Random Batch Methods (RBM) for
classical and quantum N-body interacting particle systems

Shi Jin,

Institute of Natural Sciences and School of Mathematical Sciences

Shanghai Jiao Tong University

Outline

- Random Batch method for interacting particle systems (classical N-body problems)

  with Lei Li and Jian-Guo Liu
- Molecular dynamics and quantum Monte-Carlo

  with Lei Li, Zhenli Xu, Yue Zhao and Xiantao Li
- Random Batch Method for N-body Schrodinger equation

  with Francois Golse and Thierry Paul

Classical Interacting particle systems

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = -\sum\limits_{\substack{j=1\dots N \\ j\neq i}} \nabla\varphi(x_i - x_j) \end{cases}$$

$$dX^i = -\nabla V(X^i)\,dt + \frac{1}{N-1}\sum_{j:j\neq i} K(X^i - X^j)\,dt + \sigma dB^i,\ i = 1,\dots,N.$$

- We use first-order binary interacting system as an example but the idea works for second order system (Newton's second law type)

  and more interacting particles

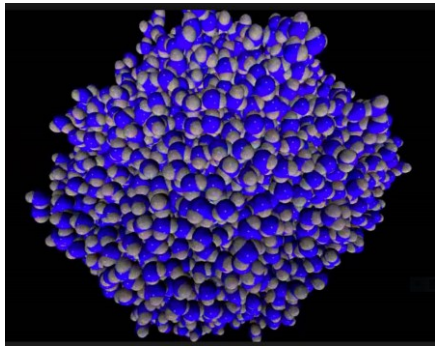
- The idea works with or without the Brownian motion term

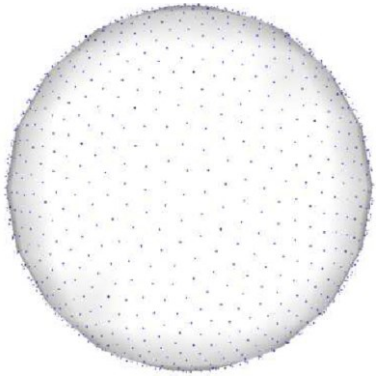
Applications

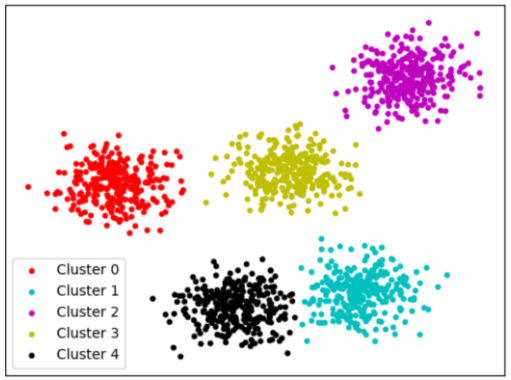
- Physics and chemistry: molecular dynamics,

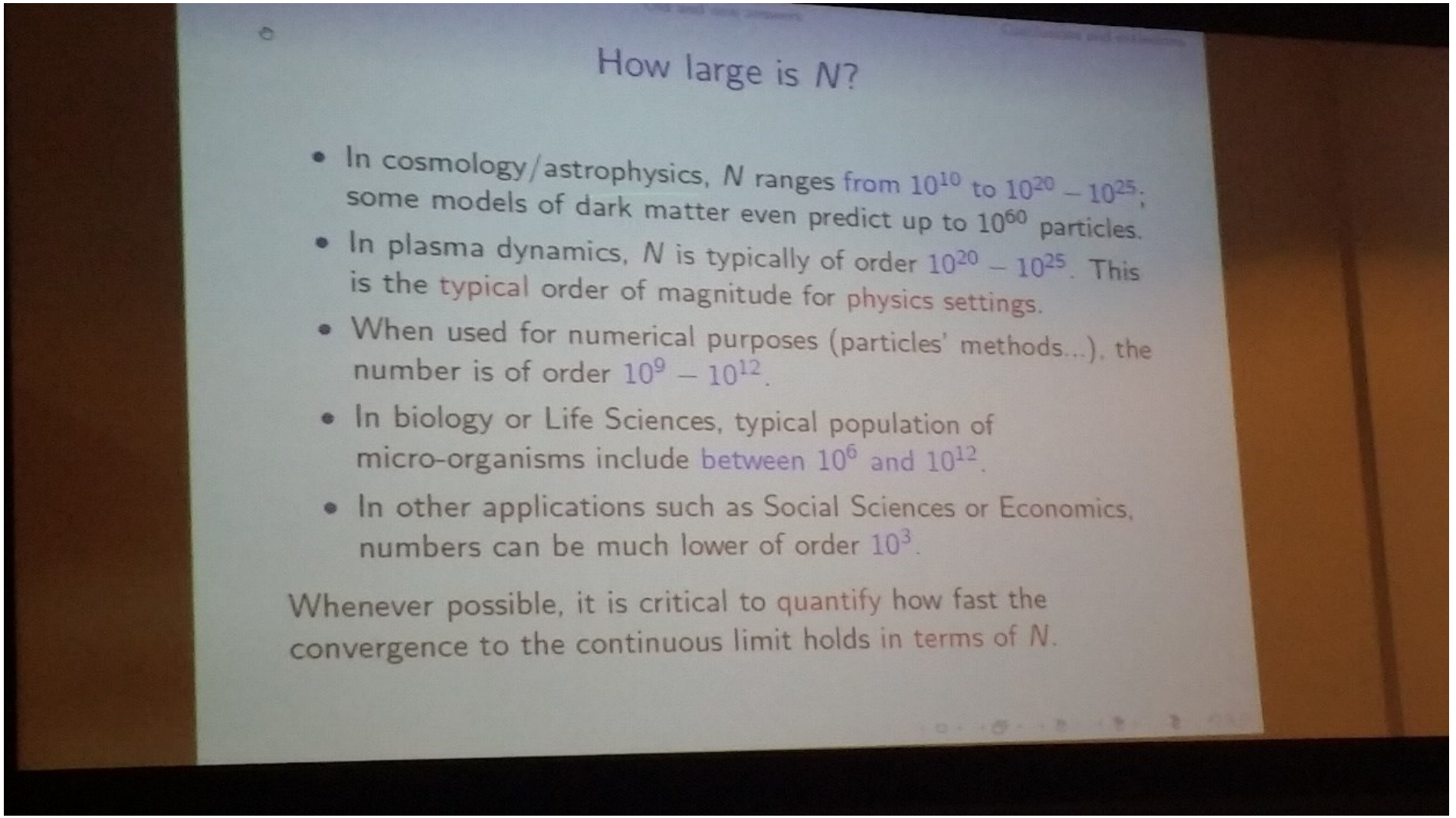
  electrostatics, astrophysics

  (stars, galaxies) …
- Biology: flocking, swarming, chemotaxis, …
- Social sciences: wealth distribution, opinion dynamics, pedestrian dynamics …
- Data sciences: clustering,…
- Numerical particle methods for kinetic/mean-field equations









• Courtesy of P.E. Jabin

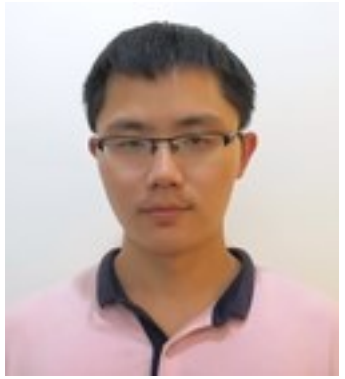
The computational cost

It is clear to be of $O(N^2)$ per time step (or $O(N^J)$ if it is J particle interaction)

