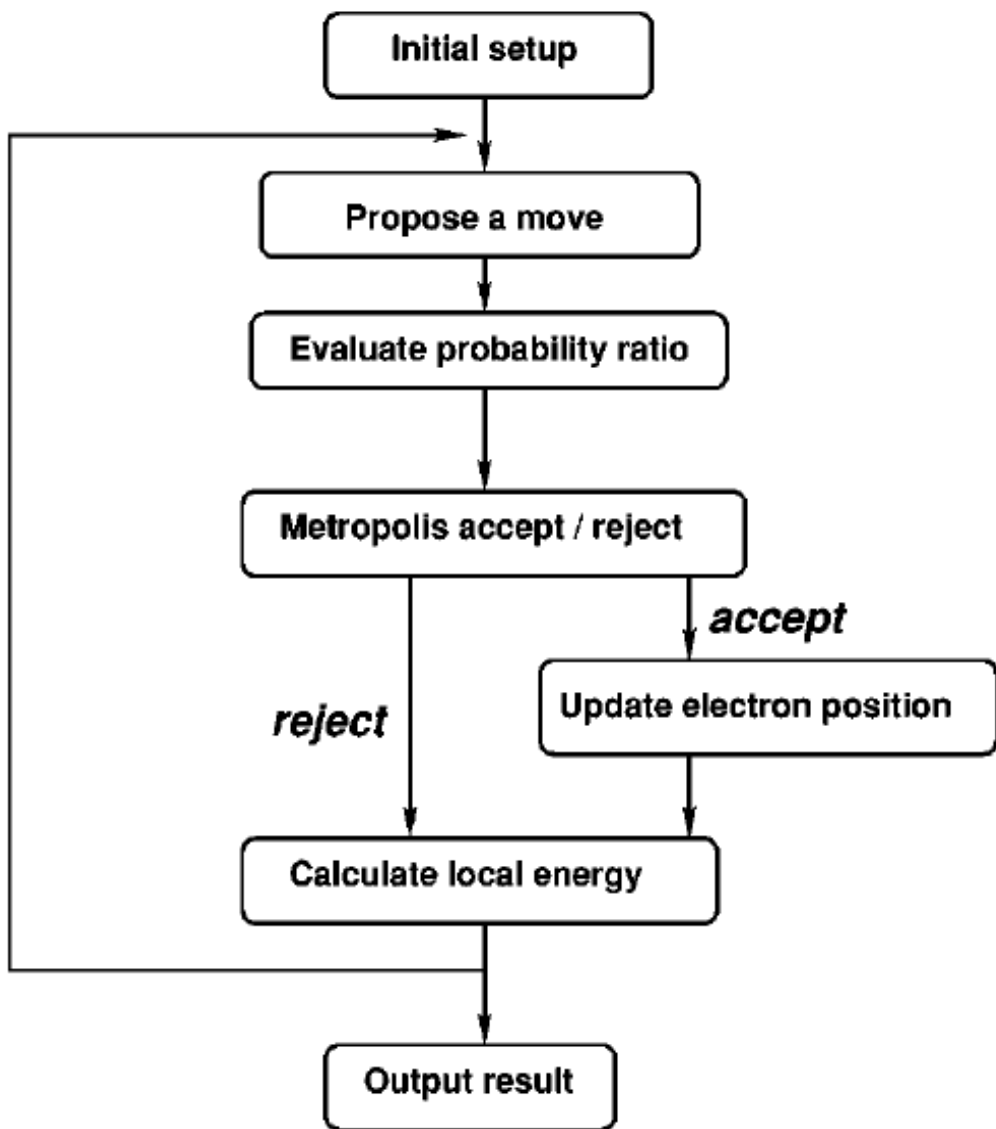
3. 重复上述步骤直至结果收敛。

$$d\mathbf{R}' A(\mathbf{R}' \leftarrow \mathbf{R}) T(\mathbf{R}' \leftarrow \mathbf{R}) \times n(\mathbf{R}) d\mathbf{R}$$

$$\frac{n(\mathbf{R})}{n(\mathbf{R}')} = \frac{A(\mathbf{R} \leftarrow \mathbf{R}') T(\mathbf{R} \leftarrow \mathbf{R}')}{A(\mathbf{R}' \leftarrow \mathbf{R}) T(\mathbf{R}' \leftarrow \mathbf{R})}$$

$$\frac{n(\mathbf{R})}{n(\mathbf{R}')} = \frac{\mathcal{P}(\mathbf{R})}{\mathcal{P}(\mathbf{R}')}$$

变分蒙特卡罗方法 (Variational MC, VMC)



变分基本公式:
$$E_V = \frac{\int \Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}} \geq E_0$$

