Efficient Interacting Particle Methods for Computing Near Singular Solutions of Keller-Segel Chemotaxis Systems and High-Dimensional Eigenvalue Problems

Zhiwen Zhang
Joint work with Drs. Zhizhang Wu, Zhongjian Wang, Junlong Lyu, and Prof. Jack Xin.

Caltech - Pasadena, California

November 2023

Outline

Motivations

- Many complex phenomena in science and engineering are modeled by PDEs, which are subject to specific boundary and initial conditions.
- Accurately solving these PDEs is crucial for simulating real-world phenomena, ranging from fluid flows and turbulent combustion to entropy production and population dynamics.
- Traditional mesh-based methods, such as finite element methods and spectral methods often encounter significant challenges when solving PDEs defined in high-dimensional spaces or with near singular solutions.
- In contrast, particle-based numerical methods that bypass the need for traditional mesh generation provide a promising solution to address these challenges.

Passive tracer models and effective diffusivities

- Consider a passive tracer model,

$$dX(t) = v(t, X)dt + \sigma dw(t), \quad X \in \mathbb{R}^d, \tag{1}$$

  where v is some spatial periodic field, $\sigma > 0$ is the molecular diffusivity, and $\{w(t)\}_{t\geq 0}$ is a standard Brownian motion (BM).
- Under certain conditions, the long-time larse-scale of $X(t)$ behaves like a BM, i.e., $\frac{X(t)}{\sqrt{2t}} \to N(0, D^E)$, $D^E \in \mathbb{R}^{d\times d}$ is called the effective diffusivity matrix.

Eulerian approach to compute D^E

- By homogenization theory, effective diffusion happens when v is mean zero and divergence-free.
- Let χ be solution of cell problem,

$$\mathcal{L}\chi = -v(y), \quad y \in \mathbb{T}^d, \tag{2}$$

where $\mathcal{L} := (v \cdot \nabla_x + D_0 \Delta_x)$ is the generator of X, $D_0 = \frac{\sigma^2}{2}$ is the elliptic coefficient. Then,

$$D^E = D_0 I + \langle v \otimes \chi \rangle_{\mathbb{T}^d} \tag{3}$$

- Numerically, if d is 3 or more and D_0 is 10^{-3} or smaller, χ will develop sharp gradients, which brings difficulties to mesh-based numerical methods.

Particle methods to compute D^E

- We can integrate along the particle path and calculate the covariance matrix of particles directly,

$$D_{ij}^E = \lim_{t\to\infty} \frac{\left\langle \left(x^i(t) - x^i(0)\right)\left(x^j(t) - x^j(0)\right)\right\rangle}{2t}, \quad 1 \le i,j \le d. \tag{4}$$

- We developed structure-preserving schemes for long-time integration and obtained uniform-in-time error analysis.

Theorem 1

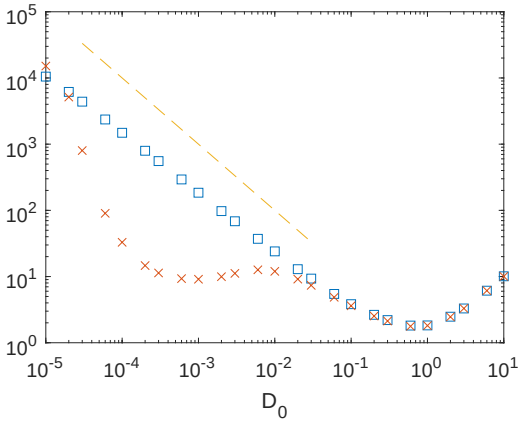
Let x_n^1, $n = 0, 1, \ldots$. be the first component of the numerical solution and Δt denote the time step. We have the convergence estimate

$$\lim_{n\to\infty} \frac{E(x_n^1)^2}{2n\Delta t} = D_{11}^E + \mathcal{O}(\Delta t), \tag{5}$$

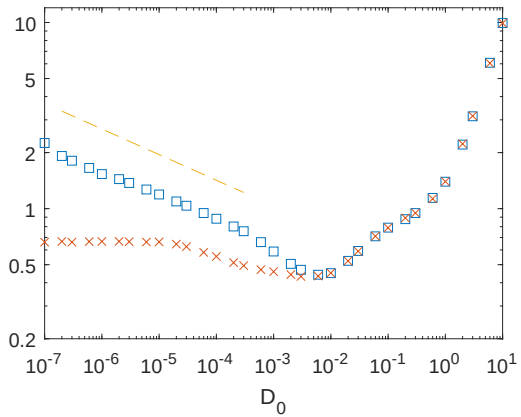
where the constant in $\mathcal{O}(\Delta t)$ may depend on the regularity of v and the constant σ but independent of T.

Convection-enhanced diffusion

- In ABC (Arnold–Beltrami– Childress) flow, $v = (\sin z + \cos y, \sin x + \cos z, \sin y + \cos x)$, and in Kolmogorov flow, $v = (\sin z, \sin x, \sin y)$.
- The calculation of effective diffusivities is new, especially in the 3D chaotic and convection-dominated flows.



(a) ABC flow, slope D_0^{-1}



(b) Kolmogorov flow, slope $D_0^{-0.19}$

Figure 1: Effective diffusivities in 3D chaotic flows. $\times$ is the result of conventional Euler-Maruyama method.

RDA equation with KPP nonlinearity

- Front propagation in complex fluid flows arises in many scientific areas such as turbulent combustion, chemical kinetics, biology, and transport in porous media.
- A fundamental problem is to analyze and compute large-scale front speeds in complex flows.
- An extensively studied model problem is the reaction diffusion advection (RDA) equation with Kolmogorov-Petrovsky-Piskunov (KPP) nonlinearity.
- To be specific, the KPP equation is

$$u_t = \kappa \Delta_{\mathbf{x}} u + (\mathbf{v} \cdot \nabla_{\mathbf{x}}) u + \tau^{-1} f(u), \quad t \in \mathbb{R}^+, \quad \mathbf{x} = (x_1, ..., x_d)^T \in \mathbb{R}^d, \tag{6}$$

where κ is diffusion constant, τ is the time scale of reaction rate, $\mathbf{v}$ is an incompressible velocity field (its precise definition will be discussed later), u is the concentration of reactant or population, and the KPP reaction term $f(u) = u(1-u)$ satisfying $f(u) \leq u f'(0)$.

Computing principal eigenvalues

- If the velocity field $\mathbf{v} = \mathbf{v}(\mathbf{x})$ in the KPP equation (6) is time-independent, the minimal front speed in direction $\mathbf{e}$ is given by the variational formula: $c^*(\mathbf{e}) = \inf_{\lambda>0} \mu(\lambda)/\lambda$, where $\mu(\lambda)$ is the principal eigenvalue of the elliptic operator

$$\mathcal{A}_1^\lambda \Phi \equiv \kappa \Delta_{\mathbf{x}} \Phi + (-2\kappa\lambda \mathbf{e} + \mathbf{v}) \cdot \nabla_{\mathbf{x}} \Phi + \left(\kappa\lambda^2 - \lambda \mathbf{v} \cdot \mathbf{e} + \tau^{-1} f'(0)\right)\Phi = \mu(\lambda)\Phi. \tag{7}$$

- If $\mathbf{v} = \mathbf{v}(t, \mathbf{x})$ in the KPP equation (6) is periodic in time t, then the variational formula $c^*(\mathbf{e}) = \inf_{\lambda>0} \mu(\lambda)/\lambda$ still holds, where $\mu(\lambda)$ is the principal eigenvalue of the time-periodic parabolic operator

$$\mathcal{A}_2^\lambda \Phi \equiv \kappa \Delta_{\mathbf{x}} \Phi + (-2\kappa\lambda \mathbf{e} + \mathbf{v}) \cdot \nabla_{\mathbf{x}} \Phi + \left(\kappa\lambda^2 - \lambda \mathbf{v} \cdot \mathbf{e} + \tau^{-1} f'(0)\right)\Phi - \Phi_t = \mu(\lambda)\Phi, \tag{8}$$

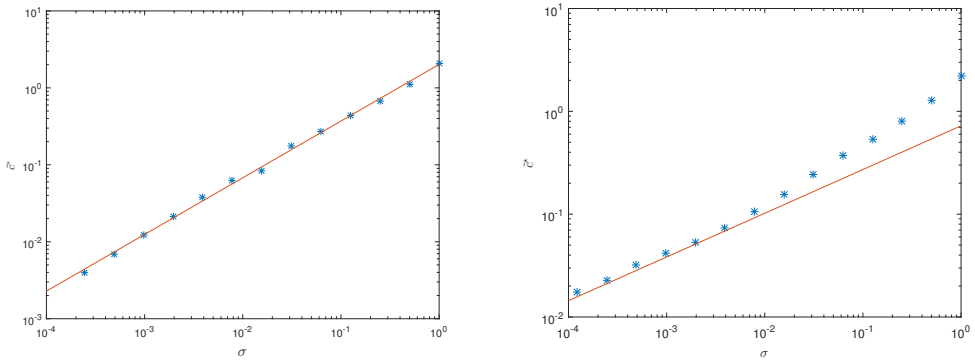
- However, when the magnitude of the velocity field is large (the problem becomes advection-dominated) and/or the dimension of spatial variables is big (e.g. $d = 3$), it is extremely expensive to compute KPP front speeds by using the FEM.