Fast Multipole Methods were developed to address this issue for binary interactions

We introduce the random batch methods to reduce the computational cost to $O(N)$ per time step and it works for general interacting potentials

The random batch methods
(with Lei Li, SJTU; Jian-Guo Liu, Duke)



at each time step,

- random shuffling: group N particles randomly to n groups (batches), each batch has p $(p \ll N, \text{often } p = 2)$ particles

- particles interacting only inside their own batches

Algorithm 1 (RBM-1)

- Let time step be τ
  discrete time $t_m = m\tau$

Algorithm 1 (RBM-1)

1: **for** m in $1 : [T/\tau]$ **do**
2: Divide $\{1, 2, \ldots, pn\}$ into n batches randomly.
3: **for** each batch $\mathcal{C}_q$ **do**
4: Update X^i's $(i \in \mathcal{C}_q)$ by solving the following SDE with $t \in [t_{m-1}, t_m)$.

$$dX^i = -\nabla V(X^i)dt + \frac{1}{p-1} \sum_{j \in \mathcal{C}_q, j \neq i} K(X^i - X^j)dt + \sigma dB^i.$$

5: **end for**
6: **end for**

The computational cost is $O(N)$

- Random shuffling algorithm:

  for example the Durstemfd's modern revision of Fisher-Yates shuffle algorithm costs $O(N)$

  In MATLAB: randperm(N)

- The summation cost is $O(N)$ due to small batch size p

RBM with replacement

- At each time step, draw a batch of size p randomly, interacting within this batch, for n independent times

Algorithm 2a (RBM-r′)

1: **for** m in $1 : [T/\tau]$ **do**
2: **for** k from 1 to N/p **do**
3: Pick a set $\mathcal{C}_k$ of size p randomly with replacement.
4: Update X^i's $(i \in \mathcal{C}_k)$ by solving the following SDE for time τ.

$$\begin{cases} dY^i = -\nabla V(Y^i)dt + \dfrac{1}{p-1} \displaystyle\sum_{j \in \mathcal{C}_k, j \neq i} K(Y^i - Y^j)dt + \sigma dB^i, \\ Y^i(0) = X^i, \end{cases}$$

i.e., solve (2.7) with initial values $Y^i(0) = X^i$, and set $X^i \leftarrow Y^i(\tau)$.

5: **end for**
6: **end for**

Remarks

- For these methods to be competitive over the deterministic solvers, the time step τ needs to be independent of N (we will prove this for special V and K)

- For Coulomb interaction, one can do a time splitting

$$dX^i = \frac{1}{p-1} \sum_{j \in \mathcal{C}, j \neq i} K(X^i - X^j)\, dt,$$

$$dX^i = -\nabla V(X^i)\, dt + \sigma dB^i.$$

then for p=2, the Coulomb interaction step can be solved analytically, avoiding stiffness due to singularity

Relevant approaches in other fields

- stochastic gradient (or coordinate) descent methods in machine learning (use small and random batches to do gradient descent)

- Direct simulation Monte-Carlo methods (Birds, Bobylev, Nanbu)—based on binary collisions-- for Boltzmann equation; and its adaptation for mean-field equations of flocking models using random binary interactions (Albi, Pareschi, Carrillo)

Theoretical analysis for RBM-1 (with p=2)

Assumption 3.1. Suppose V is strongly convex on $\mathbb{R}^d$ so that $x \mapsto V(x) - \frac{r}{2}|x|^2$ is convex, and ∇V, $\nabla^2 V$ have polynomial growth (i.e. $|\nabla V(x)| + |\nabla^2 V(x)| \leq C(1 + |x|^q)$ for some $q > 0$). Assume $K(\cdot)$ is bounded, Lipschitz on $\mathbb{R}^d$ with Lipschitz constant L and has bounded second order derivatives. Moreover,

$$r > 2L. \tag{3.1}$$

This assumption guarantees that the evolution group of the original deterministic particle system is a contraction:

$$\frac{d}{dt}\sum_{i=1}^{N}|X^i - Y^i|^2 \leq -(r-2L)\sum_{i=1}^{N}|X^i - Y^i|^2$$

In this case one needs $\sigma > 0$ to have a non-trivial equilibrium (otherwise all particles go to a single point)

Notations and definitions

- Solution of the original particle system: X^i
- Solution of RBM-1 $\tilde{X}^j$
- Assume all particles are indistinguishable
- Coupling $\tilde{X}^i(0) = X^i(0) =: X_0^i \sim \nu,$
  $B^i(t) = \tilde{B}^i(t).$

- Error process $Z^i(t) := \tilde{X}^i(t) - X^i(t)$
- Norm of error $\|v\| = \sqrt{\mathbb{E}|v|^2}$

- Wasserstein-2 distance $W_2(\mu,\nu) = \left(\inf_{\gamma \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 d\gamma \right)^{1/2}$

Error estimates

Theorem 3.1. *Suppose Assumption 3.1 holds. With the coupling constructed above,*

$$\sup_{t\geq 0}\|Z^1(t)\| \leq C\sqrt{\frac{\tau}{p-1}+\tau^2}, \tag{3.9}$$

where C is independent of N, p and t. Consequently, let $\mu_N^{(1)}(t)$ be the first marginal for (1.2) and $\tilde{\mu}_N^{(1)}$ be the first marginal for system (2.3). Then

$$\sup_{t\geq 0} W_2(\mu_N^{(1)}(t), \tilde{\mu}_N^{(1)}(t)) \leq C\sqrt{\frac{\tau}{p-1}+\tau^2} \leq C\sqrt{\tau}. \tag{3.10}$$

- Since C is independent of N, therefore τ is independent of N

A key lemma (consistency)

$$\chi_{m,i}(X(t)) = \frac{1}{p-1} \sum_{j \in \mathcal{C}_q, j \neq i} K(X^i - X^j) - \frac{1}{N-1} \sum_{j: j \neq i} K(X^i - X^j)$$

Lemma 3.1. *Consider $p \geq 2$ and a given fixed $x = (x^1, \ldots, x^N) \in \mathbb{R}^{Nd}$. Then, for all i,*

$$\mathbb{E}\chi_{m,i}(x) = 0, \tag{3.11}$$

where the expectation is taken with respect to the random division of batches. Moreover, the variance is given by

$$\mathrm{Var}(\chi_{m,i}(x)) = \left(\frac{1}{p-1} - \frac{1}{N-1} \right) \Lambda(x), \tag{3.12}$$

where

$$\Lambda(x) := \frac{1}{N-2} \sum_{j: j \neq i} \left(K(x^i - x^j) - \frac{1}{N-1} \sum_{k: k \neq i} K(x^i - x^k) \right)^2. \tag{3.13}$$

Mean-field limit

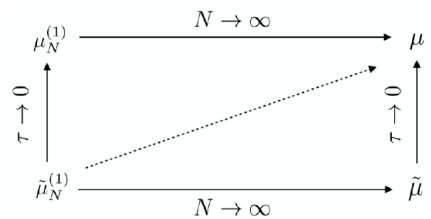
- The empirical measure

$$\mu_N(t) := \frac{1}{N}\sum_{i=1}^{N}\delta(x - X^i(t))$$

- As $N \to \infty$, under certain assumption on V and K, the marginal distribution of particles converge to the Fokker-Planck equation

$$\partial_t \mu = \nabla \cdot (\nabla V(x)\mu) - \nabla \cdot (\mu K * \mu) + \frac{1}{2}\sigma^2 \Delta \mu$$

Error to the mean-field limit



Corollary 3.1. *Suppose Assumption 3.1 holds, then*

$$W_2(\tilde{\mu}_N^{(1)}(t), \mu(t)) \leq C(\sqrt{\tau} + N^{-1/2+\varepsilon})$$

for any $\varepsilon > 0$.