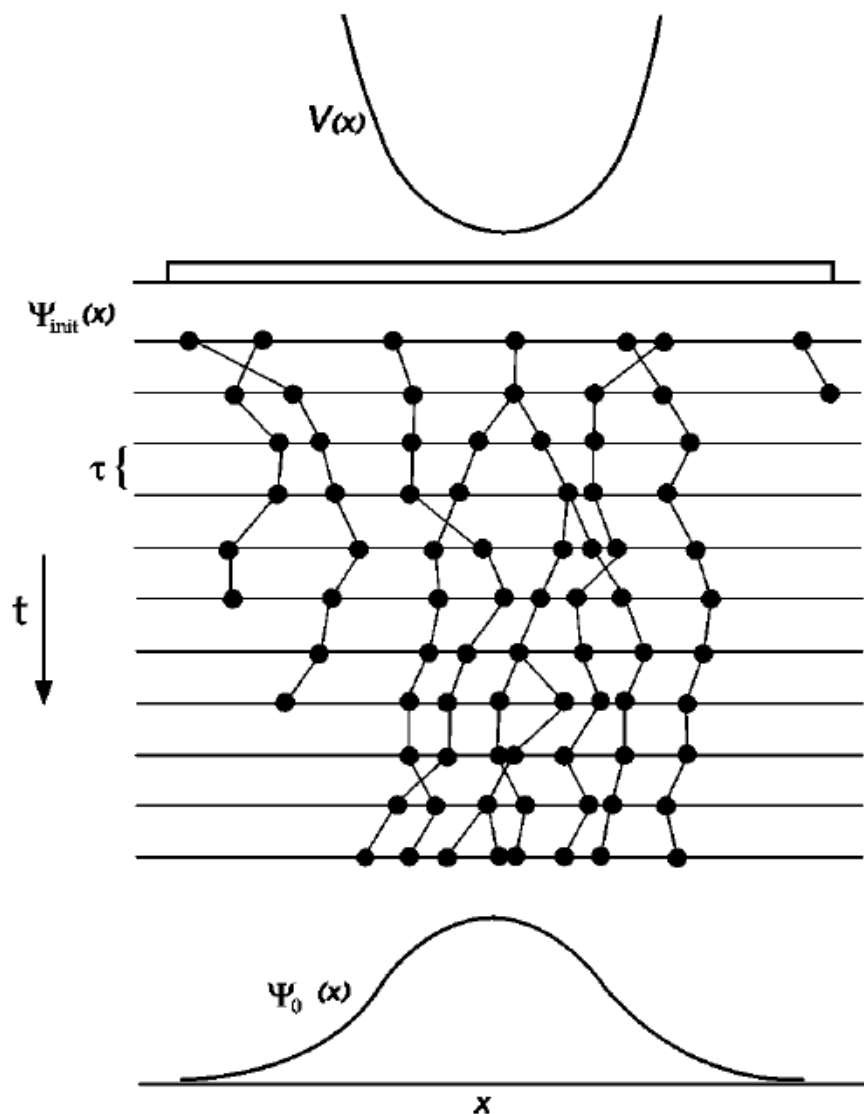
取局域能量 $E_L(\mathbf{R}) = \Psi_T(\mathbf{R})^{-1} \hat{H} \Psi_T(\mathbf{R})$

$$\mathcal{P}(\mathbf{R}) = |\Psi_T(\mathbf{R})|^2 / \int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}$$

$$E_V = \frac{\int |\Psi_T(\mathbf{R})|^2 [\Psi_T(\mathbf{R})^{-1} \hat{H} \Psi_T(\mathbf{R})] d\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}}$$

$$E_V \approx \frac{1}{M} \sum_{m=1}^M E_L(\mathbf{R}_m)$$

投影蒙特卡罗方法 (Projector MC, PMC)



哈密顿量投影: $\exp(-\tau\hat{H}) = \sum_i |\Psi_i\rangle \exp(-\tau E_i) \langle \Psi_i|$

扩散方程 (diffusion equation) :

$$\partial_t \Phi(\mathbf{R}, t) = \frac{1}{2} \sum_{i=1}^N \nabla_i^2 \Phi(\mathbf{R}, t)$$

格林函数解: $G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{-3N/2} \exp\left[\frac{-|\mathbf{R} - \mathbf{R}'|^2}{2\tau}\right]$

概率: $P = \exp[-\tau(V(\mathbf{R}) + V(\mathbf{R}') - 2E_T)/2]$

图示为出生-死亡法, 即增加粒子“出生”或“死亡”的判断。

$$M_{\text{new}} = \text{INT}(P + \eta)$$

$$E_T \leftarrow E_T - C_E \ln(M_{\text{act}}/M_{\text{ave}})$$

全组态量子蒙特卡罗方法 (Full Configuration Interaction Quantum Monte Carlo, FCIQMC)

全组态构造: $|D_i\rangle \equiv |D_{n_1, n_2, \dots, n_N}\rangle = a_{n_1}^\dagger a_{n_2}^\dagger \dots a_{n_N}^\dagger | \rangle = \frac{1}{\sqrt{N!}} |\phi_{n_1} \phi_{n_2} \dots \phi_{n_N}|, \quad n_1 < n_2 < \dots < n_N$

构造矩阵元 K_{ij} : $K_{ij} \equiv \langle D_i | K | D_j \rangle = \langle D_i | H | D_j \rangle - E_{\text{HF}} \delta_{ij}$

可给出波函数表达: $\Psi(\tau) = \sum_i C_i(\tau) |D_i\rangle$, 其中 C_i 解为 $\sum_j K_{ij} C_j = S C_i$.

模拟步骤:

生成步(spawning)



死亡/克隆步(diagonal death/cloning)



湮灭步(annihilation)

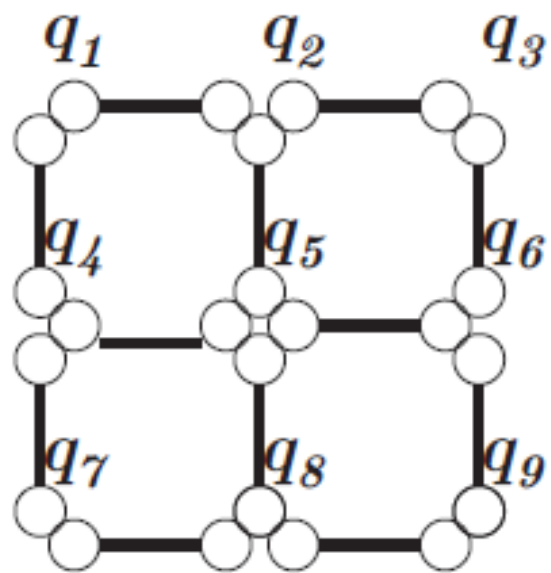
$$p_s(\mathbf{j}|\mathbf{i}_\alpha) = \frac{\delta\tau |K_{\mathbf{i}_\alpha \mathbf{j}}|}{p_{\text{gen}}(\mathbf{j}|\mathbf{i}_\alpha)}$$


$$p_d(\mathbf{i}_\alpha) = \delta\tau (K_{\mathbf{i}_\alpha \mathbf{i}_\alpha} - S)$$