Interacting particle methods (IPMs)

- We develop interacting particle methods to compute KPP front speeds via the Feynman–Kac formula.
- We obtained accurate principal eigenvalues by studying the convergence of the Feynman–Kac semigroup associated with the SDE system and the potential from the operator $\mathcal{A}$.
- We also obtained error estimates of the proposed IPMs in computing principal eigenvalues.
- In the literature, the relation between KPP front speed and effective diffusivity, i.e., $c^*(A) = O(\sqrt{D^E(A)})$ is proved for the 2D steady cellular flow, where A is the strength of the flows.
- Our numerical result shows that this relation is still true for Kolmogorov flow.

Relation of KPP front speeds and effective diffusivity

- Figure 2 shows the numerical results of $\widetilde{c}^*(\sigma)$ in the 2D steady cellular flow obtained by our method. From the numerical results, we compute regression and obtain $\widetilde{c}^*(\sigma) = O(\sigma^{0.74})$, which agrees with the theoretical result $c^*(A) = O(\sqrt{D^E(A)})$, where $A = 1/\sigma$.
- Numerical results of $\widetilde{c}^*(\sigma)$ in 3D Kolmogorov flow. The fitted slope is ≈ 0.43, which also indiates that $c^*(A) = O(\sqrt{D^E(A)})$ is true.



(a) 2D cellular flow (b) 3D Kolmogorov flow

Figure 2: Numerical results of $\widetilde{c}^*(\sigma)$ in different flows.

Parabolic-Parabolic Keller-Segel Chemotaxis Systems

- We consider the parabolic-parabolic (fully parabolic) KS system of the form:

$$\begin{aligned} \rho_t &= \nabla \cdot (\mu \nabla \rho - \chi \rho \nabla c), \\ \epsilon c_t &= \Delta c - k^2 c + \rho, \end{aligned} \tag{9}$$

  where χ, μ (ϵ, k) are positive (non-negative) constants.
- The model is called elliptic if $\epsilon = 0$ (when c evolves rapidly to a local equilibrium), and parabolic if $\epsilon > 0$.
- The ρ is the density of active particles (bacteria), and c is the concentration of chemo-attractant (e.g. food).
- Several numerical methods, including finite-volume methods and spectral methods have been developed for KS systems to date. However, to the best of our knowledge, these numerical methods are tailored for 2D cases.

A stochastic interacting particle-field (SIPF) algorithm

- We propose a stochastic interacting particle-field (SIPF) algorithm for the fully parabolic KS system (9).
- Our method takes into account the coupled stochastic particle evolution (density ρ) and the accompanying field (concentration c) in the system and allows for a self-adaptive simulation of focusing and potentially singular behavior.
- In the SIPF algorithm, we represent the active particle density ρ by empirical particles and the concentration field c is discretized by a spectral method. This is possible since the field c is smoother than density ρ.
- We demonstrate the effectiveness of our method through numerical experiments in three space dimensions (3D), which have not been systematically computed and benchmarked to the best of our knowledge.

Numerical results for 3D Euler equations

- It is worth noting that the pseudo-spectral methods were employed to compute the nearly singular solutions of the 3D Euler equations by Profs. Hou and Li.
- Subsequently, the finite-time blowup of the 3D axisymmetric Euler equations was computed using the adaptive moving mesh method by Prof. Hou's group.
- These methods represent the cutting edge in the computation of nearly singular solutions of the 3D Euler equations.
- Nevertheless, we also point out that the implementation of pseudo-spectral methods for 3D problems demands substantial computational resources, while the adaptive moving mesh method requires sophisticated design and advanced programming skills.

Implementation of the SIPF algorithm

- Since we are interested in the spatially localized aggregation behavior, we restrict the system (9) in a large domain $\Omega = [-L/2, L/2]^d$ and assume Dirichlet boundary condition for particle density ρ and Neumann boundary condition for chemical concentration c.
- As a discrete algorithm, we assume the temporal domain $[0, T]$ is partitioned by $\{t_n\}_{n=0:n_T}$ with $t_0 = 0$ and $t_{n_T} = T$.
- We approximate the density ρ by particles, i.e.

$$\rho(t) \approx \frac{M_0}{P} \sum_{j=1}^{P} \delta(x - X_t^p), \quad P \gg 1, \tag{10}$$

where M_0 is the conserved total mass (integral of ρ).

Implementation of the SIPF algorithm (cont.)

- For chemical concentration c, we approximate by Fourier basis, namely, $c(\mathbf{x}, t)$ has an series representation

$$\sum_{j,m,l\in\mathcal{H}} \alpha_{t;j,m,l} \exp(i2\pi j\, x_1/L) \exp(i2\pi m\, x_2/L) \exp(i2\pi l\, x_3/L), \tag{11}$$

  where $\mathcal{H}$ denotes index set $\{(j,m,l) \in \mathbb{N}^3 : |j|, |m|, |l| \leq \frac{H}{2}\}$, and $i = \sqrt{-1}$.

- For ease of presenting our algorithm, with a slight abuse of notation, we use $\rho_n = \frac{M_0}{P} \sum_{p=1}^{P} \delta(x - X_n^p)$, and

$$c_n = \sum_{j,m,l\in\mathcal{H}} \alpha_{n;j,m,l} \exp(i2\pi j\, x_1/L) \exp(i2\pi m\, x_2/L) \exp(i2\pi l\, x_3/L)$$

  to represent density ρ and chemical concentration c at time t_n.

Updating chemical concentration c

- Let $\delta t = t_{n+1} - t_n > 0$ be the time step. We discretize the c equation of (9) in time by an implicit Euler scheme:

$$\epsilon\,(c_n - c_{n-1})/\delta t = (\Delta - k^2)\,c_n + \rho_n. \tag{12}$$

- It follows that:

$$c_n = c(\mathbf{x}, t_n) = -\mathcal{K}_{\epsilon,\delta t} * (\epsilon\, c_{n-1}/\delta t + \rho_n) = -\mathcal{K}_{\epsilon,\delta t} * (\epsilon\, c(\mathbf{x}, t_{n-1})/\delta t + \rho(x, t_n)) \tag{13}$$

  where $*$ is spatial convolution operator, and $\mathcal{K}_{\epsilon,\delta t}$ is the Green's function of the operator $\Delta - k^2 - \epsilon/\delta t$.

- In case of $\mathbb{R}^3$, the Green's function $\mathcal{K}_{\epsilon,\delta t}$ reads as follows

$$\mathcal{K}_{\epsilon,\delta t} = \mathcal{K}_{\epsilon,\delta t}(\mathbf{x}) = -\frac{\exp\{-\beta|\mathbf{x}|\}}{4\pi|\mathbf{x}|}. \quad \beta^2 = k^2 + \epsilon/\delta t, \tag{14}$$

Updating chemical concentration c (cont.)

- The Green's function admits a closed-form Fourier transform,

$$\mathcal{F}\mathcal{K}_{\epsilon,\delta t}(\omega) = -\frac{1}{|\omega|^2 + \beta^2}. \tag{15}$$

- For the term $-\mathcal{K}_{\epsilon,\delta t} * c_{n-1}$ in (13), by Eq.(15) it is equivalent to modify Fourier coefficients $\alpha_{j,m,l}$ to $\alpha_{j,m,l}/(4\pi^2 j^2/L^2 + 4\pi^2 m^2/L^2 + 4\pi^2 l^2/L^2 + \beta^2)$.
- For the second term $\mathcal{K}_{\epsilon,\delta t} * \rho$, we first approximate $\mathcal{K}_{\epsilon,\delta t}$ with $\cos$ series expansion, then according to the particle representation of ρ in (10),

$$(\mathcal{K}_{\epsilon,\delta t} * \rho)_{j,m,l} \approx \frac{M_0}{P} \sum_{p=1}^{P} \frac{\exp(-2\pi j X^p_{n,1}/L - 2\pi m X^p_{n,2}/L - 2\pi l X^p_{n,l}/L)(-1)^{j+m+l}}{4\pi^2 j^2/L^2 + 4\pi^2 m^2/L^2 + 4\pi^2 l^2/L^2 + \beta^2}. \tag{16}$$

Updating density of active particles ρ

- In the one-step update of density ρ_n represented by particles $\{X_n^p\}_{p=1:P}$, we apply Euler-Maruyama scheme to evolve particles:

$$X_{n+1}^p = X_n^p + \chi \nabla_x c(X_n^p, t_n)\delta t + \sqrt{2\mu\delta t} N_n^p, \tag{17}$$

  where N_n^p's are i.i.d. standard normal distributions.
- For $n > 1$, substituting (13) in (17) gives:

$$X_{n+1}^p = X_n^p - \chi \nabla_x \mathcal{K}_{\epsilon,\delta t} * (\epsilon\, c_{n-1}(\mathbf{x})/\delta t + \rho_n(\mathbf{x}))|_{x=X_n^p} \delta t + \sqrt{2\,\mu\,\delta t}\, N_n^p, \tag{18}$$

  from which $\rho_{n+1}(\mathbf{x})$ is constructed via (10).