(Cattiaux-Guillin-Malrieu)

- If the deterministic flow is not contracting, one may still get

$$W_2(\tilde{\mu}_N^{(1)}(t), \mu(t)) \leq C(T)(\sqrt{\tau} + N^{-1/2}), \ \forall t \in [0, T].$$

(Dobrushin, Benachour-Roynette-Talay-Vallois, Jabin-Wang)

Applications

- Dyson Brownian motion from random matrix theory

$$d\lambda_j(t) = -\beta\lambda_j(t) + \frac{1}{N}\sum_{k:k\neq j}\frac{1}{\lambda_j - \lambda_k}dt + \frac{1}{\sqrt{N}}dB_j$$

- Charged particles on sphere: Thomson's problem (Brownian motion on sphere):

  F Coulomb

$$dX^i = \mathbb{P}_S\left(\frac{1}{N-1}\sum_{j\neq i}F(X^i - X^j)\right)dt + \sqrt{2D_1}dB_S^i$$

- Stochastic dynamics of wealth (long range interaction, multiplicative noise)

$$dY^i = -\frac{\kappa}{N-1}\sum_{k:k\neq i}\partial_y\phi(Y^i - Y^k)\,dt + \sqrt{2D}Y^i dB^i \qquad \phi(y) = \frac{1}{2}y^2$$

- Opinion dynamics
- Data clustering and stochastic block model, reordering sparse matrix

The Dyson Brownian motion

- The eigenvalues of a Hermitian matrix valued Ornstein-Ulenbeck process satisfies Dyson Brownian motion:

$$d\lambda_j(t) = -\beta\lambda_j(t) + \frac{1}{N}\sum_{k:k\neq j}\frac{1}{\lambda_j - \lambda_k}dt + \frac{1}{\sqrt{N}}dB_j \qquad (1 \le j \le N)$$

  (Tao, Erdos-Yau)

- It has a mean-field limit

$$\partial_t\rho(x,t) + \partial_x(\rho(u-\beta x)) = 0,\ \ u(x,t) = \pi(H\rho)(x,t)$$

- Analytical solution (for $\beta = 1$):

$$\rho(x,t) = \frac{\sqrt{2\sigma(t) - x^2}}{\sigma(t)\pi},\ \ \sigma(t) = 1 + e^{-2t}$$

- Invariant measure (semi-circle law):

$$\rho(x) = \frac{1}{\pi}\sqrt{2 - x^2}$$

Comparison between RBM-1 and RMB-r

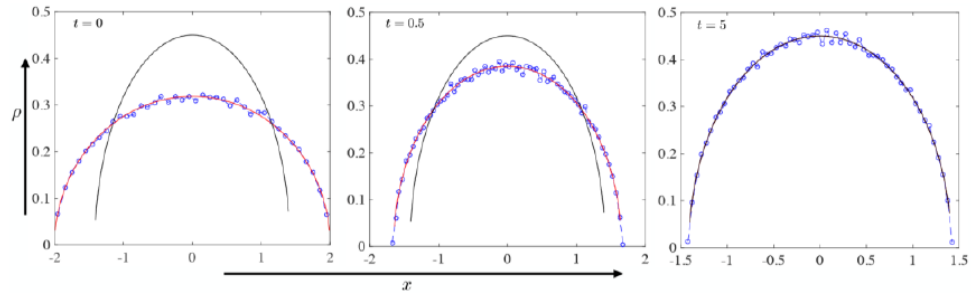


Figure 4: The RMB-1 simulation of the Dyson Brownian motion. The empirical densities at various times are plotted. The red curve is the density distribution predicted by the analytic solution (4.8). The black curve is the equilibrium semicircle law (4.7).

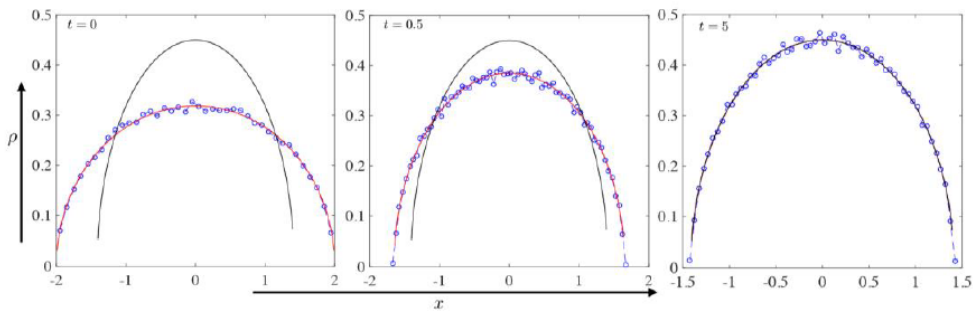


Figure 5: The RBM-r simulation of Dyson Brownian motion. The 'time' is regarded as $\tau = 10^{-3}$ for $N/2$ iterations. The red curve is the density distribution predicted by analytic solution (4.8). The black curve is the equilibrium semicircle law (4.7).

Charged particles on sphere

Thomson problem is to determine the stable configuration of N electrons on a sphere. When N becomes large, this could lead to the so-called spherical crystals. The configuration may have some meta-states (local minimizers of the energy surface). When the number of particles is large, the spherical crystals have defects due to the topology of the sphere.
In the $N \to \infty$ limit, the mean field energy

$$E(\rho) = \frac{1}{2} \iint_{S \times S} \frac{1}{|x-y|} \rho(x)\rho(y)\, dS_x dS_y$$

is minimized

In the overdamped limit and with suitable scaling, we then have interacting particle system on sphere

$$dX^i = \mathbb{P}_S \left(\frac{1}{N} \sum_{j \neq i} F(X^i - X^j) \right) dt + \sqrt{2D_1} dB_S^i$$

F: Coulomb

RBM-r (consider $D_1 = 0$)

- Randomly picking two indices. Then for $t \in [t_{m-1}, t_m)$ solve

$$dX^i = \sum_{j:I(i,j)=1} \frac{X^i - X^j}{|X^i - X^j|^3}\, dt,$$

  where $I(i,j) = 1$ means that i, j are in the same batch.
- Project the obtained points back to the sphere by dividing its magnitude.