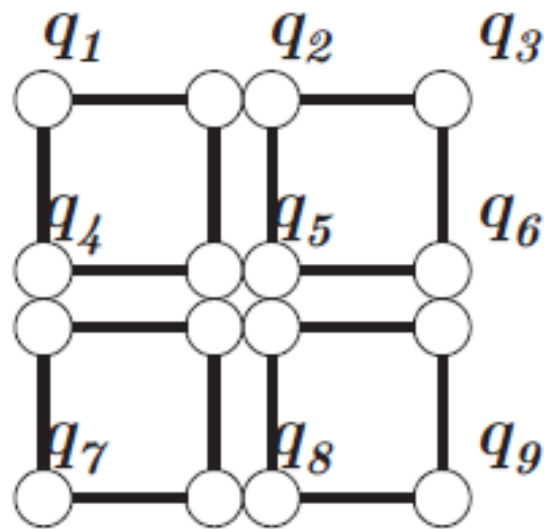
$$\mathcal{O}[N_s \ln(N_s N_w)]$$

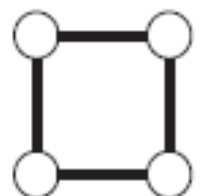
关联子直积态方法（correlator product state, CPS）

波函数表达: $|\Psi\rangle = \sum_{\{q\}} \Psi^{q_1 q_2 \dots q_L} |q_1 q_2 \dots q_L\rangle$ 关联子表示 \longrightarrow $|\Psi\rangle = \sum_{\{q\}} \prod_{\langle ij \rangle} C^{q_i q_j} |q_1, \dots, q_L\rangle$



 $= C^{q_i q_j}$



 $= C^{q_i q_j q_k q_l}$

固定体系总粒子数

$|\Psi\rangle = \sum_{\{q\}} \sum_{\langle ij \rangle} C^{q_i q_j} \hat{P}_N |q_1, \dots, q_L\rangle$

$\langle q_i q_j | \hat{C}^{ij} | q'_i q'_j \rangle = \delta_{q_i q'_i} \delta_{q_j q'_j} C^{q_i q_j}$

$|\Psi\rangle = \prod_{i>j} \hat{C}^{ij} |\Phi\rangle$

二、非线性波函数的量子蒙特卡罗方法原理

投影蒙特卡罗模拟: $|\Psi_0\rangle = \lim_{k \rightarrow \infty} [1 - \tau(\hat{H} - \hat{I}E_0)]^k |\psi^{(0)}\rangle$

Ritz变分泛函: $\langle \Psi | \hat{H} | \Psi \rangle / \langle \Psi | \Psi \rangle$

这两种方法本质上大同小异，我们使用正定的拉格朗日函数将两者结合起来：

$$\mathcal{L}[\Psi(Z_\sigma)] = \langle \Psi | \hat{H} | \Psi \rangle - E_0(\langle \Psi | \hat{I} | \Psi \rangle - A)$$

可以证明，这里所求出的泛函和Ritz泛函给出的最小值一致。

随机梯度下降 (stochastic gradient descent, SGD)

考虑一个对所有变参量都适用的梯度下降方案：

$$Z_{\sigma}^{(k+1)} = Z_{\sigma}^{(k)} - \tau_k \frac{\partial \mathcal{L}[\Psi^{(k)}]}{\partial Z_{\sigma}}$$

将该方案投影到全希尔伯特 (Hilbert) 空间，有：

$$Z_{\sigma}^{(k+1)} = Z_{\sigma}^{(k)} - \tau_k \sum_{\mathbf{n}\mathbf{m}} \left\langle \frac{\partial \Psi^{(k)}}{\partial Z_{\sigma}} \middle| \mathbf{m} \right\rangle \times (H_{\mathbf{mn}} - E^{(k)} \delta_{\mathbf{mn}}) \langle \mathbf{n} | \Psi^{(k)} \rangle$$

如果这里的波函数为线性无关构型的展开，便和投影量子蒙特卡罗模拟的基本公式一致。

随机梯度下降 (stochastic gradient descent, SGD)

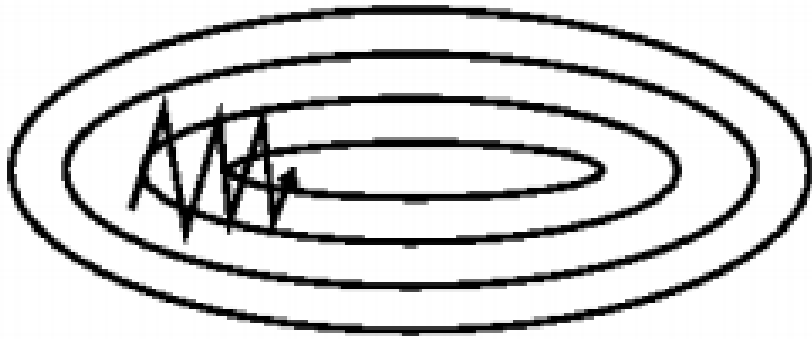
优势:

1. 不依赖于Hilbert空间尺寸，适用于大规模体系计算；
2. 避免了切空间中的矩阵构造；
3. Krylov子空间技术使得不利的条件可投影至可控空间；
4. 性质由密度泛函推导而来，而非局部能量 (e.g. VMC)。