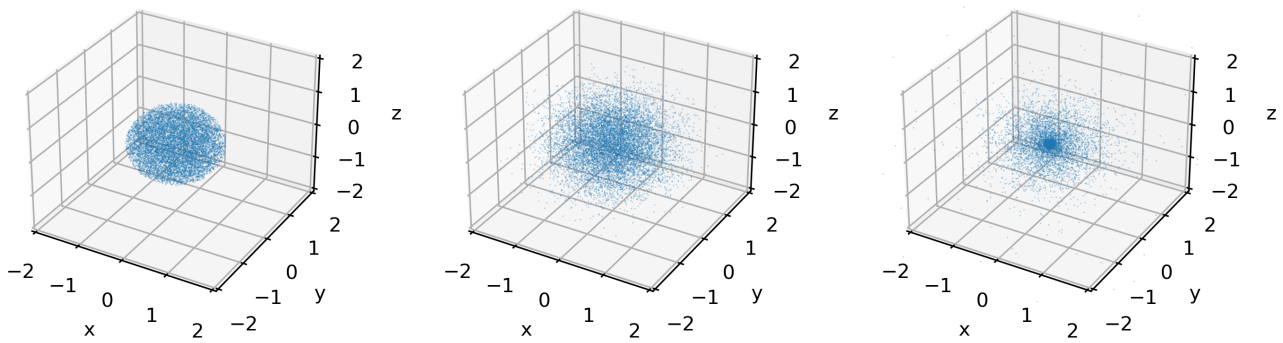
Updating density of active particles ρ (cont.)

- For $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * c_{n-1}(X_n^p)$, to avoid the singular points of $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$, we evaluate the integral with the quadrature points that are away from 0.
- Denote the standard quadrature point in Ω with $x_{j,m,l} = (jL/H, mL/H, jL/H)$, where j, m, l are integers ranging from $-H/2$ to $H/2-1$.
- When computing the integral $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * c_{n-1}(X_n^p)$, we evaluate $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$ at $\{X_n^p + \bar{X}_n^p - x_{j,m,l}\}_{j,m,l}$ where a small spatial shift $\bar{X}_n^p = \frac{H}{2L} + \lfloor \frac{X_n^p}{H/L} \rfloor \frac{H}{L} - X^p$ and c at $\{x_{j,m,l} - \bar{X}_n^p\}_{j,m,l}$ correspondingly.
- The latter one is computed by inverse Fourier transform of shifted coefficients, with $\alpha_{j,m,l}$ modified to $\alpha_{j,m,l}\exp(-i2\pi j\bar{X}_{n;1}^p/L - i2\pi m\bar{X}_{n;2}^p/L - i2\pi l\bar{X}_{n;3}^p/L)$, where $(\bar{X}_{n;i}^p)$ denotes the i-th component of $\bar{X}_n^p$.
- The term $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * \rho(X_n^p, t_n)$ is straightforward thanks to the particle representation of $\rho(X_n^p, t_n)$ in (10):

$$\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * \rho_n(X_n^p) = \int \mathcal{K}_{\epsilon,\delta t}(X_n^p - y)\rho(y) \approx \sum_{q=1, q\neq p}^{P} \frac{M}{P}\mathcal{K}_{\epsilon,\delta t}(X_n^p - X_n^q).$$

Aggregation Behaviors

- The initial distribution ρ_0 is assumed to be a uniform distribution over a ball centered at $(0, 0, 0)$ with radius 1, see Fig.3(a).
- We assume the following model parameters, $\mu = \chi = 1$, $\epsilon = 10^{-4}$ and $k = 10^{-1}$.
- The choice is made so that the model exhibits comparable behavior as the corresponding parabolic-elliptic KS system whose blow-up behavior is known.



(a) $T = 0$ (b) $T = 0.1, M_0 = 20$ (c) $T = 0.1, M_0 = 80$

Figure 3: Density ρ approximated by empirical distribution at $T = 0.1$: the mass effect on focusing.

An indicator of a possible blow-up

- If we assume, there exists a self-similar profile of ρ at origin $\rho(x,t) \sim \frac{1}{|x|^2}$, by (9), the Fourier coefficients of c has the asymptotics,

$$\mathcal{F}c(\omega) \sim \frac{1}{|\omega|^2 + k^2}\hat{\rho} \sim \frac{1}{(|\omega|^2 + k^2)|\omega|}. \tag{20}$$

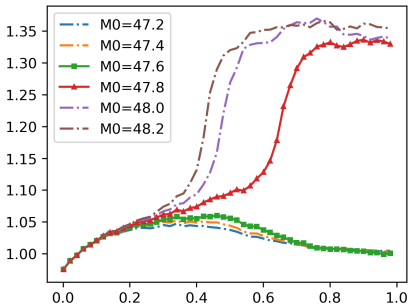
- Then the maximum of c in the computation shall vary vs the discretization parameter H. More precisely, we note at the origin,

$$c(0) \sim \int \frac{1}{(|\omega|^2 + k^2)|\omega|} e^{i\omega x} d\omega|_{x=0} = \int \frac{1}{(|\omega|^2 + k^2)|\omega|} d\omega. \tag{21}$$

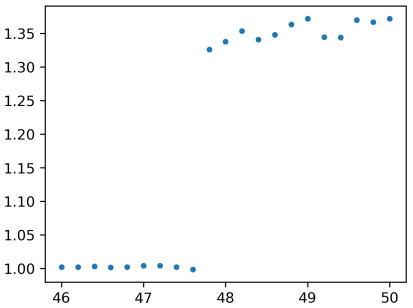
- In practical discretization, the range of integral (21) is related to the maximum frequency, namely $[-\frac{\pi}{L}(\frac{H}{2} - 1), \frac{\pi}{L} \cdot \frac{H}{2}]^3$. Then, for the type of $\frac{1}{|x|^2}$ profile blow up, $\|c\|_\infty = \mathcal{O}(\ln(H))$.
- Similarly for the type of $\delta(x)$ profile blow up, $\|c\|_\infty = \mathcal{O}(H)$.

Mass dependence

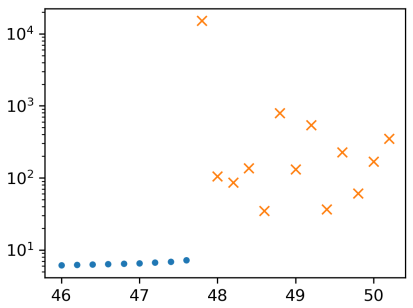
- The critical mass M_0 which plays the dominant role in the simple 2D parabolic elliptic system.
- We initialize the algorithm with uniform distribution over the unit ball centered at the origin and $c(0,x)=0$, and apply the SIPF with two different H to compute the density and chemical concentration.
- We validate the accuracy of SIPF by comparison with the result of the radial solution obtained by FDM.
- Our SIPF applies directly to more general (non-radial) KS systems.



(a) $\frac{|c|_{\infty,H=24}}{|c|_{\infty,H=12}}$ vs. computation time T



(b) $\frac{|c|_{\infty,H=24}}{|c|_{\infty,H=12}}$ at $T=1$ vs. M_0.



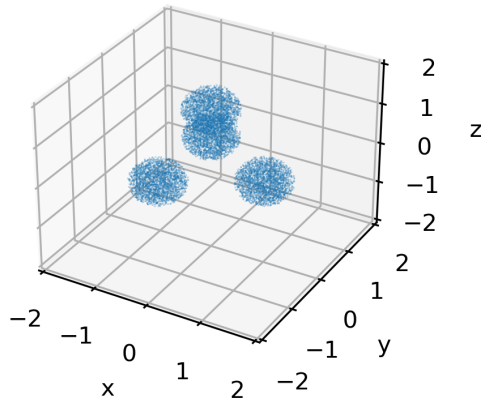
(c) $|c|_{\infty,FDM}$ at $T=1$ vs. M_0. $\times$ denotes unstable results.