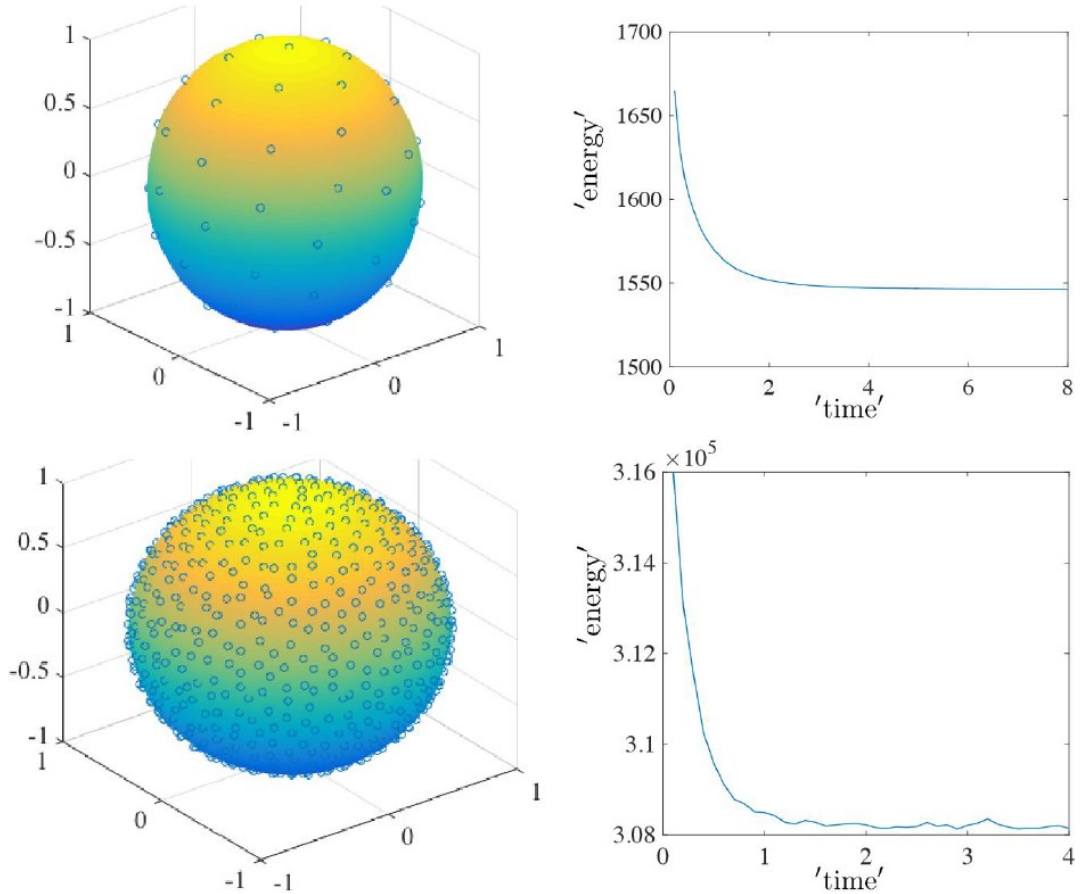


Figure 6: Charged particles on sphere. The first row is for $N = 60$ while the second row corresponds to $N = 800$. The first column shows the distributions at the end of simulation while the second column shows how the energy changes with 'time'.

Stochastic dynamics of wealth

- A mean-field game model by Degond, Liu and Ringhofer:
  Considering N market agents with two attributes:
  the economic configurations X^i and its wealth Y^j

$$\dot{X}^i = V(X^i, Y^i),$$
$$dY^i = -\frac{1}{N}\sum_{k:k\neq i} \xi_{ik}\Psi(|X^i - X^k|)\partial_y\phi(Y^i - Y^k)\,dt + \sqrt{2D}Y^i dB^i. \quad (3)$$

The second equation describes the evolution of the wealth containing two mechanisms: trading model and geometric Brownian motion.

Comparison with equilibrium distribution of the mean-field equation for $\phi(y) = \frac{1}{2}y^2$. (D=1, multiplicative noise and long-range potential)

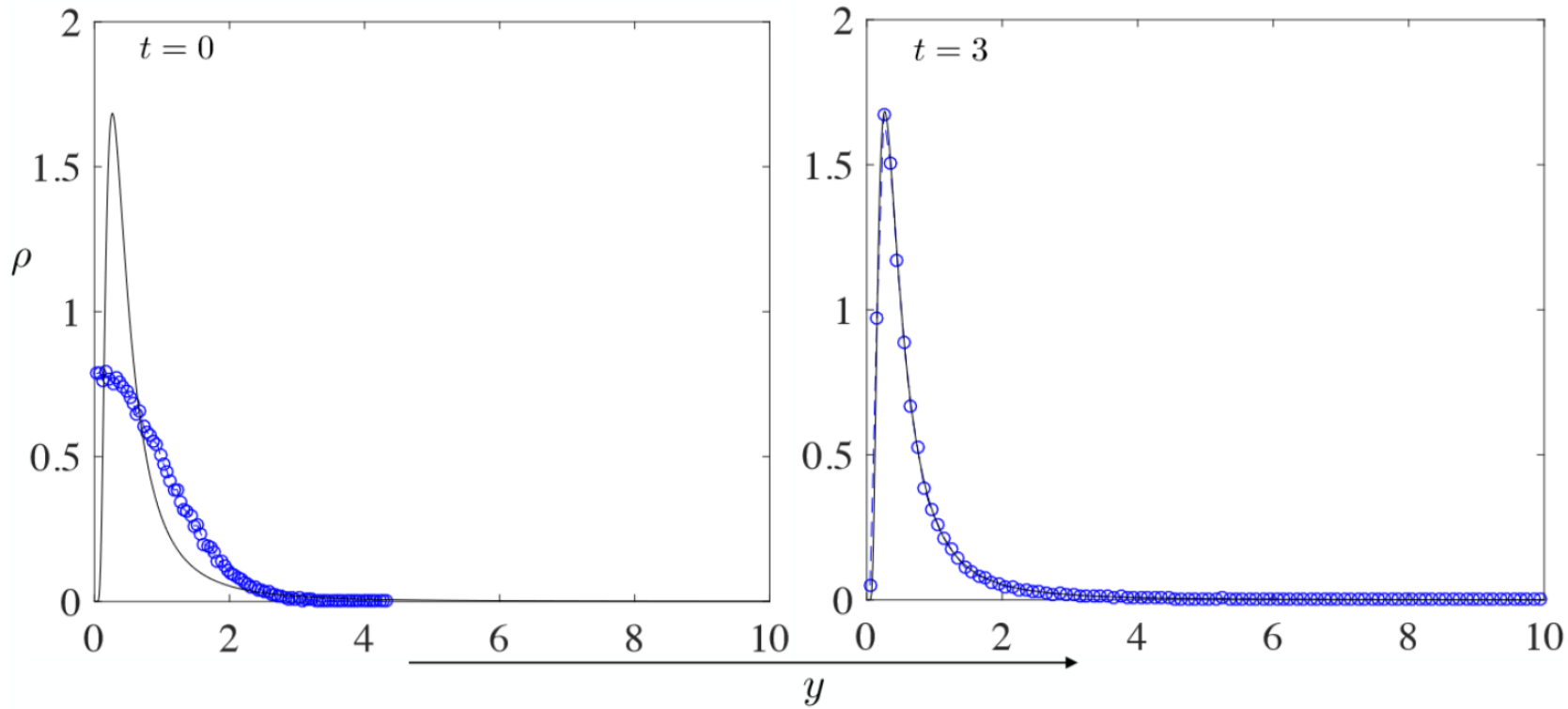


Figure 8: Wealth distribution for stochastic dynamics

Stochastic opinion dynamics (for consensus of opinions)

An opinion dynamics (Motsch-Tadmor)

$$\frac{d}{dt}X^i = \alpha \frac{1}{N} \sum_{j \neq i} \phi(|X^j - X^i|)(X^j - X^i)$$

Here, ϕ is called the influence function.