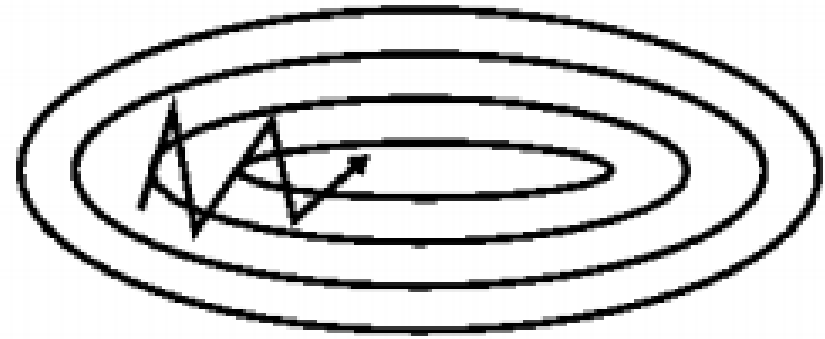
核心问题：收敛过慢，计算资源需求量大。

收敛速度： $\mathcal{O}[(1/k) + (\sigma/\sqrt{k})]$

Accumulated Scheme: original SGD and Momentum



(a) SGD without momentum



(b) SGD with momentum

Accumulated Scheme: Nesterov's Accumulated Approach

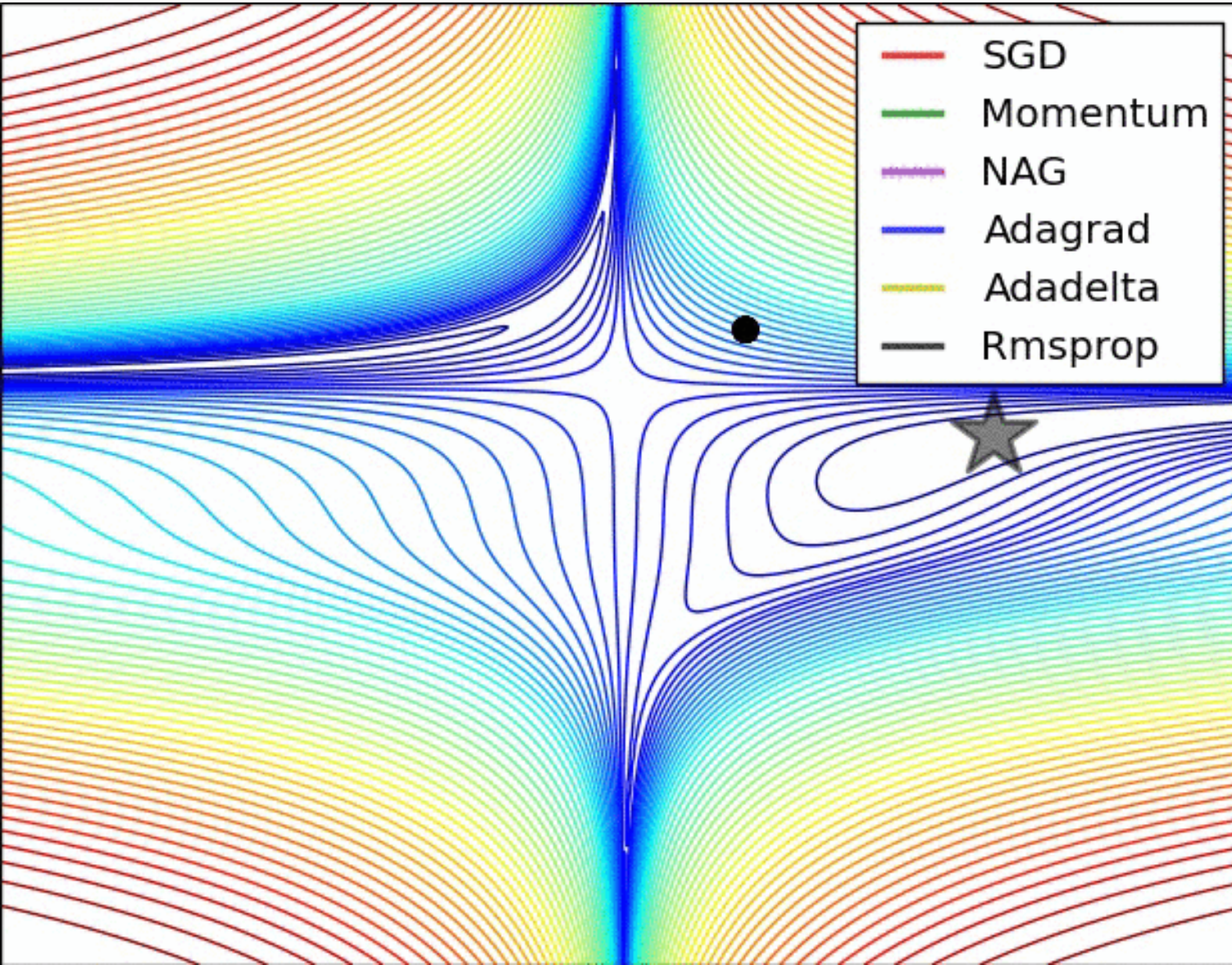
定义序列 $\left\{ \begin{array}{l} \lambda_0 = 0 \\ \lambda_k = \frac{1}{2} + \frac{1}{2} \sqrt{1 + 4\lambda_{k-1}^2} \\ \gamma_k = (1 - \lambda_k) / \lambda_{k+1} \end{array} \right.$

设置起始点: $Z_\sigma^{(1)} = Y_\sigma^{(1)}$

迭代方程 $\left\{ \begin{array}{l} Y_\sigma^{(k+1)} = Z_\sigma^{(k)} - \tau_k \frac{\partial \mathcal{L}[\Psi^{(k)}]}{\partial Z_\sigma} \\ Z_\sigma^{(k+1)} = (1 - \gamma_k) Y_\sigma^{(k+1)} + \gamma_k Y_\sigma^{(k)} \end{array} \right.$