Aggregation behaviors from non-radial initial data

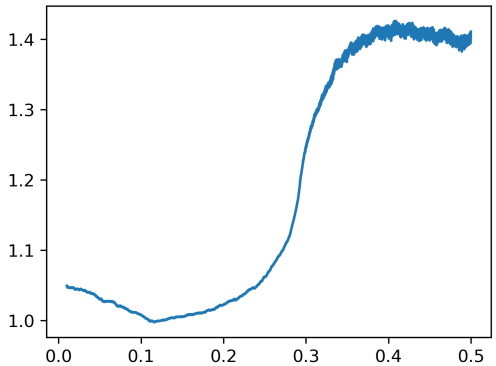
- We consider a more practical scenario where the initial distribution ρ models several separated clusters of organisms and the mass in each individual cluster is below the critical mass while the total mass is super-critical.
- We assume the initial distribution is a uniform distribution on four balls with a radius 0.5 and centered at four vertices of a regular tetrahedron, namely, $(1,0,0)$, $(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$, $(-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0)$ and $(0,0,\sqrt{2})$.
- We assume the total mass to be $M_0 = 80$ and so each cluster has a mass of 20 which is below the critical mass for a ball with radius $r = 0.5$.

Aggregation behaviors from non-radial initial data

- Then we apply the algorithm to compute the KS system up to $T = 0.5$ with $H = 24$ and $H = 12$ while keeping the rest of the configurations.
- In Fig.5(b), we compute the ratio between the maxima of c vs time with two different spatial discretizations. We can see the singularities formed in the system at around $T = 0.3$.



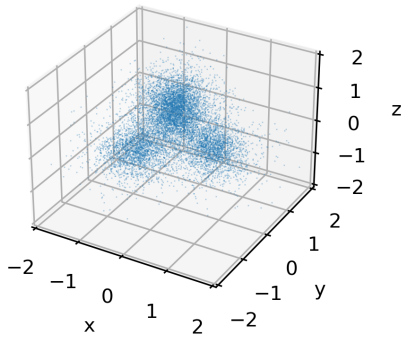
(a) Initial distribution.



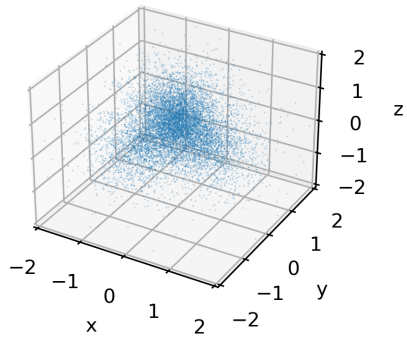
(b) $\frac{|c|_{\infty,H=24}}{|c|_{\infty,H=12}}$ vs. time.

Figure 5: Identifying the formation of a finite time singularity at $t \approx 0.3$ in non-radial solutions.

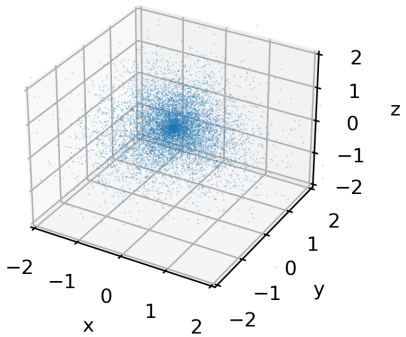
Aggregation behaviors from non-radial initial data



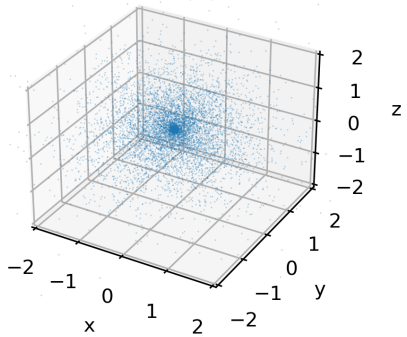
(a) $T = 0.1$



(b) $T = 0.2$



(c) $T = 0.3$



(d) $T = 0.4$

Figure 6: Scatter plots of four clusters merging and a singularity formation.

Entropy production rate

- We are interested in computing the entropy production rate to quantify the time reversal of an SDE of the form:

$$d\mathbf{X}_s = -\nabla V(\mathbf{X}_s)ds + \mathbf{b}(\mathbf{X}_s)ds + \sqrt{2\varepsilon}d\mathbf{W}_s, \quad \mathbf{X}_s \in \mathbb{R}^d, \tag{22}$$

  where $\varepsilon > 0$, V is a smooth potential function with at least quadratic growth at infinity, $\mathbf{b}$ is a bounded vector field and nonconservative, $\mathbf{b}$ and ∇V are orthogonal, and $\mathbf{W}_s$ is a standard Brownian motion.
- The problem of time reversibility in diffusion processes was first studied by Kolmogorov in 1937.
- The time-reversed diffusion process to map the terminal distribution back to the initial distribution is in general governed by a different SDE unless $\mathbf{b} = 0$.
- Interestingly, the concept of time-reversed diffusion (i.e., diffusion models) has been adopted in recent years as an effective way to generate high-quality images in computer vision.

Moment generating function

- One can quantify time reversal through the entropy integral $S_t = \varepsilon^{-1} \int_0^t \langle \mathbf{b}(\mathbf{X}_s), \circ d\mathbf{X}_s \rangle$, which is the work done by the non-gradient part of the Hodge decomposition of the drift force in (22).
- Here $\circ d\mathbf{X}_s$ denotes the Stratonovich integral with respect to $\mathbf{X}_s$. Let $P_t^{\mu,\varepsilon}$ denote the probability measure of (22) from the initial measure μ (i.e., $\mathbf{X}_0 \sim \mu$).
- Its moment generating function is $\chi_t^\varepsilon(\alpha) = \int_{C_t} \exp(-\alpha S_t^\varepsilon) dP_t^{\mu,\varepsilon}$, where $\alpha \in \mathbb{R}$ and C_t is the space $C([0,t];\mathbb{R}^d)$ of continuous paths in $\mathbb{R}^d$ over the time interval $[0,t]$.

principal eigenvalue

- The representation holds: $\chi_t^{\varepsilon}(\alpha) = \int_{\mathbb{R}^d} g^{-\alpha}\big(\exp(tA^{\varepsilon,\alpha})g^{\alpha}\big)d\mu$, where g is a continuous function and the infinitesimal operator $A^{\varepsilon,\alpha}$ acts on smooth and compactly supported functions f as follows:

$$A^{\varepsilon,\alpha}f = \varepsilon\Delta f + \big(-\nabla V + (1-2\alpha)\mathbf{b}\big)\cdot\nabla f - \frac{\alpha(1-\alpha)}{\varepsilon}|\mathbf{b}|^2 f + \frac{\alpha}{\varepsilon}(\mathbf{b}\cdot\nabla V)f - \alpha(\nabla\cdot\mathbf{b})f \tag{23}$$

  over $\alpha \in [-\delta, 1+\delta]$ for some small $\delta > 0$. Let $\lambda^{\varepsilon,\alpha}$ denote the principal eigenvalue of $A^{\varepsilon,\alpha}$.
- Then $\lim_{t\to\infty} t^{-1}\log\chi_t^{\varepsilon}(\alpha) = \lambda^{\varepsilon,\alpha}$, which is convex in α and symmetric about $\alpha = 1/2$. The Legendre transform of $\lambda^{\varepsilon,\alpha}$ in α is the large deviation rate function of $t^{-1}S_t^{\varepsilon}$.
- Hence, it characterizes the stochastic growth rate of entropy.