- We consider a stochastic (RBM-1) version

$$\frac{d}{dt}X^i = \alpha \phi(|X^\theta - X^i|)(X^\theta - X^i) + \epsilon_N dB^i$$

$$\phi(r) = \chi_{[0,1]}$$

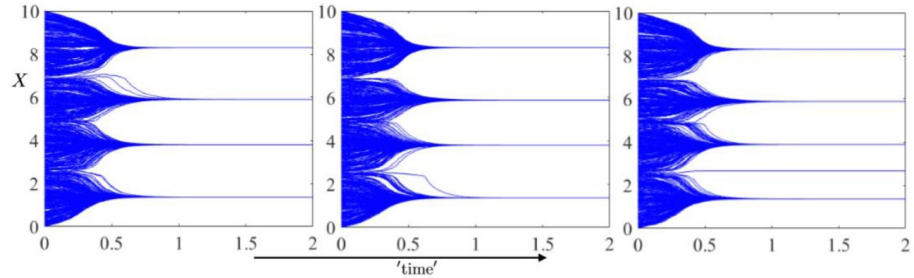


Figure: Stochastic opinion dynamics versus time. Plots of three experiments of the stochastic dynamics with the same initial data. No Brownian motion ($\epsilon_N = 0$).

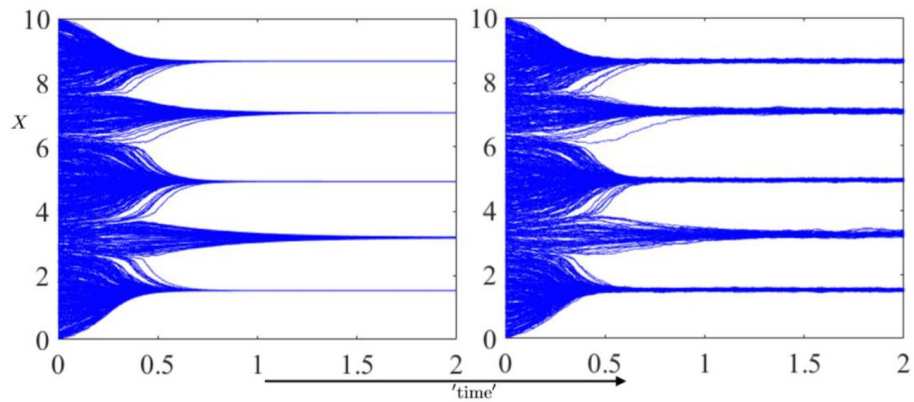


Figure: Opinion dynamics versus time. The first figure has no Brownian motion. The second is with $\epsilon_N = N^{-1/3}$. The two figures are with the same initial data.

Some other applications/extensions

- Sampling:

  Stein variational gradient descent

  L. Li, Y. Li, J.-G. Liu, J.F. Lu

  Sampling of Gibbs measure that corresponds to particle systems with Lenard-Jones potential

  L. Li, Z. Xu and X. Zhao

- Poisson-Boltzmann (particle formulation)

  L. Li, J.-G. Liu, Y. Tang

- Control of synchronalization in particle system

  E. Zuazua, etc.

- agend-based models for collective behavior (flocking, swarming, synchronization)

  S.Y. Ha, S Jin, D. Kim, D. Ko

- Molecular dynamics

  classical Coulomb: *Jin-L. Li-Z. Xu, Y. Zhao*

  quantum Monte-Carlo: *Jin-X. Li*

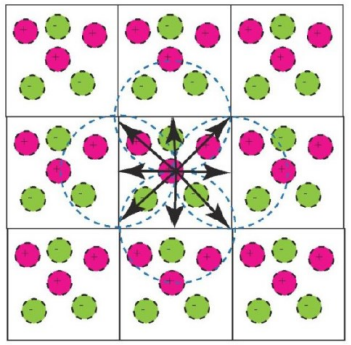
Random batch Ewald method for particle systems with Coulomb interactions (with Lei Li, Zhenli Xu and Yue Zhao)

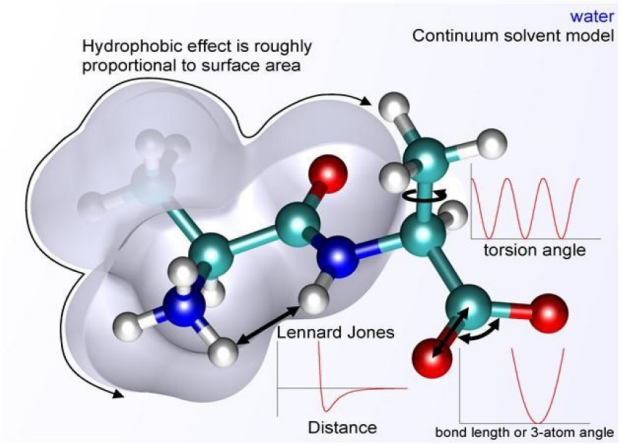
Potential energy function for particle simulations

$$E = \sum_{bonds} k_b(b-b_0)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 + \sum_{torsions} k_w[cos(nw+\gamma)+1]$$
$$+ \sum_{atom\ pairs} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + \sum_{atom\ pairs} \frac{q_i q_j}{r_{ij}}$$

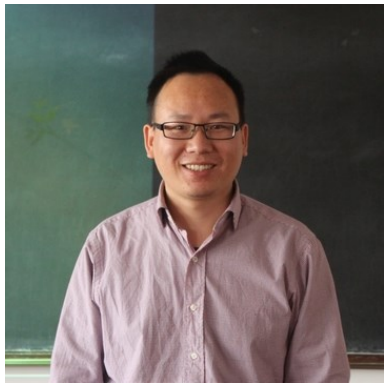
Long Range $O(N^2)$

Short range cut-off $O(N)$





Monte Carlo: Sampling from Boltzmann distribution
Molecular dynamics: Solving the Newton's equation



RBM sampling in the Fourier space

RBMD for Coulomb particles

- Ewald splitting is used:

Split potential by a short-range and a long-range parts:

$$\frac{1}{r} = \frac{\mathrm{erfc}(\alpha r)}{r} + \frac{\mathrm{erf}(\alpha r)}{r}$$

Electrostatic potential due to the charges becomes,

$$\Phi(\boldsymbol{r}) = \sum_{i=1}^{N} \sum_{\boldsymbol{n}} \frac{q_i \mathrm{erfc}(\alpha |\boldsymbol{r}_i - \boldsymbol{r} + \boldsymbol{n}L|)}{|\boldsymbol{r}_i - \boldsymbol{r} + \boldsymbol{n}L|} + \frac{4\pi}{V} \sum_{i=1}^{N} \sum_{\boldsymbol{k}\neq 0} \frac{\exp(-k^2/4\alpha^2)}{k^2} e^{j\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}_i)}$$

$$\boldsymbol{F}_i = -\nabla_{\boldsymbol{r}_i} U = -\sum_{\boldsymbol{k}\neq 0} \frac{4\pi q_i \boldsymbol{k}}{Vk^2} e^{-k^2/(4\alpha)} \mathrm{Im}(e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i} \rho(\boldsymbol{k})) - q_i \sum_{j,\boldsymbol{n}}{}' q_j G(|\boldsymbol{r}_{ij} + \boldsymbol{n}L|) \frac{\boldsymbol{r}_{ij} + \boldsymbol{n}L}{|\boldsymbol{r}_{ij} + \boldsymbol{n}L|} =: \boldsymbol{F}_{i,1} + \boldsymbol{F}_{i,2},$$

where we recall $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$, pointing towards particle j, and

$$G(r) := \frac{\mathrm{erfc}(\sqrt{\alpha} r)}{r^2} + \frac{2\sqrt{\alpha} e^{-\alpha r^2}}{\sqrt{\pi} r}.$$

$$\boldsymbol{F}_{i,1} \approx \boldsymbol{F}^*_{i,1} := -\sum_{\ell=1}^{p} \frac{S}{p} \frac{4\pi \boldsymbol{k}_\ell q_i}{V k_\ell^2} \mathrm{Im}(e^{-i\boldsymbol{k}_\ell \cdot \boldsymbol{r}_i} \rho(\boldsymbol{k}_\ell))$$

Charge inversion in a salty environment

(All-atom simulation for 17736 water molecules)

	α	r_c	n_c	Time (s)
Ewald ($c = 0$ mM)	0.0014	90.0	8.7	16698
RBE ($c = 0$ mM)	0.0072	40.0	$p = 100$	1167
Ewald ($c = 196$ mM)	0.0014	90.0	8.7	137217
RBE ($c = 196$ mM)	0.0072	40.0	$p = 100$	15258

Table 2: Computational time per 1e5 simulation steps. The RBE samples from all frequencies and it shows p values in the n_c column.