稳定性修正: $\gamma_k \rightarrow \gamma_k e^{-(1/d)(k-1)}$

Accumulated Scheme: RMSprop

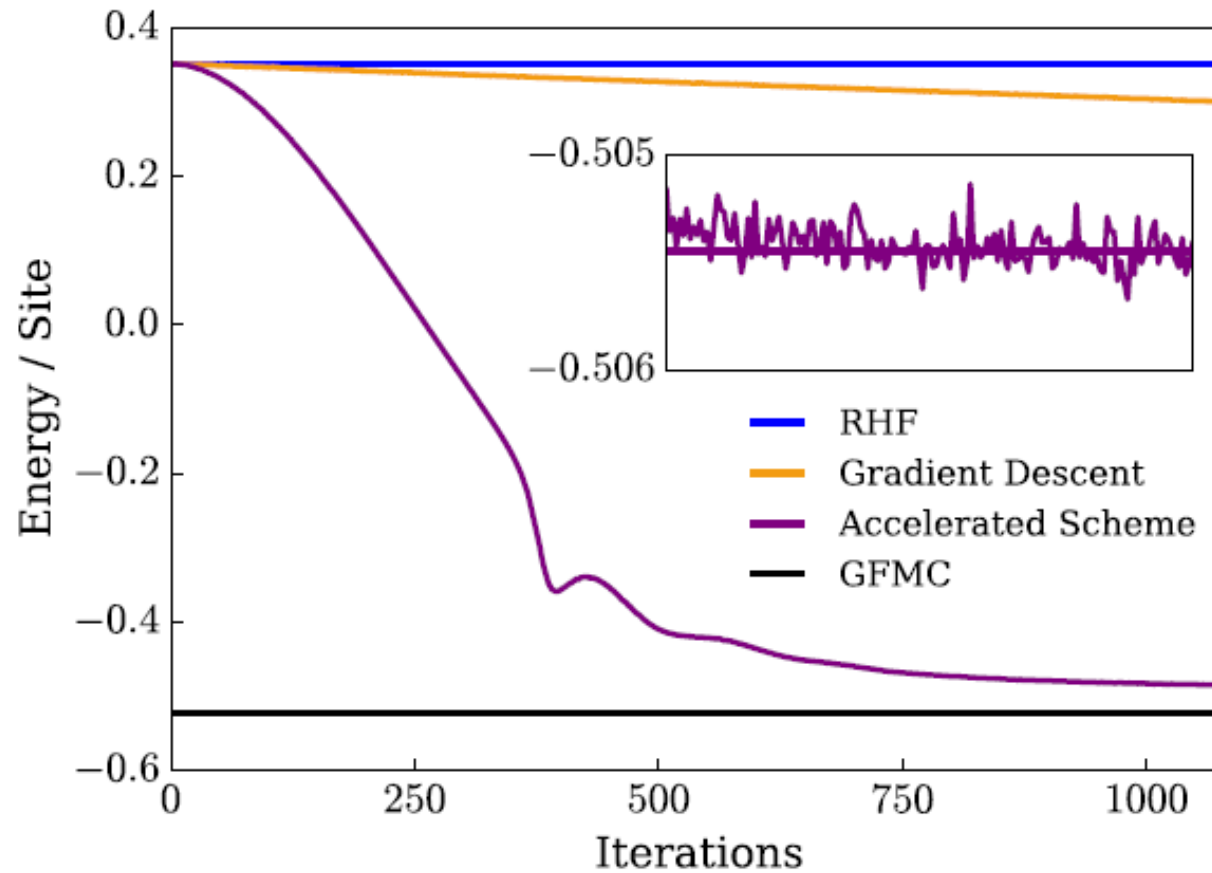


$$\tau_{Z_\sigma}^{(k)} = \eta (\text{RMS}[g_{Z_\sigma}]^{(k)})^{-1}$$

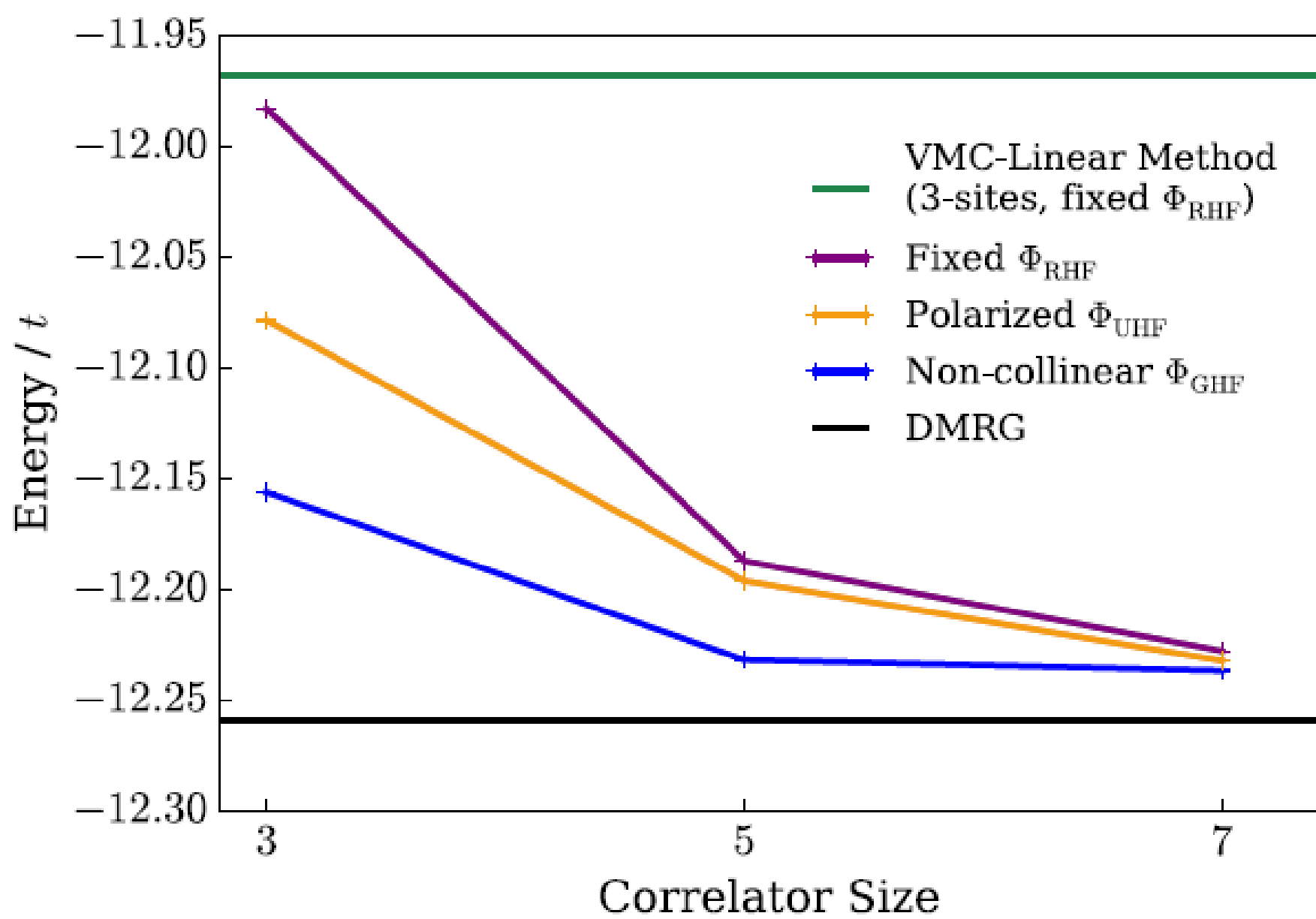
$$\text{RMS}[g_{Z_\sigma}]^{(k)} = \sqrt{E[g_{Z_\sigma}^2] + \epsilon}$$

$$E[g_{Z_\sigma}^2]^{(k)} = \rho E[g_{Z_\sigma}^2]^{(k-1)} + (1 - \rho) g_{Z_\sigma}^2$$