- The operator $A^{\varepsilon,\alpha}$ is unitarily equivalent to

$$\begin{aligned}\mathcal{A}^{\varepsilon,\alpha} f &:= \exp\left((-2\varepsilon)^{-1}V\right) A^{\varepsilon,\alpha}\left(\exp((2\varepsilon)^{-1}V)\right) f \qquad (24)\\ &= \varepsilon\Delta f + (1-2\alpha)\mathbf{b}\cdot\nabla f - \frac{1}{4\varepsilon}|\nabla V|^2 f + \frac{1}{2\varepsilon}(\mathbf{b}\cdot\nabla V) f\\ &- \frac{\alpha(1-\alpha)}{\varepsilon}|\mathbf{b}|^2 f + \frac{1}{2}(\Delta V) f - \alpha(\nabla\cdot\mathbf{b}) f.\end{aligned}$$

- $\mathcal{A}^{\varepsilon,\alpha}$ shares the same principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and shows that the quadratic approximation near the zeros of $\frac{1}{4}|\nabla V|^2 - \frac{1}{2}\mathbf{b}\cdot\nabla V + \alpha(1-\alpha)|\mathbf{b}|^2$ is crucial in studying vanishing-noise limit as $\varepsilon \to 0$.
- There are several difficulties involved in computing the principal eigenvalue $\lambda^{\varepsilon,\alpha}$. (1) The operators $A^{\varepsilon,\alpha}$ and $\mathcal{A}^{\varepsilon,\alpha}$ are defined in high-dimensional spaces and non-self-adjoint. (2) When studying the vanishing-noise limit as $\varepsilon \to 0$, the operators $A^{\varepsilon,\alpha}$ and $\mathcal{A}^{\varepsilon,\alpha}$ become singularly perturbed.

Feymann-Kac semigroup formulation

- The Feynman-Kac formula establishes a connection between PDEs and SDEs, which inspires us to develop particle-based methods to compute the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ of $\mathcal{A}^{\varepsilon,\alpha}$.
- For simplicity, we will temporarily suppress the parameters ε and α. We decompose the operator $\mathcal{A}$ into $\mathcal{A} = \mathcal{L} + \mathcal{U}$, where $\mathcal{L} := \varepsilon\Delta + (1-2\alpha)\mathbf{b}\cdot\nabla$ and $\mathcal{U} := -\frac{1}{4\varepsilon}|\nabla V|^2 + \frac{1}{2\varepsilon}\mathbf{b}\cdot\nabla V - \frac{\alpha(1-\alpha)}{\varepsilon}|\mathbf{b}|^2 + \frac{1}{2}(\Delta V) - \alpha(\nabla\cdot\mathbf{b})$.
- We consider the SDE with $\mathcal{L}$ as the associated infinitesimal generator: $d\mathbf{X}_t = (1-2\alpha)\mathbf{b}dt + \sqrt{2\varepsilon}d\mathbf{B}_t$, where $\mathbf{B}_t$ is a d-dimensional Brownian motion.
- Then, we define an evolution operator $P_t^{\mathcal{U}}$ as $P_t^{\mathcal{U}}\varphi(x) = \mathbb{E}\left[\varphi(\mathbf{X}_t)\exp\left(\int_0^t \mathcal{U}(\mathbf{X}_s)ds\right)|\mathbf{X}_0 = x\right]$, where $\mathbb{E}$ is the expectation with respect to the Brownian motion and φ is a measurable function.
- Note that $P_t^{\mathcal{U}} = \exp(t(\mathcal{L}+\mathcal{U})) = \exp(t\mathcal{A})$ by the Feynman-Kac formula.

- We consider the Feynman-Kac semigroup associated with $P_t^{\mathcal{U}}$ as

$$\Theta_t(\mu)(\varphi) = \frac{\mu(P_t^{\mathcal{U}}\varphi)}{\mu(P_t^{\mathcal{U}}\mathbb{1})} = \frac{\mathbb{E}\big[\varphi(\mathbf{X}_t)\exp\big(\int_0^t \mathcal{U}(\mathbf{X}_s)ds\big)|\mathbf{X}_0 \sim \mu\big]}{\mathbb{E}\big[\exp\big(\int_0^t \mathcal{U}(\mathbf{X}_s)ds\big)|\mathbf{X}_0 \sim \mu\big]}, \quad \mu \in \mathcal{P}(\mathbb{R}^d), \tag{25}$$

  where μ is an initial distribution and $\mathcal{P}(\mathbb{R}^d)$ is the space of all probability measures over $\mathbb{R}^d$.
- Under certain assumptions for V and $\mathbf{b}$, we prove that there exists a unique invariant measure $\mu_{\mathcal{U}}^\star \in \mathcal{P}(\mathbb{R}^d)$ such that, for any $\mu \in \mathcal{P}(\mathbb{R}^d)$ and suitable bounded function φ, we have $|\Theta_t(\mu)(\varphi) - \mu_{\mathcal{U}}^\star(\varphi)| \leq C_\mu \exp(-\kappa t)||\varphi||_{L^\infty}$, where $C_\mu > 0$ and $\kappa > 0$.
- In addition, we prove that $\lambda = \lim_{t\to\infty} t^{-1}\log\mathbb{E}\big[\exp\big(\int_0^t \mathcal{U}(\mathbf{X}_s)ds\big)|\mathbf{X}_0 \sim \mu\big]$.
- This elegant result enables us to develop particle-based numerical methods to compute the principal eigenvalue λ.

Numerical dicretization

- To compute the leading eigenvalue λ, we need to consider the discretization of the operator P_t^U, which consists of two steps, an operator splitting scheme for P_t^U and subsequently an Euler-Maruyama scheme for the SDE.
- With a time step size $\Delta t > 0$, define an evolution operator $\widetilde{P}_{\Delta t}^U$ as

$$\widetilde{P}_{\Delta t}^U \varphi(x) = \exp(\Delta t U(x)) \mathbb{E}\left[\varphi(X_{\Delta t}) | X_0 = x\right], \tag{26}$$

  where $X_{\Delta t}$ satisfies the SDE and φ is a measurable function.
- Note that if we define an operator P_t as

$$P_t \varphi(x) = \mathbb{E}\left[\varphi(X_t) | X_0 = x\right], \text{ for any } \varphi \text{ measurable}, \tag{27}$$

  then $P_t = \exp(t\mathcal{L})$. Hence, $\widetilde{P}_{\Delta t}^U = \exp(\Delta t U) \exp(\Delta t \mathcal{L})$ can be seen as an approximation of $P_{\Delta t}^U$ using an operator splitting scheme.

The corresponding Feynman-Kac semigroup

- Analogously, if we define $\widehat{\lambda}_{\Delta t} = \frac{1}{\Delta t}\log(\widehat{\Lambda}_{\Delta t})$, $\widehat{\lambda}_{\Delta t}$ is also an approximation of λ.
- We then define the corresponding Feynman-Kac semigroup of $Q^U_{\Delta t}$ as

$$\Phi_{k,\Delta t}(\mu)(\varphi) = \frac{\mu((Q^U_{\Delta t})^k \varphi)}{\mu((Q^U_{\Delta t})^k \mathbb{1})} = \frac{\mathbb{E}\left[\varphi(X_k)\exp\left(\Delta t \sum_{j=0}^{k-1} U(X_j)\right)\middle| X_0 \sim \mu\right]}{\mathbb{E}\left[\exp\left(\Delta t \sum_{j=0}^{k-1} U(X_j)\right)\middle| X_0 \sim \mu\right]} \tag{28}$$

for any initial measure μ and any measurable function φ.

- We have the following theorem that shows the stability of $\Phi_{k,\Delta t}$ and gives a representation of $\widehat{\lambda}_{\Delta t}$.

Theorem 2

Under certain assumptions for V and b, there exists an invariant measure $\widehat{\mu}^{\star}_{U,\Delta t} \in \mathcal{P}(\mathbb{R}^d)$ and $\widehat{\beta} \in (0,1)$ such that for any initial measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, there is C_μ for which

$$|\Phi_{k,\Delta t}(\mu)(\varphi) - \widehat{\mu}^{\star}_{U,\Delta t}(\varphi)| \leq C_\mu \widehat{\beta}^k ||\varphi||_{L^\infty}, \quad \forall \varphi \in L^\infty(\mathbb{R}^d), \quad \forall k \geq 1, \tag{29}$$

where $\widehat{\beta} \in (0,1)$ and $\Phi_{k,\Delta t}(\widehat{\mu}^{\star}_{U,\Delta t}) = \widehat{\mu}^{\star}_{U,\Delta t}$. Moreover,

$$\widehat{\lambda}_{\Delta t} = \lim_{k \to +\infty} \frac{1}{k\Delta t} \log \mathbb{E}\left[\exp\left(\Delta t \sum_{j=1}^{k-1} U(X_j) \right) \middle| X_0 \sim \mu \right]. \tag{30}$$

Implementation of the interacting particle method

- The leading eigenvalue λ can be represented as a scaled cumulant generation function and can be approximated using a particle approach.
- However, if we compute the quantity using a direct Monte Carlo simulation, the variance of the estimator will increase exponentially in time, which leads to numerical instability.
- Therefore, we need to supplement the computation with a multinomial resampling technique, which acts as a variance reduction technique.