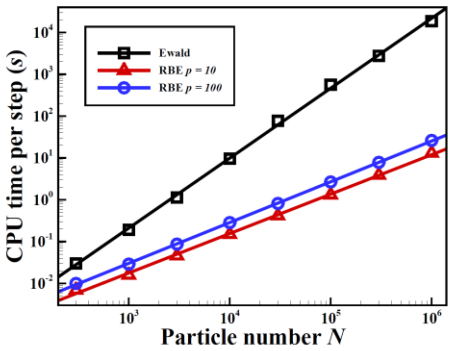
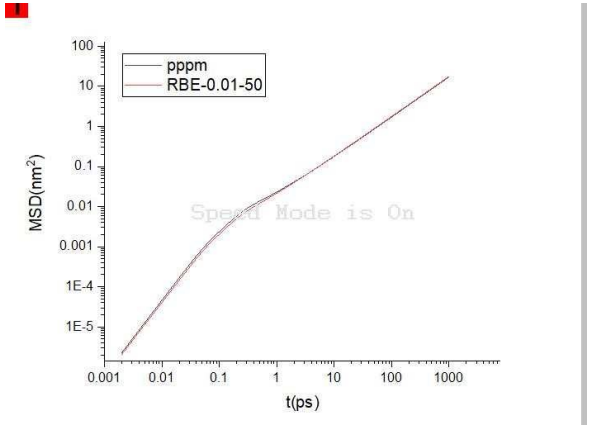
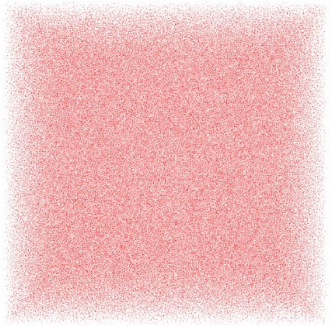


Figure 3: CPU time per step for the classical Ewald and the RBE methods with increasing N.



With L. Hong

Performance of all-atom MD simulation of water based on Random batch Ewald method

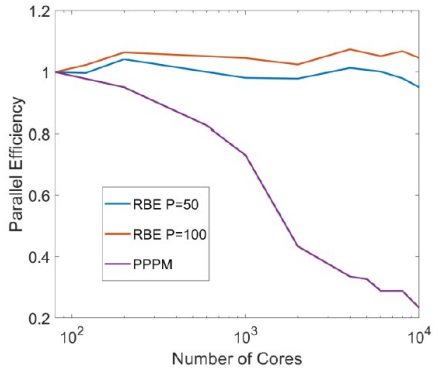


100nm

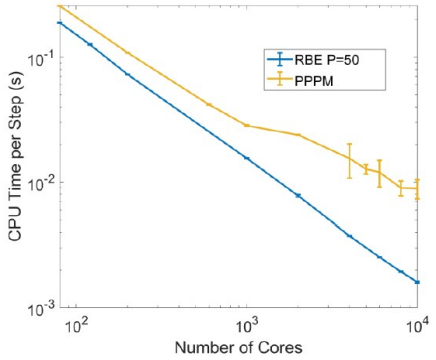
SJTU Π supercomputer

10,000,000 Water molecules

Much higher parallel efficiency



An order of magnitude faster (computational speed)



Quantum Monte-Carlo for Boson systems (with Xiantao Li)

$$\widehat{H}=\sum_{i=1}^{N}-\frac{\hbar^2}{2m}\triangle_{r_i}+\sum_{i\neq j}W(\boldsymbol{r}_i-\boldsymbol{r}_j)+\sum_{i=1}^{N}V_{ext}(\boldsymbol{r}_i).$$

$$E=\min_{\Phi}\frac{\int_{\mathbb{R}^{3N}}\Phi\widehat{H}\Phi d\boldsymbol{r}_1\cdots d\boldsymbol{r}_N}{\int_{\mathbb{R}^{3N}}|\Phi|^2 d\boldsymbol{r}_1\cdots d\boldsymbol{r}_N}$$

- Hamiltonian
- Ground state

$$E=\langle E_{\text{tot}}(\cdot)\rangle=\int p(\boldsymbol{r})E_{\text{tot}}(\boldsymbol{r})d\boldsymbol{r},\quad p(\boldsymbol{r})\propto|\Phi_0(\boldsymbol{r})|^2\quad E_{\text{tot}}(\boldsymbol{r})=\frac{\widehat{H}\Phi_0}{\Phi_0}\qquad \Phi_0=e^{-J(r)}\Pi_{i=1}^{N}\phi(\boldsymbol{r}_i),\quad u(r)=\left(\frac{a}{r}\right)^5+\frac{b^2}{r^2+c^2}.$$

$$J(\boldsymbol{r})=\sum_{i<j}u(|\boldsymbol{r}_i-\boldsymbol{r}_j|)\qquad p(\boldsymbol{r})\propto e^{-2V},\quad V=-\ln\Phi_0=-\sum_i\log\phi(\boldsymbol{r}_i)+\frac{1}{2}\sum_i\sum_{j\neq i}u(|\boldsymbol{r}_i-\boldsymbol{r}_j|)$$

$$d\boldsymbol{r}_i=\nabla\log\phi(\boldsymbol{r}_i)dt-\sum_{\substack{j=1\\j\neq i}}^{N}\nabla_{r_i}u(|\boldsymbol{r}_i-\boldsymbol{r}_j|)dt+dW_i(t),\quad 1\leq i\leq N.$$

- Variational MC: based on over-damped Langevin equation

$$\begin{cases}\boldsymbol{r}_i(t+\Delta t)=\boldsymbol{r}_i(t)+\nabla\log\phi(\boldsymbol{r}_i)\Delta t+(N-1)\nabla_{r_i}u(|\boldsymbol{r}_i-\boldsymbol{r}_j|)\Delta t+\Delta W_i,\\ \boldsymbol{r}_j(t+\Delta t)=\boldsymbol{r}_j(t)+\nabla\log\phi(\boldsymbol{r}_j)\Delta t+(N-1)\nabla_{r_j}u(|\boldsymbol{r}_i-\boldsymbol{r}_j|)\Delta t+\Delta W_j.\end{cases}$$

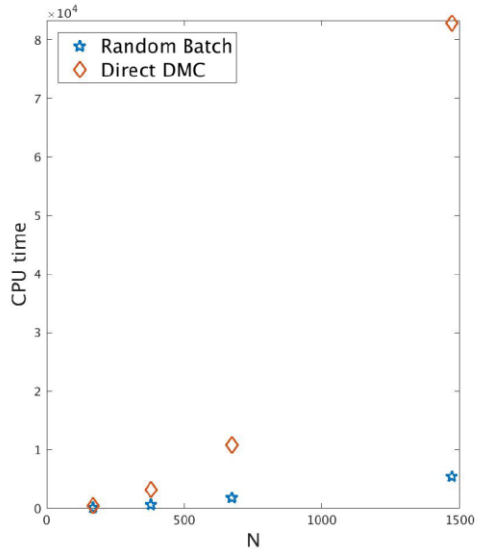
- RBM (p=2)