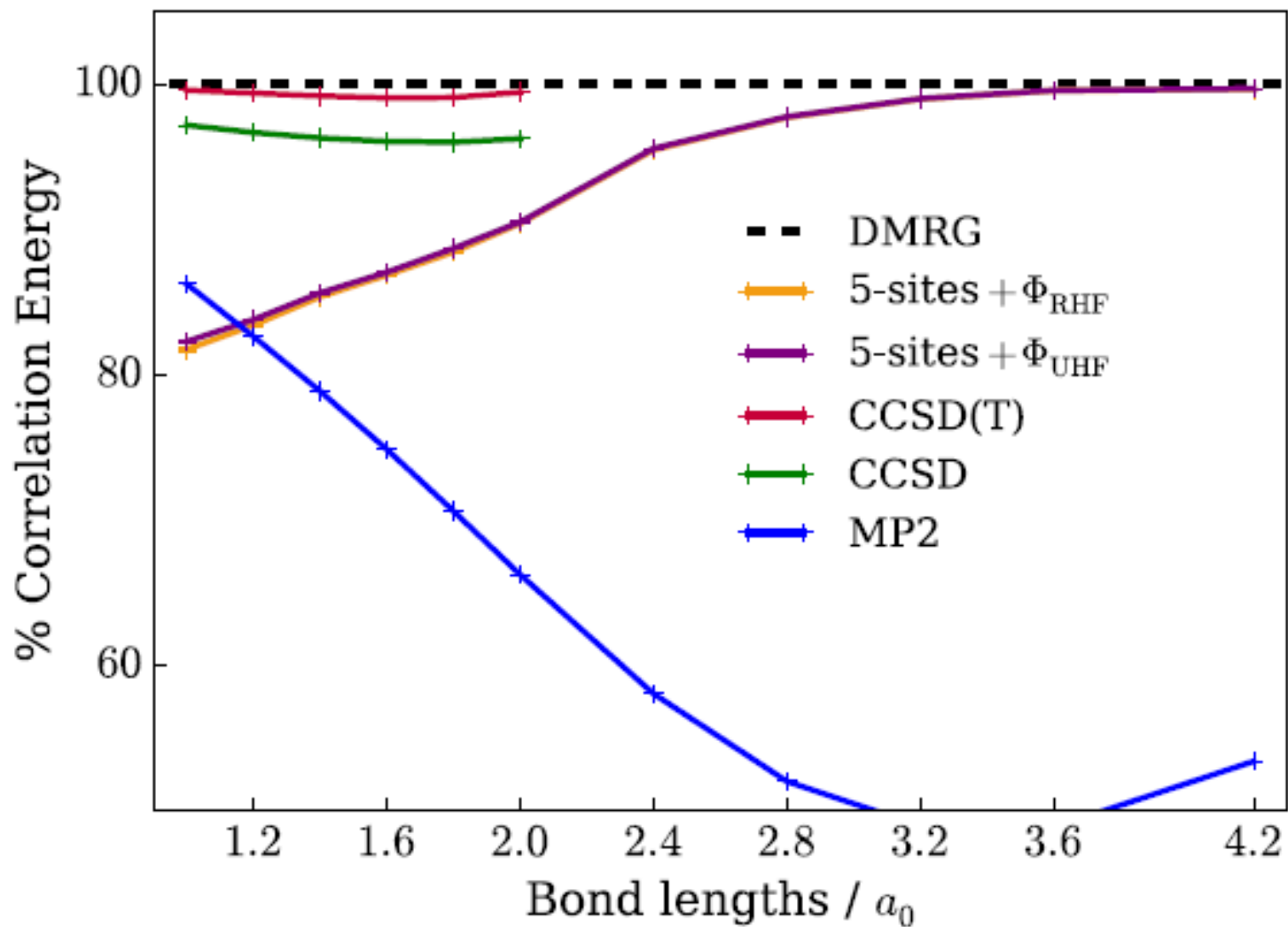
三、结果讨论与对比



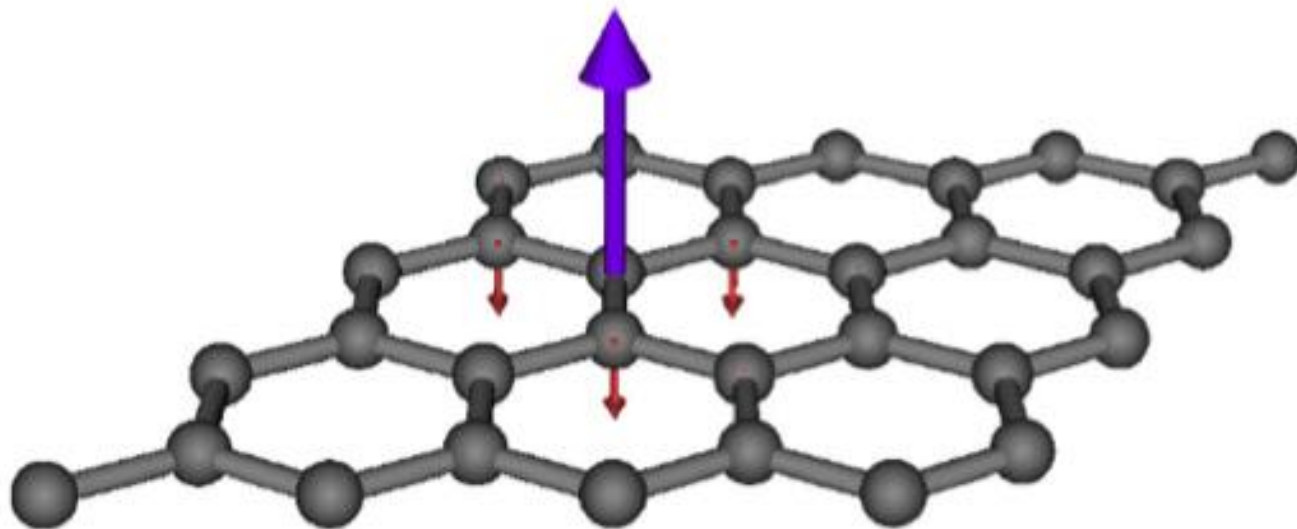
实验1：有 $O[10^5]$ 个参数的SGD方法的CPS收敛性和在 $U = 8t$ 时98个格点(倾斜)的二维Hubbard模型的RMSProp加速方案收敛性对比。



