Lecture II

Efficient Numerical Methods for Particle Systems

Shi Jin

Shanghai Jiao Tong University, China

CIRM-Jean-Morlet Chair Program: Research School

January 2021

Newton's equations—interacting particle systems

F=ma

$$dX^i = V^i\,dt,$$
$$dV^i = \left[b(X^i) + \alpha_N \sum_{j:j\neq i} K(X^i - X^j) - \gamma V^i\right] dt + \sigma\, dW^i$$

$\{W^i\}_{i=1}^N$ are i.i.d. Wiener processes, or the standard Brownian motions

Mean-field limit

$$\alpha_N = \frac{1}{N-1} \tag{1.4}$$

so that as $N \to \infty$ the empirical distribution $\mu^{(N)} := N^{-1}\sum_{i=1}^N \delta(x - X^i) \otimes \delta(v - V^i)$ converges almost surely under the weak topology to the solutions of the limiting PDE

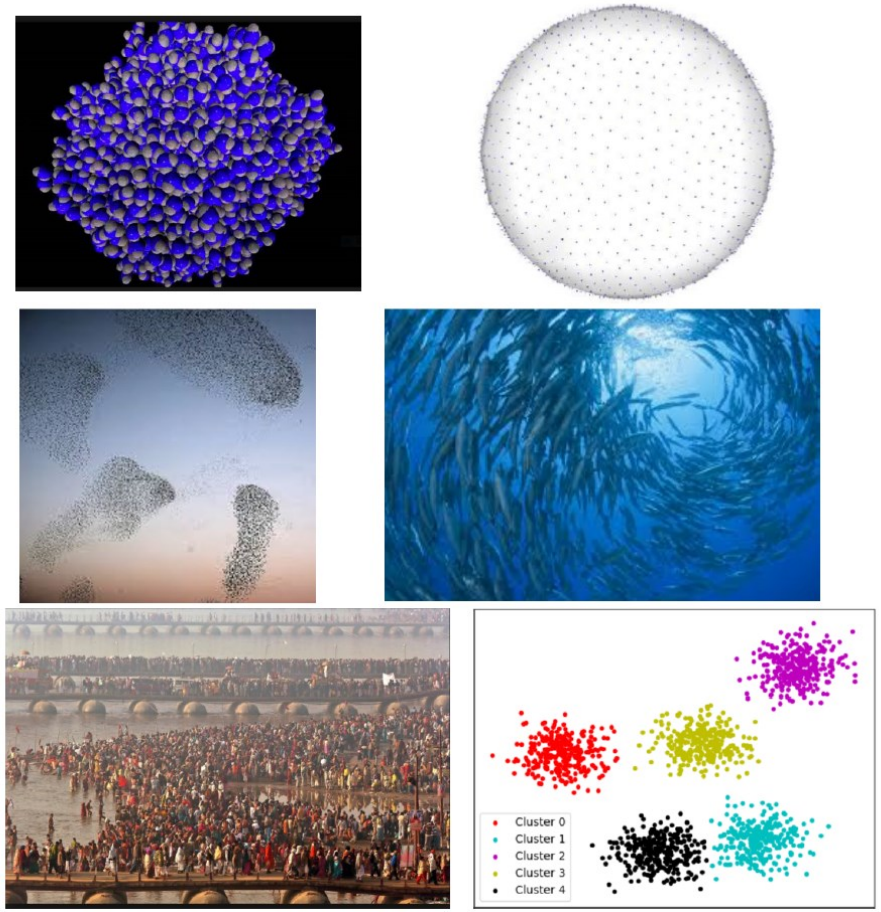
$$\partial_t f = -\nabla_x \cdot (vf) - \nabla_v \cdot \Big((b(x) + K *_x f - \gamma v)f\Big) + \frac{1}{2}\sigma^2 \Delta_v f. \tag{1.5}$$

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



- Course by P. Degond

How many particles?

Denote by N the number of particles or agents.

- In cosmology/astrophysics, N ranges from 10^{10} to $10^{20} - 10^{25}$; some models of dark matter even predict up to 10^{60} particles.
- In plasma dynamics, N is typically of order $10^{20} - 10^{25}$. This is the typical order of magnitude for physics settings.
- When used for numerical purposes (particles' methods...), the number is of order $10^{9} - 10^{12}$.
- In biology or Life Sciences, typical population of micro-organisms include between 10^{6} and 10^{12}.
- In other applications such as collective dynamics, Social Sciences or Economics, numbers can be much lower of order 10^{3}.