- Move 400 Markov Chains for 1000 steps

Table 2: Comparison of several VMC methods

	Metropolis-Hastings	Euler-Maruyama	Random Batch
CPU time for a 1000-step sampling period	1503	469	54

- Direct DMC vs RBM



Diffusion QMC

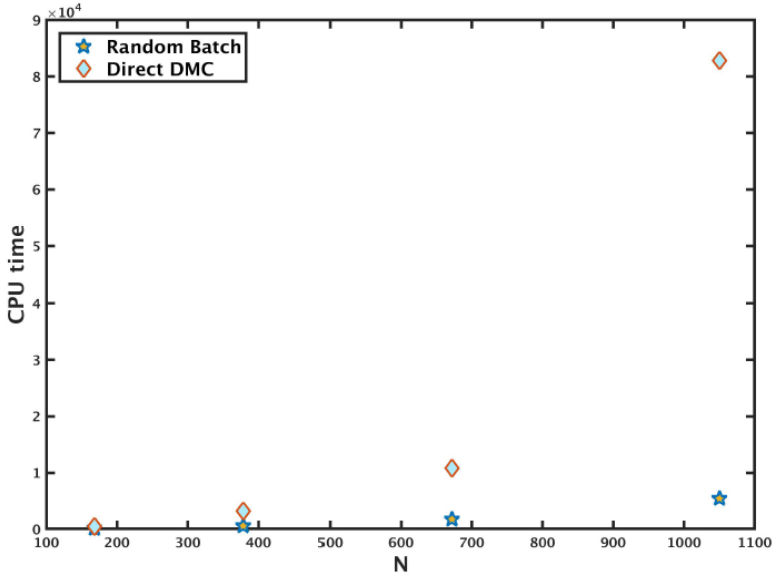


Figure 8: A comparison of the CPU runtime (in seconds) for running 1000 steps of DMC.

Quantum N-body problem
(with Francois Golse and Thierry Paul, Ecole Polytechnique-Paris)

- The N-body Schrodinger equation

$$i\hbar\partial_t\Psi(t,x_1,\ldots,x_N)=\mathcal{H}_N\Psi(t,x_1,\ldots,x_N)\,,\qquad \Psi\big|_{t=0}=\Psi^{in}\qquad x_1,\ldots,x_N\in\mathbf{R}^d$$

$$\mathcal{H}_N:=\sum_{m=1}^{N}-\tfrac{1}{2}\hbar^2\Delta_{x_m}+\tfrac{1}{N-1}\sum_{1\le l<n\le N}V(x_l-x_n)$$

- Main challenge in quantum simulation: large N, small $\hbar$
- One can use RBM to reduce the computational cost of the interacting potential terms from $O(N^2)$ to $O(N)$
- But unlike the classical case, the right hand side of the Schrodinger equation is of $O(N)$ thus any error introduced from the potential term will be amplified by N! so one cannot get an error for Ψ uniformly in N. Ever worse, the error is also divided by $\hbar$

The RBM (for p=2)

- Let $\sigma_1, \sigma_2, \ldots, \sigma_j, \ldots$ be the random permutation of {1, 2, ⋯ N}
- RBM operator:

$$\mathbf{T}_t(l,n) := \begin{cases} 1 & \text{if } \{l,n\} = \left\{\sigma_{[\frac{t}{\Delta t}]+1}(2k-1), \sigma_{[\frac{t}{\Delta t}]+1}(2k)\right\} \text{ for some } k = 1, \ldots, \frac{N}{2} \\ 0 & \text{otherwise,} \end{cases}$$

- RBM Hamiltonian

$$\mathcal{H}_N(t) := \sum_{m=1}^{N} -\tfrac{1}{2}\hbar^2 \Delta_{x_m} + \sum_{1 \le l < n \le N} \mathbf{T}_t(l,n) V(x_l - x_n)$$

Mathematical setting

$$\mathfrak{H} := L^2(\mathbf{R}^d; \mathbf{C}) \qquad \mathfrak{H}_N = \mathfrak{H}^{\otimes N} \simeq L^2((\mathbf{R}^d)^N; \mathbf{C})$$

- $\mathcal{L}(\mathfrak{H})$: The algebra of bounded operator on $\mathfrak{H}$
- Density operator: $R(t) = |\Psi(t)\rangle\langle\Psi(t)|$
- The N-body von Neumann equation

$$i\hbar\partial_t R(t) = \mathcal{H}_N R(t) - R(t)\mathcal{H}_N =: [\mathcal{H}_N, R(t)], \quad R(0) = R^{in}$$

- Solution of Schrodinger: $\Psi(t,\cdot) = e^{-it\mathcal{H}_N}\Psi^{in}$
- Solution of von Neumann: $R(t) = e^{-it\mathcal{H}_N/\hbar} R^{in} e^{it\mathcal{H}_N/\hbar}$

The random Batch von Neumann equation

$$i\hbar\partial_t \widetilde{R}(t) = [\mathcal{H}_N(t), \widetilde{R}(t)], \qquad \widetilde{R}(0) = R^{in}$$

Analytic solution

$$\widetilde{R}(t) = U(t,0)\widetilde{R}(0)U(0,t)$$

where, for each $0 \le s \le t$,

$$U(s,t) := e^{-\frac{i(s-[s/\Delta t]\Delta t)}{\hbar}\mathcal{H}_N([\frac{s}{\Delta t}]\Delta t)} \prod_{j=[s/\Delta t]}^{[t/\Delta t]-1} e^{\frac{i\Delta t}{\hbar}\mathcal{H}_N(j\Delta t)} e^{\frac{i(t-[t/\Delta t]\Delta t)}{\hbar}\mathcal{H}_N([\frac{t}{\Delta t}]\Delta t)},$$

$$U(t,s) := U(s,t)^* .$$

Notation: $\mathcal{U}(t,s)A := U(t,s)AU(s,t)$

Then $\widetilde{R}(t) = \mathcal{U}(t,0)R^{in}, \qquad t \ge 0$

Difficulties to obtain an N-independent error

- error: $\widetilde{R}(t) - \hat{R}(t)$
- What norm to use? trace norm (used in justifying the mean-field limit of N-body Schrodinger (Bardos-Golse-Mauser, 2000)) does not give an N nor $\hbar$ -independent error;

$$\|\mathbf{E}\widetilde{R}_1(t) - R_1(t)\|_1 = \sup_{\|A\|\leq 1} |\mathrm{trace}((\mathbf{E}\widetilde{R}_1(t) - R_1(t))A)| \quad \leq \frac{2N}{\hbar}\Delta t\|V\|_{L^\infty(\mathbf{R}^d)}\left(1 + \frac{Nt}{\hbar}\|V\|_{L^\infty(\mathbf{R}^d)}\right)$$

- Quantity of interest (QoI)? Reduced density operator.
- Loss of symmetry: if R^{in} has an integral kernel

$$r^{in} \equiv r^{in}(x_1, \ldots, x_N; y_1, \ldots, y_N)$$

such that $r^{in}(x_1, \ldots, x_N; y_1, \ldots, y_N) = r^{in}(x_{\sigma(1)}, \ldots, x_{\sigma(N)}; y_{\sigma(1)}, \ldots, y_{\sigma(N)})$

then $R(t)$ has the same symmetry.