实验2：1 x 22哈伯德模型的一系列CPS波函数的能量收敛。



实验3：在STO-6G的基础上，由CPS波函数捕获的包含50个氢原子的线性链对称离解的DMRG相关能的百分比，包括数值精确的DMRG，以及MP2、CCSD和CCSD(T)的强相关量子化学方法。图中所有键长的总能量与DMRG的最大偏差为每个原子1.1千卡/摩尔。



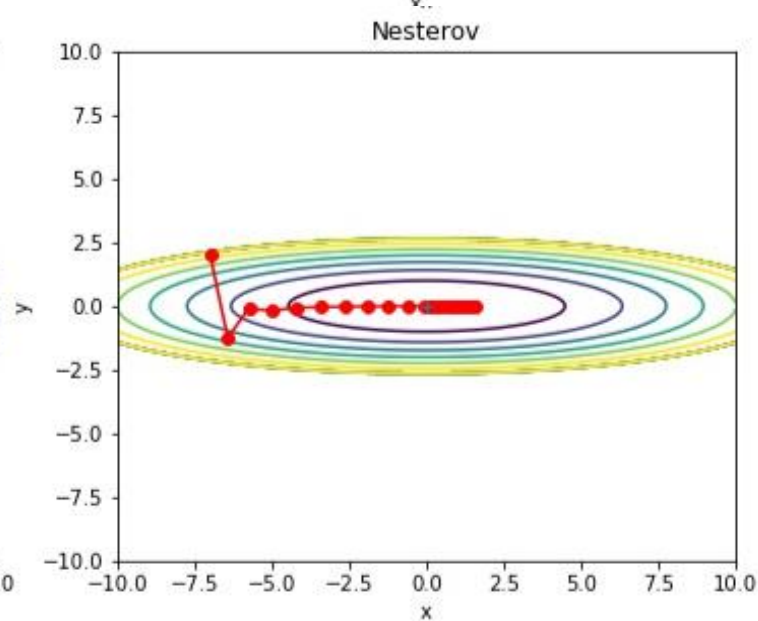
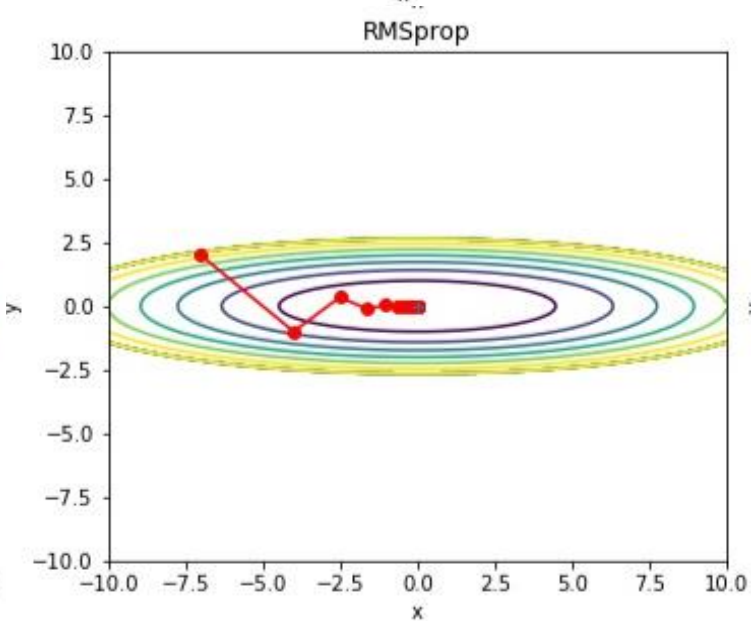
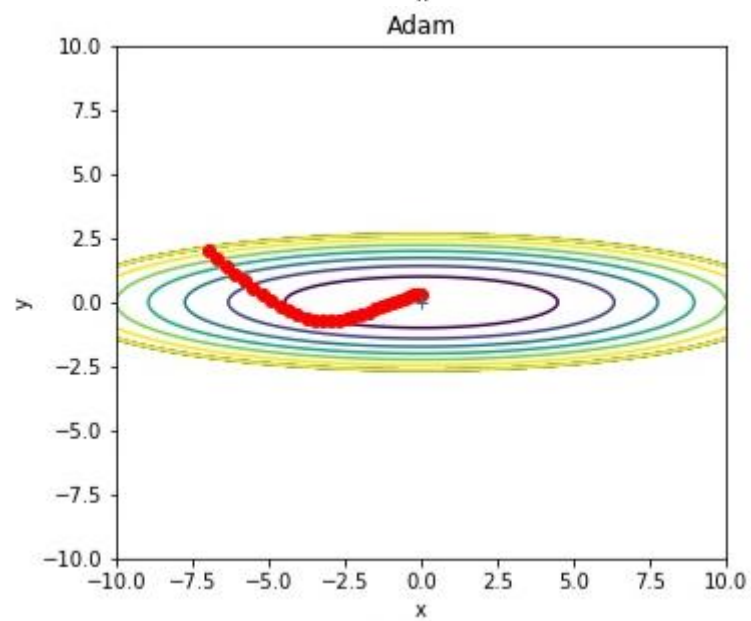
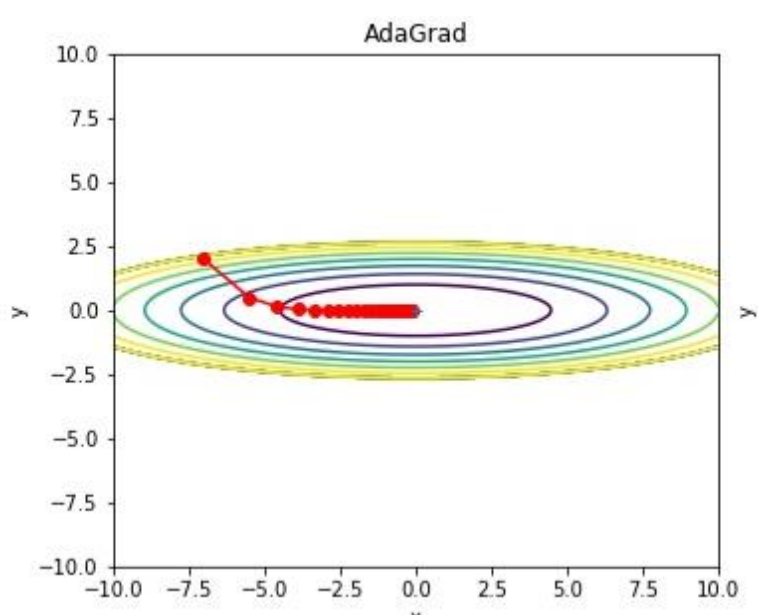
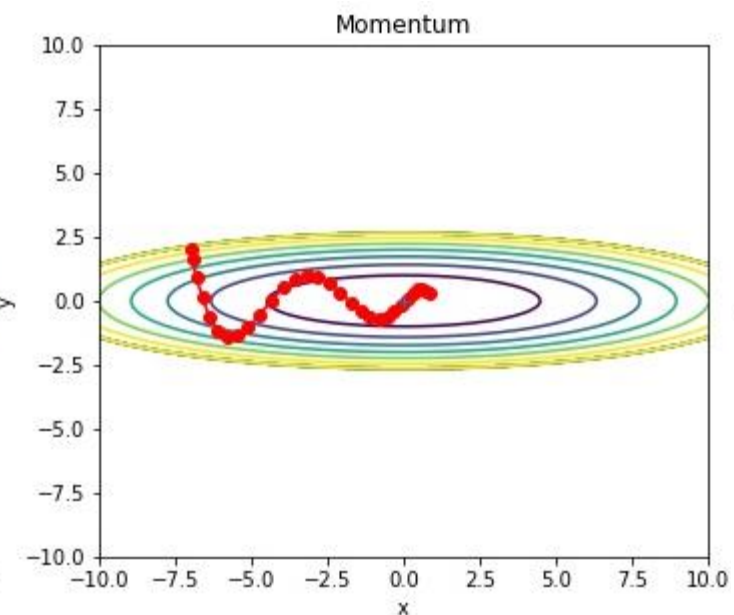
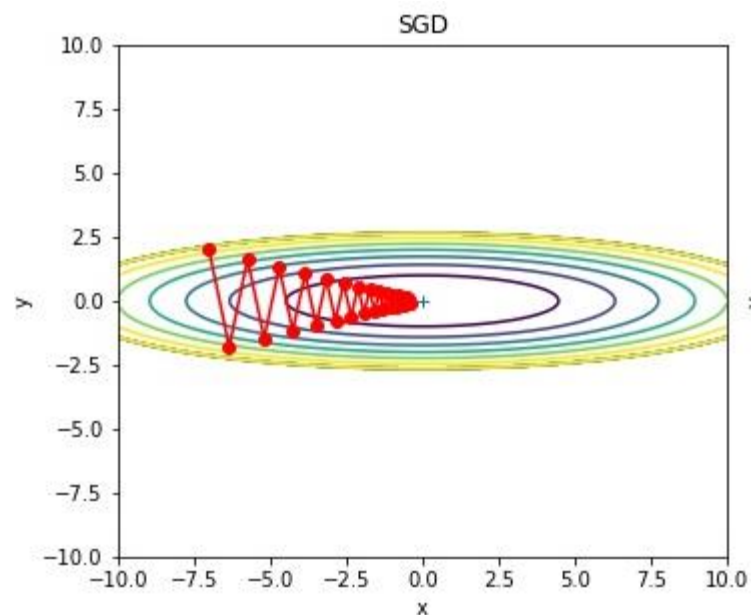
实验4：考虑一个 4×4 k-point采样的无限周期石墨烯薄片，从双 ζ 周期高斯基出发，以每个碳原子为中心选择一个定域且平动不变的 $2p_z$ 轨道，给出一个含67584个参数的波函数——相当于一个完整的活动空间的32个轨道的量子化学计算。