- 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 Method (*J-L. Li-J.-G. Liu, JCP '20*)

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-I (RBM-1)

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

$$dX^i = V^i\, dt,$$

$$dV^i = \left[b(X^i) + \frac{\alpha_N (N-1)}{p-1} \sum_{j \in \mathcal{C}_q, j \neq i} K(X^i - X^j) - \gamma V^i \right] dt + \sigma\, dW^i.$$

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

Algorithm II: RBM-r

RBM with replacement

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

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 $X^i \quad V^i$ $(i \in \mathcal{C}_k)$ by solving the following SDE for time τ

$$dX^i = V^i\,dt,$$
$$dV^i = \left[b(X^i) + \frac{\alpha_N(N-1)}{p-1} \sum_{j \in \mathcal{C}_q, j \neq i} K(X^i - X^j) - \gamma V^i\right] dt + \sigma\,dW^i$$

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)

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)

A strong convergence analysis for RBM-1

Notations

Analysis solution: (X_i, V_i)

RBM solution: $(\tilde{X}_i, \tilde{V}_i)$

synchronization condition: $X^i(0) = \tilde{X}^i(0) \sim \mu_0,\ W^i = \tilde{W}^i$

Norm of error: $\|v\| = \sqrt{\mathbb{E}|v|^2}$.

Mean-field scaling: $\alpha_N = 1/(N-1)$

Proposition 3.1. *Let $b(\cdot)$ be Lipschitz continuous, and $|b|, |\nabla b|$ have polynomial growth. The interaction kernel K is Lipschitz continuous. Then,*

$$\sup_{t\in[0,T]} \sqrt{\mathbb{E}|\tilde{X}^1 - X^1|^2 + \mathbb{E}|\tilde{V}^1 - V^1|^2} \le C(T)\sqrt{\frac{\tau}{p-1} + \tau^2}, \tag{3.5}$$

where $C(T)$ is independent of N.

A long-time result (*J-L. Li-Y. Sun* '20)

Assumption 3.1. Suppose $b = -\nabla U$ for some U that is bounded below $\inf_x U(x) > -\infty$, and there exist $\lambda_M \geq \lambda_m > 0$ such that the eigenvalues of $H := \nabla^2 U$ satisfy

$$\lambda_m \leq \lambda_i(x) \leq \lambda_M, \ \forall\ 1 \leq i \leq d, x \in \mathbb{R}^d.$$

The interaction kernel K is bounded and Lipschitz continuous. Moreover, the friction γ and the Lipschitz constant L of $K(\cdot)$ satisfy

$$\gamma > \sqrt{\lambda_M + 2L}, \ \lambda_m > 2L. \tag{3.6}$$

Theorem

$$\sup_{t\geq 0} \sqrt{\mathbb{E}|\tilde{X}^1(t) - X^1(t)|^2 + \mathbb{E}|\tilde{V}^1(t) - V^1(t)|^2} \leq C\sqrt{\frac{\tau}{p-1} + \tau^2},$$

where the constant C does not depend on p and N.

A key lemma (consistency)

$$\chi_i(\mathtt{x}) := \frac{1}{p-1}\sum_{j \imath \mathcal{C}} K(x^i - x^j) - \frac{1}{N-1}\sum_{j:j\neq i} K(x^i - x^j).$$

Lemma 3.2. *It holds that*

$$\mathbb{E}\chi_i(\mathtt{x}) = 0.$$

Moreover, the second moment is given by

$$\mathbb{E}|\chi_i(\mathtt{x})|^2 = \left(\frac{1}{p-1} - \frac{1}{N-1}\right)\Lambda_i(\mathtt{x}),$$

where

$$\Lambda_i(\mathtt{x}) := \frac{1}{N-2}\sum_{j:j\neq i}\left|K(x^i - x^j) - \frac{1}{N-1}\sum_{\ell:\ell\neq i} K(x^i - x^\ell)\right|^2.$$

stability

Introduce the filtration $\{\mathcal{F}_k\}_{k\geq 0}$ by

$$\mathcal{F}_{k-1} = \sigma(X_0^i, W^i(t), \mathcal{C}^{(j)}; t \leq t_{k-1}, j \leq k-1). \tag{3.4}$$

Thus, $\mathcal{F}_{k-1}$ is the σ-algebra generated by the initial values X_0^i $(i = 1, \ldots, N)$, $W^i(t)$, $t \leq t_{k-1}$, and $\mathcal{C}^{(j)}$, $j \leq k-1$. Clearly, $\mathcal{F}_{k-1}$ contains the information on how batches are constructed for $t \in [t_{k-1}, t_k)$.