This is no longer true for $\widetilde{R}(t)$

The one-body reduced density operator

- The 1-particle reduced density operator for $R(t)$ is $R_1(t)$ defined by its integral kernel

$$r_1(t,x,y) := \int_{(\mathbf{R}^d)^{N-1}} r(t;x,z_2,\dots,z_N;y,z_2,\dots,z_N)dz_2\dots dz_N$$

- Since the random batch potential $\sum_{1\le l<n\le N} \mathbf{T}_t(l,n)V(x_l-x_n)$ is not invariant under permutation, we define the 1-particle reduced density operator $\widetilde{R}_1(t)$ of $R_1(t)$ defined by its integral kernel

$$\widetilde{r}_1(t,x,y) := \frac{1}{N}\sum_{j=1}^{N}\int_{(\mathbf{R}^d)^{N-1}} \widetilde{r}(t;Z_{j,N}[x],Z_{j,N}[y])d\hat{Z}_{j,N}$$

where

$$Z_{j,N}[x] := z_1,\dots,z_{j-1},x,z_{j+1}\dots,z_N\,,\quad d\hat{Z}_{j,N} = dz_1\dots dz_{j-1}dz_{j+1}\dots dz_N$$

The norm we use

The Wigner transform

$$W_\hbar[S](x,\xi) := \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} s(x+\tfrac{1}{2}\hbar y, x-\tfrac{1}{2}\hbar y)e^{-i\xi\cdot y}dy$$

The dual norm

$$|||f|||_{-M} := \sup\left\{ \left| \iint_{\mathbf{R}^d\times\mathbf{R}^d} f(x,\xi)\overline{a(x,\xi)}dxd\xi \right| \;\middle|\; \begin{array}{l} a\in C_c(\mathbf{R}^d\times\mathbf{R}^d), \quad \text{and} \\ \max\limits_{\substack{|\alpha|,|\beta|\le M \\ |\alpha|+|\beta|>0}} \|\partial_x^\alpha\partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d\times\mathbf{R}^d)} \le 1 \end{array} \right\}$$

$$M \ge 1$$

The main result

Theorem 2.1. *Assume that $N \geq 2$ and that $V \in C(\mathbf{R}^d)$ is a real-valued function such that*

$$V(z) = V(-z) \text{ for all } z \in \mathbf{R}^d\,, \qquad \lim_{|z|\to+\infty} V(z) = 0\,, \qquad \text{and} \quad \int_{\mathbf{R}^d} (1+|\omega|^2)|\hat{V}(\omega)|d\omega < \infty\,.$$

Then there exists a constant $\gamma_d > 0$ depending only on the dimension d of the configuration space such that, for each $t > 0$, one has

$$\begin{aligned} &|||W_\hbar[\mathbf{E}\widetilde{R}_{N,\mathbf{1}}(t)] - W_\hbar[R_{N,\mathbf{1}}(t)]|||_{-[d/2]-3} \\ &\leq 2\gamma_d \Delta t e^{6t\max(1,\sqrt{d}L(V))}\Lambda(V)(2 + 3t\Lambda(V)\max(1,\Delta t) + 4\sqrt{d}L(V)t\Delta t)\,. \end{aligned} \tag{14}$$

where $\mathbf{E}$ is the mathematical expectation and Δt the reshuffling time-step in the definition of the random batch Hamiltonian (3), while

$$L(V) := \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} |\omega|^2|\hat{V}(\omega)|d\omega\,, \qquad \Lambda(V) := \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d}\sum_{\mu=1}^{d}|\omega^\mu||\hat{V}(\omega)|d\omega\,,$$

where ω^ν is the ν-th component of ω.

Remark

The error estimate is

(a) independent of N, and
(b) uniform in $\hbar \in (0, 1)$.

A new metric for density operators

Definition 4.1. *For $R, S \in \mathcal{D}(\mathfrak{H})$, set $D := -i\partial$ and*

$$d_\hbar(R,S) := \sup\left\{ |\operatorname{trace}_{\mathfrak{H}}((R-S)A)| \;\middle|\; \begin{array}{l} A \in \mathcal{L}(\mathfrak{H}) \text{ and for all } 1 \le \mu,\nu \le d \\ \hbar\|[x^\mu, A]\| + \hbar\|[\hbar D_\mu, A]\| + \|[x^\nu,[x^\mu,A]]\| \\ + \|[\hbar D_\nu,[x^\mu,A]]\| + \|[\hbar D_\nu,[\hbar D_\mu,A]]\| \le 5\hbar^2 \end{array} \right\}$$

This is analogous to Monge-Kantorovich or Wasserstein distance of exponent 1, note (Connes '89)

$$MK_1^\hbar(R,S) := \sup_{\substack{A\in\mathcal{L}(\mathfrak{H}) \\ \max_{1\le\mu\le d}(\|[x^\mu,A]\|,\|[\hbar D_\mu,A]) \le \hbar}} |\operatorname{trace}_{\mathfrak{H}}((R-S)A)| .$$

New idea: We need iterated commutators due to random batching

propositions

$$|||W_\hbar[R] - W_\hbar[S]|||_{-[d/2]-3} \leq \gamma_d d_\hbar(R,S)$$

$$\begin{aligned} d_\hbar(\mathbf{E}\widetilde{R}_{N,\mathbf{1}}(t), R_{N,\mathbf{1}}(t)) \leq & 10\Lambda(V)^2 t\Delta t \max(1,\Delta t) e^{6t\max(1,\sqrt{d}L(V))} \\ & + 10\Lambda(V)(1+2\sqrt{d}L(V)t\Delta t)\Delta t e^{t\max(1,2\sqrt{d}L(V))} \end{aligned}$$

Summary of key new ideas in the proof

- Use of $\|[f,T]\| \le \Lambda(f) \max_{1\le\nu\le d} \|[x^\nu,T]\|$ to transform the commutator with V to that with x

- Choice of $B_N(t) = \frac{1}{N}\sum_{k=1} J_k A$

  This gives rise to

$$[x^\mu_m, J_k A] = [-i\hbar\partial_{x^\mu_m}, J_k A] = 0 \text{ for all } \mu = 1,\dots,d \text{ and all } m = 1,\dots,N \text{ unless } m = k.$$

  this helps to get rid of N

- The use of new metric and choice of A (Weyl operator) makes

$$\|[x^\mu_m, A]\| = O(\hbar) \text{ and } \|[-i\hbar\partial_{x^\mu_m}, J_k A]\| = O(\hbar) \text{ for all } m = 1,\dots,N \text{ and } \mu = 1,\dots,d,$$

$$\text{so that } M_1(t) = O(\hbar)$$

  this helps to get rid of $\hbar$.

conclusions

The Newton and Schrodinger equations for N-particle system are two of the most fundamental equations in physics (with many other applications)

- We proposed simple random batch methods to reduce the computational cost of N-body interactions from $O(N^2)$ to $O(N)$
- We established rigorous error estimates uniform in N for classical particles, and uniformly in N and $\hbar$ for quantum systems
- Used it for both classical MD (with Coulomb interactions) and QMC—achieving the cost of O(N) with significant improvement on paralleliability
- Further questions (for both classical and quantum systems): convergence study for time-discretizations, convergence toward global equilibrium, theoretical justifications for mean-field and semiclassical limits, weaker regularity for the potentials, and other quantum systems, and more applications (QMC for Fermionic systems, for example)