The complete algorithm of the interacting particle method

Algorithm 1 The stochastic interacting particle method for computing the entropy

Input: velocity field b, potential V, the number of particles M, initial distribution μ, final time T, time step size $\Delta t = \frac{T}{N}$.

1: Generate M i.i.d. μ-distributed particles $\{\mathbf{q}^{0,m} = (q_1^{0,m}, \ldots, q_d^{0,m})\}_{m=1}^M$.
2: **for** n = 1:N **do**
3: For $m = 1, \ldots, M$, compute $\widetilde{\mathbf{q}}^{n,m}$ using the Euler Maruyama scheme (3.7) with $\mathbf{q}^{n-1,m}$ as the initial value.
4: For $m = 1, \ldots, M$, compute the weight $w^{n-1,m} = \exp(\Delta t U(\mathbf{q}^{n-1,m}))$ and the probability $p^{n-1,m} = w^{n-1,m}/P^{n-1}$, where $P^{n-1} = \sum_{m=1}^M w^{n-1,m}$.
5: Resample $\{\widetilde{\mathbf{q}}^{n,m}\}_{m=1}^M$ according to the multinomial distribution associated with the probability $\{p^{n-1,m}\}_{m=1}^M$, which gives the particle $\{\mathbf{q}^{n,m}\}_{m=1}^M$ at the n-th time step.
6: Compute the quantity $\widehat{\lambda}^{n-1} = \log(P^{n-1}/M)$.
7: **end for**
8: Compute the approximation of the leading eigenvalue $\widehat{\lambda}_{\Delta t} = \frac{1}{T}\sum_{n=0}^{N-1} \widehat{\lambda}^n$.

Output: the approximation of the leading eigenvalue $\widehat{\lambda}_{\Delta t}$.

A 16D result

- We consider a 16D mixed single-well and double-well potential

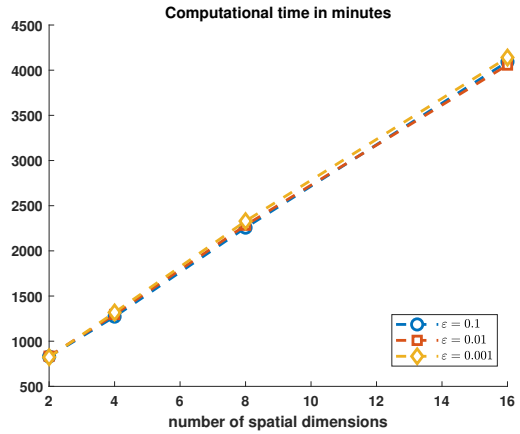
$$V(x_1, x_2, ..., x_{15}, x_{16}) = \sum_{i=1}^{4} \left(\frac{4x_{2i-1}^2 + x_{2i-1}^4}{8} + \frac{4x_{2i}^2 + x_{2i}^4}{8} \right)$$
$$+ \sum_{i=1}^{4} \left(x_{2i+7}^4 - 2x_{2i+7}^2 + (1 + a_i(x_{2i+7} - 1)^2)x_{2i+8}^2 + x_{2i+8}^4 \right),$$

  where $a_1 = 0.2$, $a_2 = 0.7$, $a_3 = 0.5$, $a_4 = 0.3$.

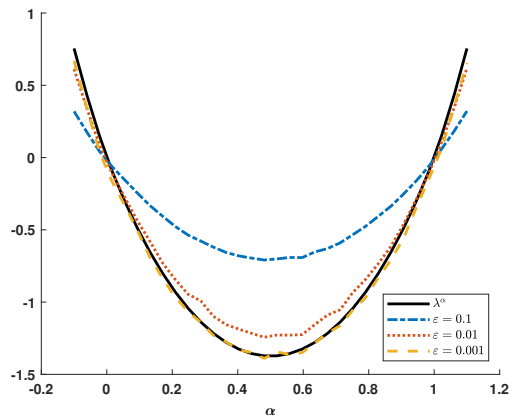
- The $\mathbf{b} = (b_1, b_2, ..., b_{15}, b_{16})$, where $b_{2i-1} = \pi^{-1} \cos(\pi x_{2i-1}) \sin(\pi x_{2i})$ and $b_{2i} = -\pi^{-1} \sin(\pi x_{2i-1}) \cos(\pi x_{2i})$, for $i = 1, ..., 8$.

A 16D result

- We choose $M = 500,000, T = 2048, \Delta t = 2^{-8}$ in our method and choose the initial distribution to be the standard Gaussian distribution.
- We also let $\alpha \in [-0.1, 1.1]$.
- We can see that the computational time grows linearly with the number of spatial dimensions and does not change much as ε varies.
- Right is the computed eigenvalues for $\varepsilon = 0.1, 0.01, 0.001$ and different α's, where a convergent empirical distribution for a larger ε serves as the initial distributions for a smaller ε.



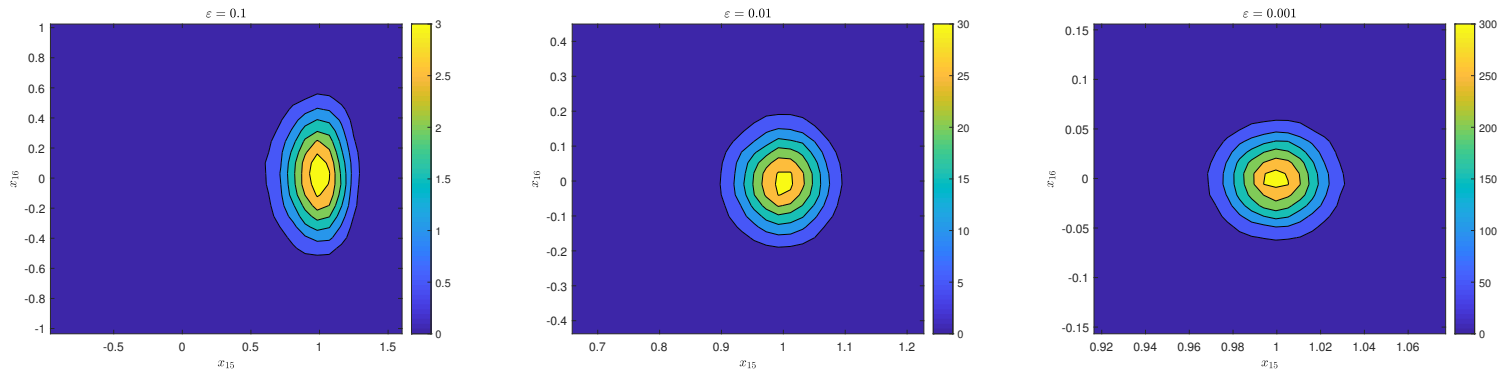
(a) Computational times



(b) Computed eigenvalues

A 16D result

- We note that the maximum density of the distribution is inversely proportional to ε, which indicates the singularity of distributions as $\varepsilon \to 0^+$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 8: At $T = 2048$ and with $\alpha \approx 0.2097$, the convergent empirical distribution of particles is projected onto the $x_{15}x_{16}$-plane.

Penalty functions: discrete Wasserstein distance

- Given distributions μ and ν defined on metric spaces X and Y, we aim to construct a transport map $f_*^0 : X \to Y$ such that $f_*^0(\mu) = \nu$.
- Given function $f : X \to Y$, the p-Wasserstein distance between $f_*(\mu)$ and ν is

$$W_p(f_*(\mu), \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} dist(f(x), y)^p \, \mathrm{d}\gamma(x, y) \right)^{1/p}, \tag{31}$$

  where $\Gamma(\mu, \nu)$ denotes the collection of all measures on $X \times Y$ with marginals μ and ν on the first and second factors respectively.
- In our work, we study the case with $p = 2$.

Empirical distributions and doubly stochastic matrix

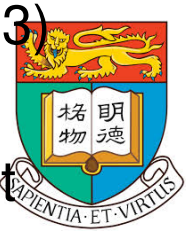
- In practice, the closed-form solution of μ and ν may be unknown, instead only N independent and identically distributed (i.i.d.) samples of μ and ν are available.
- We approximate the probability measures μ and ν by empirical distribution functions:

$$\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i} \quad \text{and} \quad \nu = \frac{1}{N}\sum_{j=1}^{N}\delta_{y_j}. \tag{32}$$

- Any element in $\Gamma(\mu,\nu)$ can clearly be represented by an $N\times N$ doubly stochastic matrix, denoted as transition matrix, $\gamma = (\gamma_{ij})_{i,j}$ satisfying:

$$\gamma_{ij} \geq 0; \qquad \forall j,\ \sum_{i=1}^{N}\gamma_{ij} = 1; \qquad \forall i,\ \sum_{j=1}^{N}\gamma_{ij} = 1. \tag{33}$$

- The empirical distribution functions allow us to approximate different measures.