Lemma 3.4. *Under Assumption 3.1, it holds for $q \geq 1$ that*

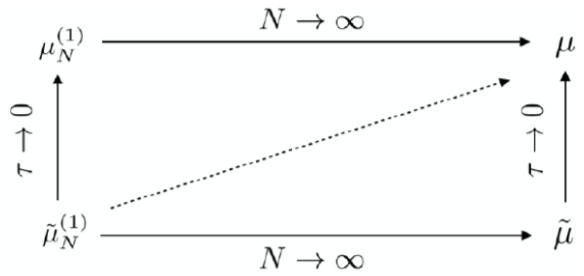
$$\sup_{t>0} \left(\mathbb{E}(|X^i(t)|^q + |V^i(t)|^q) + \mathbb{E}(|\tilde{X}^i(t)|^q + |\tilde{V}^i(t)|^q) \right) \leq C_q.$$

Besides, for any $k > 0$ and $q \geq 2$,

$$\sup_{t\in[t_{k-1},t_k)} \left| \mathbb{E}(|\tilde{X}^i(t)|^q + |\tilde{V}^i(t)|^q | \mathcal{F}_{k-1}) \right| \leq C(1 + |\tilde{X}^i(t_{k-1})|^q + |\tilde{V}^i(t_{k-1})|^q)$$

holds almost surely.

Mean-field limit and AP diagram (for first order systems)



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)

- Similar theorem for disparate species and weights (J-Li-J.G. Liu)

Example 1: 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

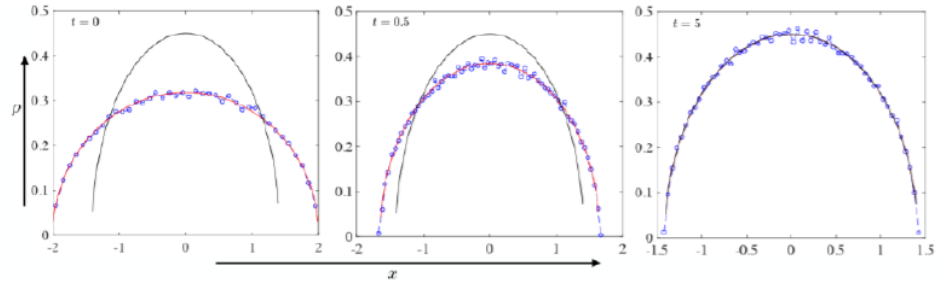


Figure 4: The RMB-1 simulation of the Dyson Brownian motion. The empirical densities at various times are plotted. The red curve is the density distribution predicted by the analytic solution (4.8). The black curve is the equilibrium semicircle law (4.7).

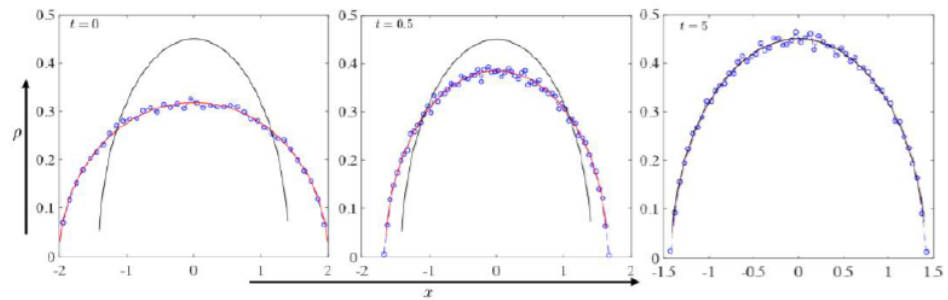


Figure 5: The RBM-r simulation of Dyson Brownian motion. The 'time' is regarded as $\tau = 10^{-3}$ for $N/2$ iterations. The red curve is the density distribution predicted by analytic solution (4.8). The black curve is the equilibrium semicircle law (4.7).