四、总结与思考

1. 综合了PMC (FCIQMC)和VMC两种方法的优点，使用正定的拉格朗日泛函对能量进行变分处理；
2. 对传统的SGD算法进行优化，保证算法稳定性的同时加快收敛速度；
3. 在格点系统和从头算体系中应用系列算法并进行对比，表明新算法在大体系中能够大幅度节省计算资源，同时保持原有的稳定性。

思考

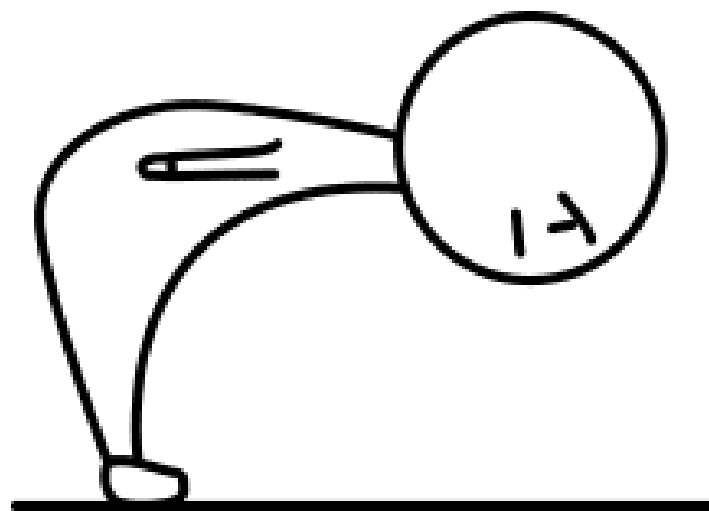
1. 在算法优化的手段中，除去RMSprop算法外，还有收敛速度更快，计算资源使用更少的Adam算法。能否将Adam算法应用于蒙特卡罗模拟中，从而进一步提高收敛的速度与稳定性？
2. 除了对SGD学习率的更新外，增加并行算法是否可行？（与上一条类似）
3. 能否在每次迭代过程后进行“洗牌”，即全部打乱，使得在较为稀疏的数据中拥有更大的概率快速选取到最优构型？



参考文献

1. Schwarz, L. R., Alavi, A., & Booth, G. H. (2017). *Projector Quantum Monte Carlo Method for Nonlinear Wave Functions*. *Physical Review Letters*, 118(17)
2. Foulkes, W. M. C., Mitas, L., Needs, R. J., & Rajagopal, G. (2001). *Quantum Monte Carlo simulations of solids*. *Reviews of Modern Physics*, 73(1), 33–83
3. Booth, G. H., Thom, A. J. W., & Alavi, A. (2009). *Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space*. *The Journal of Chemical Physics*, 131(5), 054106

感谢观看



谢谢