Network Training Objective

- The DeepParticle method does not assume the knowledge of closed form distribution of μ and ν, instead we have i.i.d. samples of μ and ν namely, x_i and y_j, $i,j = 1, \cdots, N$, as training data.
- Then a discretization of (31) is:

$$\hat{W}(f) := \left(\inf_{\gamma \in \Gamma^N} \sum_{i,j=1}^{N} dist(f(x_i), y_j)^2 \gamma_{ij} \right)^{1/2}, \tag{34}$$

  where Γ^N denotes all $N \times N$ doubly stochastic matrices.
- Let the DNN map be $f_\theta(x;\eta)$, where x is the input, η is the shared physical parameter and θ denotes all the trainable parameters in the network.
- In case of $X = Y = \mathbb{R}^d$ equipped with Euclidean metric, the training loss function is

$$\hat{W}^2(f_\theta) := \sum_{r=1}^{n_\eta} \left(\inf_{\gamma_r \in \Gamma^N} \sum_{i,j=1}^{N} |f_\theta(x_{i,r};\kappa_r) - y_{j,r}|^2 \gamma_{ij,r} \right). \tag{35}$$

Iterative method in finding transition matrix γ

- To minimize the loss function (35), we update parameters θ of f_θ with the classical Adams stochastic gradient descent, and alternate with updating the transition matrix γ.
- The problem (35) is a linear program on the bounded convex set Γ^N of vector space of real $N \times N$ matrices. By Choquet's theorem, this problem admits solutions that are extremal points of Γ^N.
- Set of all doubly stochastic matrix Γ^N can be referred to as Birkhoff polytope. The Birkhoff–von Neumann theorem states that such polytope is the convex hull of all permutation matrices.
- So given N (typically, 2000) samples of distribution, the DOF of minimization is N^2 (4M).
- Hence, it is natural to seek a localized optimization method that decreases the loss function monotonously in each step and converges to some permutation matrix.

The mini-batch linear programming algorithm

- We present a mini-batch linear programming algorithm to find the best γ for each inner sum of (35).
- In each iteration, we select columns and rows and solve a sub-problem under the constraint that maintains column-wise and row-wise sums of the corresponding sub-matrix of γ.
- Let $\{i_k\}_{k=1}^M$, $\{j_l\}_{l=1}^M$ $(M \ll N)$ denote the index chosen from $\{1, 2, \cdots, N\}$ without replacement. The cost function of the sub-problem is

$$C(\gamma^*) := \sum_{k,l=1}^{M} |f_\theta(x_{i_k}) - y_{j_l}|^2 \gamma^*_{i_k j_l} \tag{36}$$

subject to

$$\begin{cases} \sum_{k=1}^M \gamma^*_{i_k,j_l} = \sum_{k=1}^M \gamma_{i_k,j_l} & \forall l = 1, \cdots, M \\ \sum_{l=1}^M \gamma^*_{i_k,j_l} = \sum_{l=1}^M \gamma_{i_k,j_l} & \forall k = 1, \cdots, M \\ \gamma^*_{i_k j_l} \geq 0 & \forall k, l = 1, \cdots, M, \end{cases} \tag{37}$$

where γ_{i_k,j_l} are from the previous step.

The mini-batch linear programming algorithm (cont.)

- Since optimization goal of γ and network f are Min-Min, we can solve alternately.
- The cost of finding optimal γ increases as N increases, however, the network itself is independent of γ.
- After training, our network acts as a sampler from some target distribution ν without assumption of closed-form distribution of ν.
- At this stage, the input data is no longer limited by training data, an arbitrarily large amount of samples approximately obeying ν can be generated through μ (uniform distribution).
- The DeepParticle method can be viewed as a generative model.

Par-Net: physical parameter-dependent network

- The network takes on two kinds of input, particle position **x** and physical parameter η.
- Fixing η, the network of **x** is a push-forward map, which should enjoy better regularity (stronger regularizer during training).

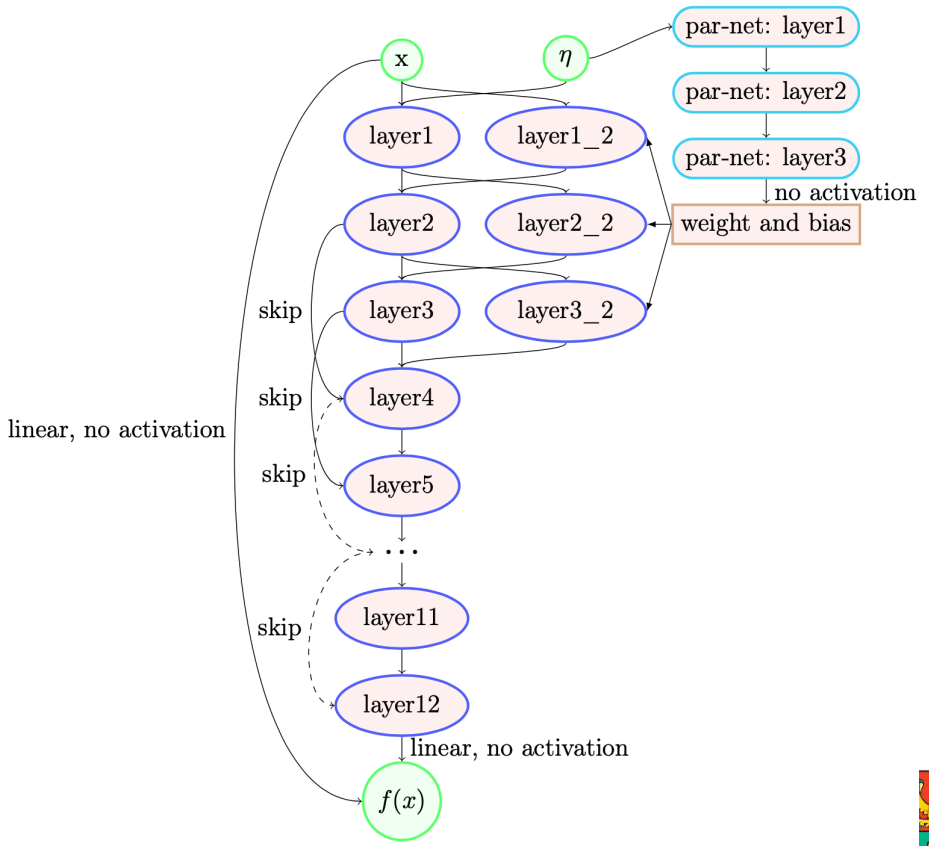


Figure 9: η is physical parameter input

Learning and generating aggregation patterns in Keller-Segel chemotaxis systems

- Keller-Segel equation models small organisms direct their movements towards the gradient of some chemical concentration.
- A common form of KS model in 2D space:

$$\rho_t = \nabla \cdot (\mu \nabla \rho - \chi \rho \nabla c), \quad \epsilon c_t = \Delta c + \rho, \tag{38}$$

  where χ, μ, and ϵ are positive constants.
- Physically, the chemical is faster than the organism, hence $\epsilon \to 0$.
- In addition, we assume an extra advection term **v** that indicates the fluid medium of the organism has its own current.
- Taking $\chi = \mu = 1$, we arrive at,

$$\rho_t = \Delta \rho + \nabla \cdot \big(\rho \nabla (\mathcal{K} * \rho)\big) - \nabla \cdot \big(\rho \mathbf{v}\big), \tag{39}$$

  where $\mathcal{K} = 1/(2\pi) \log |x|$ is the Green's function of Poisson equation..