Example 2: 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.

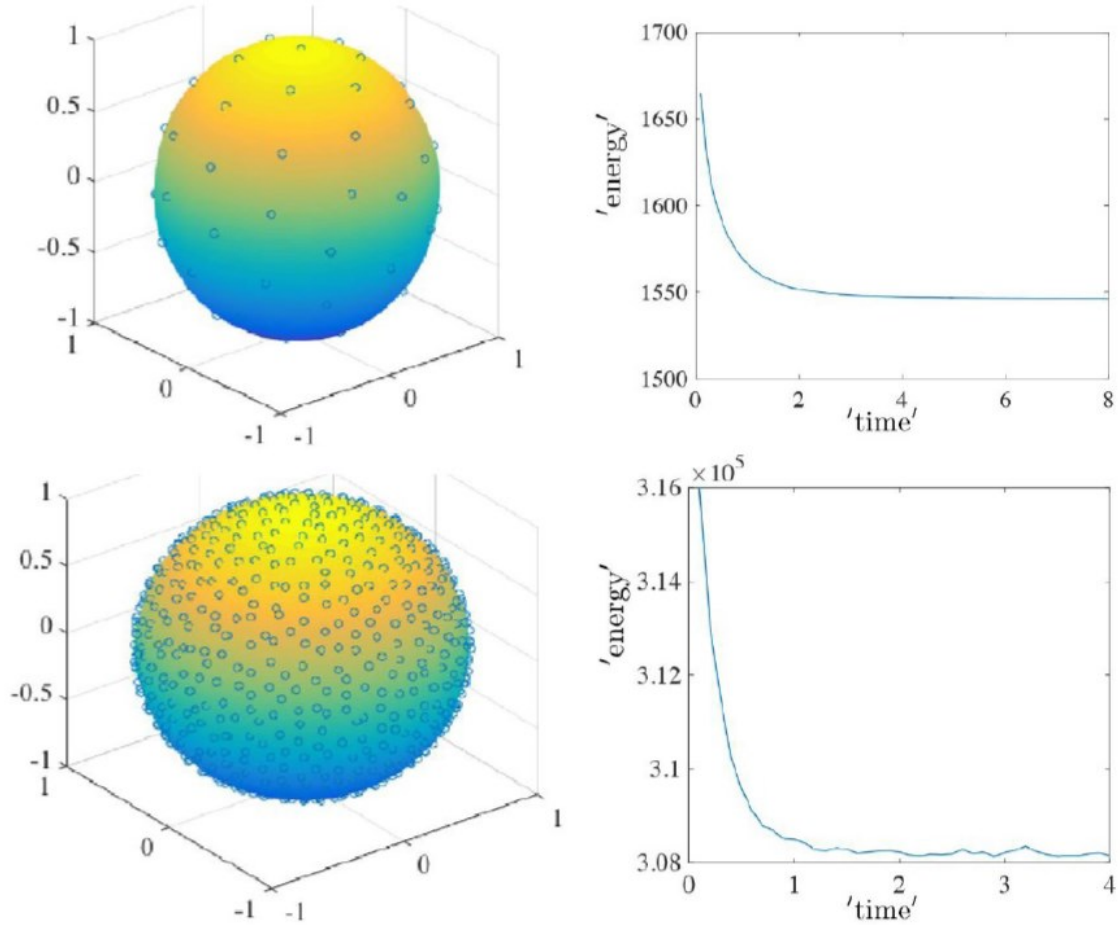
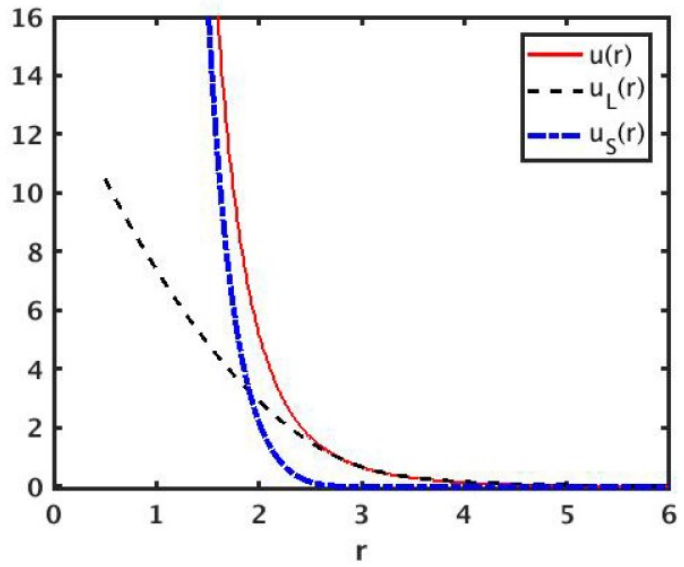


Figure 6: Charged particles on sphere. The first row is for $N = 60$ while the second row corresponds to $N = 800$. The first column shows the distributions at the end of simulation while the second column shows how the energy changes with 'time'.

Singular potential K

If K has singularity (for example Lenard-Jones potential at r=0)

decompose K into a short range, singular part (with cut-off) and long-range smooth part, and use RBM on the long-range part



Random Batch Ewald for Molecular Dynamics Simulation (with L. Li, Z. Xu and Y. Zhao)

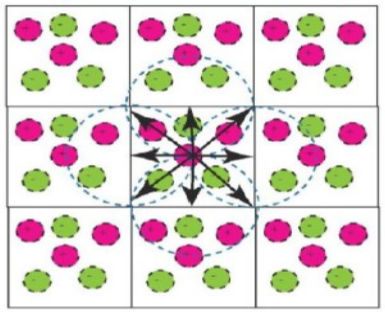
Potential energy function for particle simulations

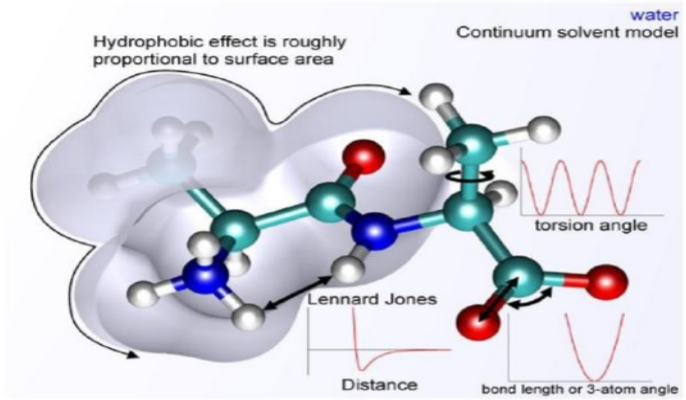
$$E = \sum_{bonds} k_b(b-b_0)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 + \sum_{torsions} k_w[cos(nw+\gamma)+1]$$

$$+ \sum_{atom\ pairs} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + \sum_{atom\ pairs} \frac{q_i q_j}{r_{ij}}$$

Long Range $O(N^2)$

Short range cut-off $O(N)$





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

Consider a periodic lattice

$$U = \frac{1}{2}\sum_{\boldsymbol{n}}{}' \sum_{i,j=1}^{N} q_i q_j \frac{1}{|\boldsymbol{r}_{ij}+\boldsymbol{n}L|} := U_1 + U_2$$

$$U_1 = \frac{1}{2}\sum_{\boldsymbol{n}}{}' \sum_{i,j} q_i q_j \frac{\mathrm{erf}(\sqrt{\alpha}|\boldsymbol{r}_{ij}+\boldsymbol{n}L|)}{|\boldsymbol{r}_{ij}+\boldsymbol{n}L|},$$

$$U_2 = \frac{1}{2}\sum_{\boldsymbol{n}}{}' \sum_{i,j} q_i q_j \frac{\mathrm{erfc}(\sqrt{\alpha}|\boldsymbol{r}_{ij}+\boldsymbol{n}L|)}{|\boldsymbol{r}_{ij}+\boldsymbol{n}L|}.$$

Change to Fourier space

$$U_1 = \frac{2\pi}{V}\sum_{\boldsymbol{k}\neq 0} \frac{1}{k^2}|\rho(\boldsymbol{k})|^2 e^{-k^2/4\alpha} - \sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} q_i^2,$$

$$\rho(\boldsymbol{k}) := \sum_{i=1}^{N} q_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}$$

Force:

$$\boldsymbol{F}_i = -\nabla_{\boldsymbol{r}_i} U = -\sum_{\boldsymbol{k}\neq 0} \frac{4\pi q_i \boldsymbol{k}}{Vk^2} e^{-k^2/(4\alpha)} \mathrm{Im}(e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}\rho(\boldsymbol{k}))$$

$$- q_i \sum_{j,\boldsymbol{n}}{}' q_j G(|\boldsymbol{r}_{ij}+\boldsymbol{n}L|)\frac{\boldsymbol{r}_{ij}+\boldsymbol{n}L}{|\boldsymbol{r}_{ij}+\boldsymbol{n}L|} =: \boldsymbol{F}_{i,1} + \boldsymbol{F}_{i,2},$$

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

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

Random Batch Ewald

PME (particle mesh Ewald) or PPPM (Particle-Particle-Particle Mesh): Cut off in Fourier space for high wave number

Random-batch Ewald: does't cut off high k, but uses *random mini-natch* sampled k from Gaussian distribution (importance sampling which reduces the variance)

$$\mathscr{P}_{\boldsymbol{k}} := S^{-1} e^{-k^2/(4\alpha)}$$

(sampling can be done offline and then randomly drawn p samples for each iteration)

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

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.

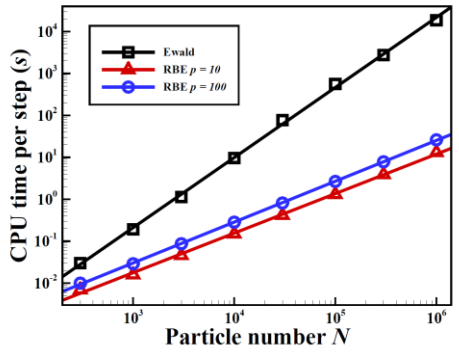
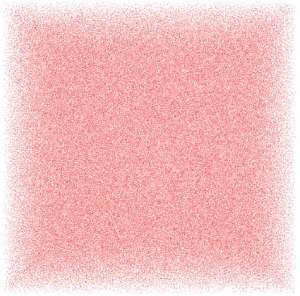


Figure 3: CPU time per step for the classical Ewald and the RBE methods with increasing N.

With L. Hong's group

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

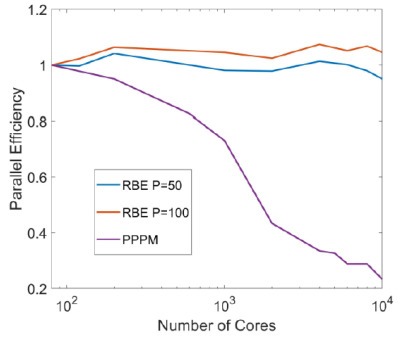


100nm

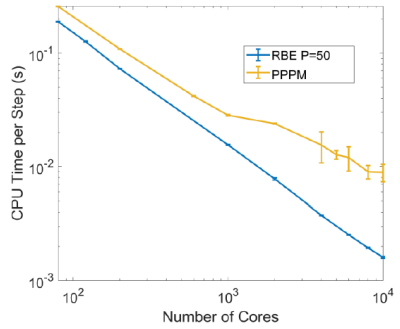
SJTU Π supercomputer

10,000,000 Water molecules

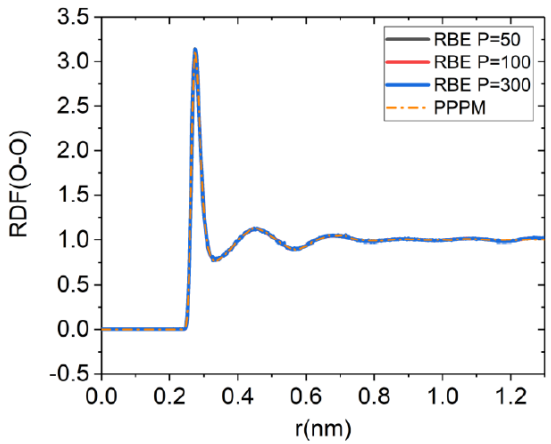
Much higher parallel efficiency



An order of magnitude faster (computational speed)

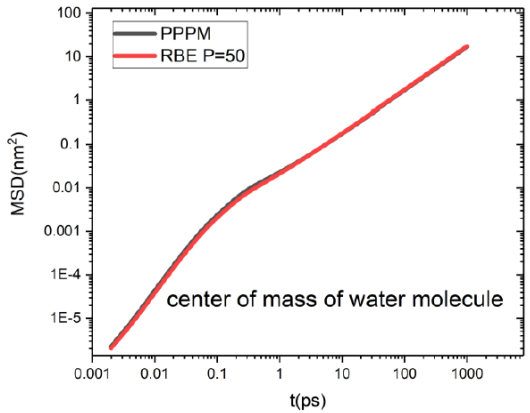


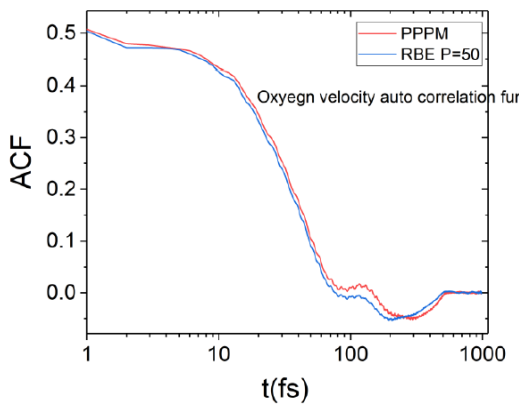
Accuracy of the structure and dynamics of the simulated water system



Radial distribution function of oxygen atoms between water molecules

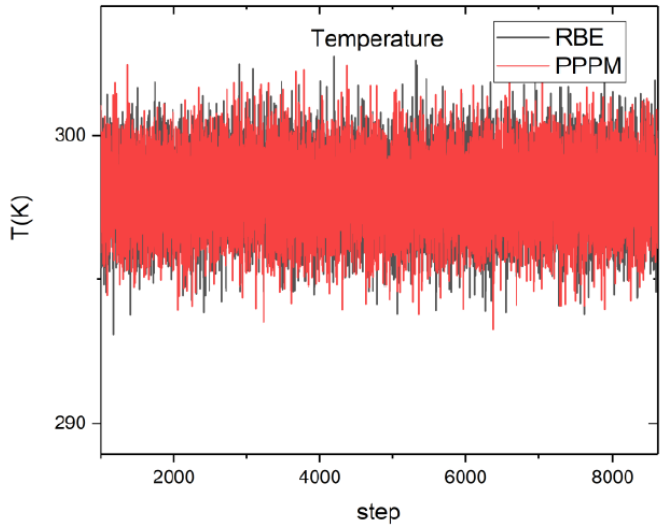
Mean-squared atomic displacement (from 2 fs to 1 ns)

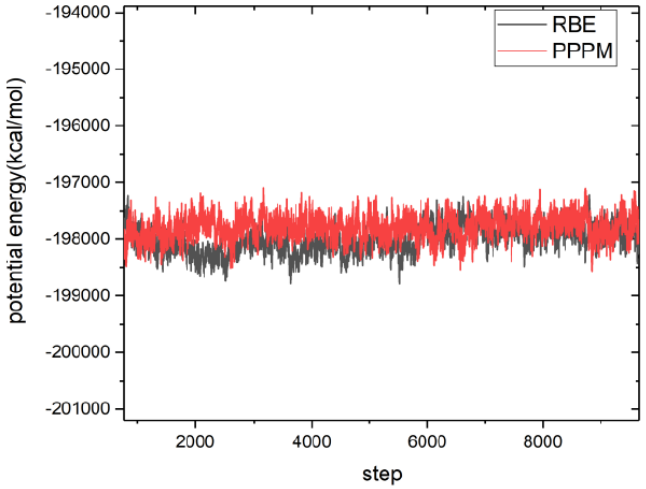




Velocity auto correlation function (from 1 fs to 1 ps)

Kinetic and potential energy fluctuations





Both fluctuations of the kinetic and potential energies agree with the PPPM

Other applications and 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
- Control of synchronalization in particle system
  E. Zuazua, etc.
- agent-based models for collective behavior (flocking, swarming, synchronization)
  S.Y. Ha, S Jin, D. Kim, D. Ko
- Quantum Monte-Carlo:
  Jin-X. Li
- Quantum N-body Schrodinger: uniform in $\hat{N}$ and $\hbar$ for reduced (1-body) density operator
  Golse-J-Paul