Interacting Particle Method

- The density connects to McKean-Vlasov equation as $N \uparrow \infty$,

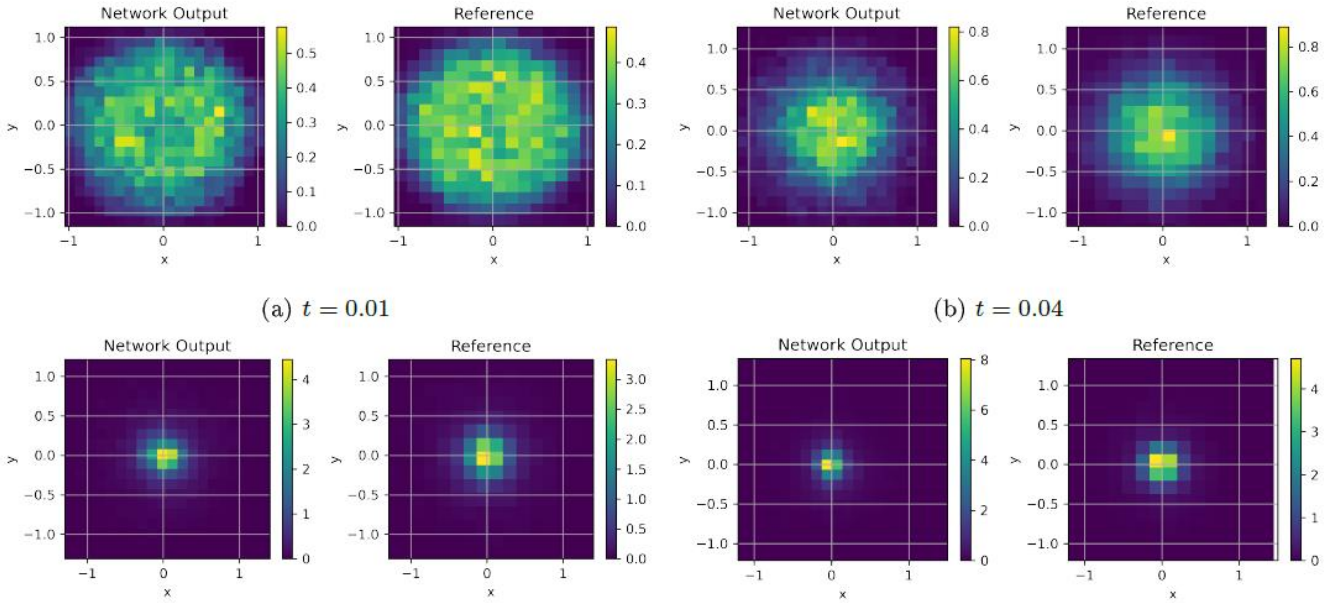
$$dX^j = -\frac{M}{N}\nabla_{X^j} \sum_{i=1:J, i\neq j} \mathcal{K}(|X^j - X^i|)dt + \mathbf{v}(X^j)dt, +\sqrt{2}dW^j$$

$$j = 1, 2, \cdots, N;$$

  where M is the conserved total mass (integral of ρ), W^j's are independent BM.
- Numerical difficulties/instabilities:
  - In each step, computational cost is $\mathcal{O}(N^2)$
  - $\mathcal{K}(X^j - X^i)$ is singular when particles are near.
- We use the DeepParticle method for learning and generating aggregation patterns in multi-dimensional Keller-Segel chemotaxis systems.

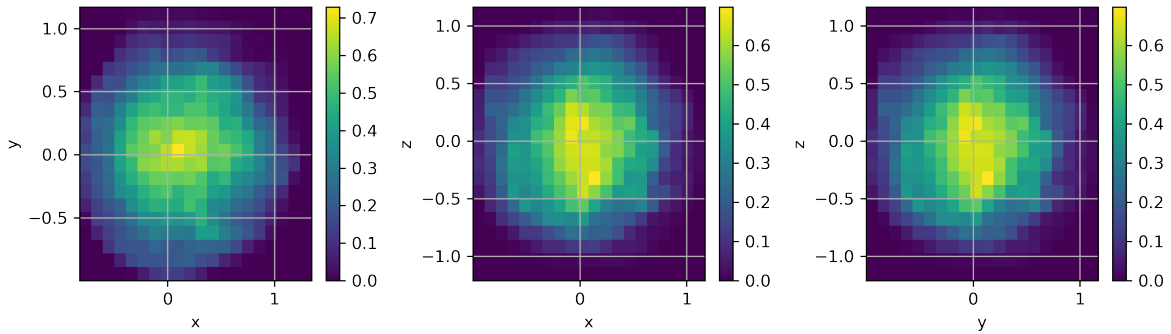
A 2D problem with $\mathbf{v} = 0$

- In case of no convection, blow up at $t = 0.125$ is predicted by second-order moment.
- The training data are snapshots of empirical distribution with different t within $[0, 0.1]$.
- It is generated by IPM with 1000 particles, and includes a regularization term in interaction.

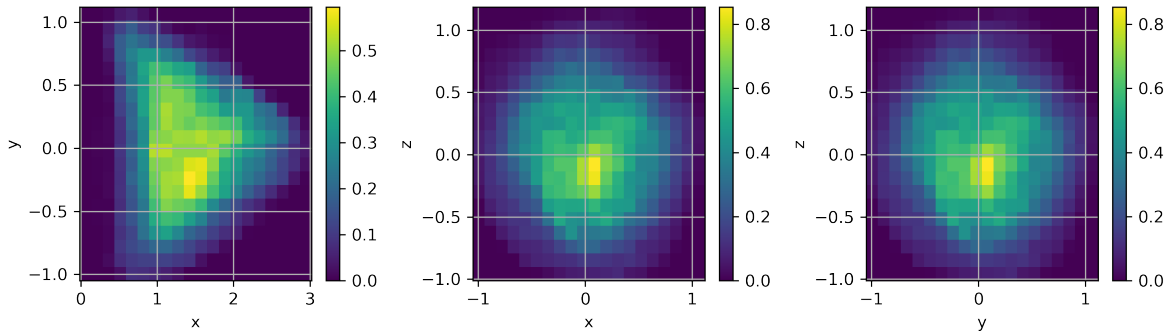


3D Laminar, with different A, fixed time

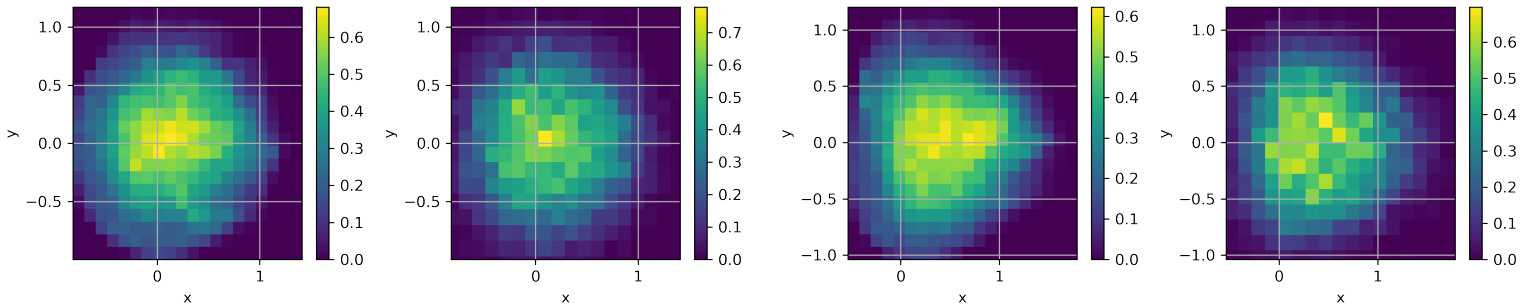
- Both IPM and DeepParticle generalize trivially to higher dimensions.
- Laminar flow: $\mathbf{v} = A(\exp(-y^2 - z^2), 0, 0)^T$



(a) $A = 10$

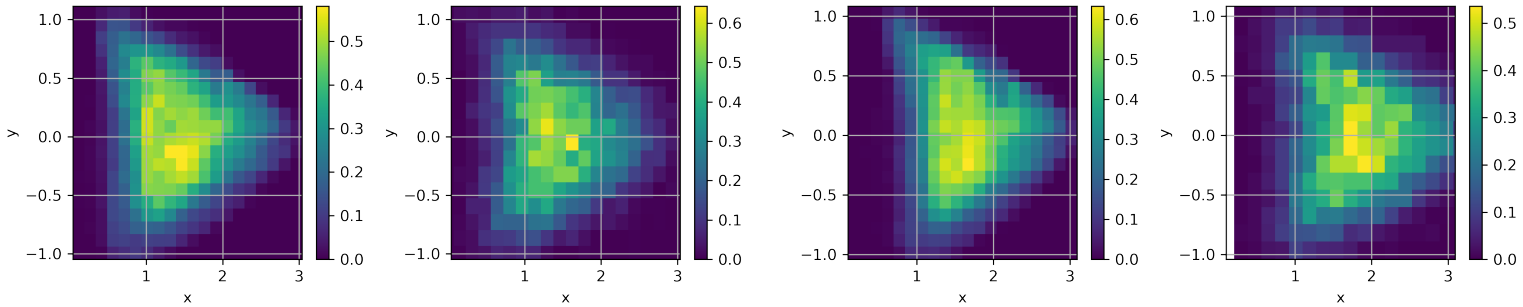


(b) $A = 100$



(a) $A = 10$

(b) $A = 30$



(c) $A = 100$

(d) $A = 130$ (extrapolation)

Figure 11: Comparison between reference and predicted density, projected to xy plane.

Conclusion

- We developed structure-preserving schemes to compute effective diffusivities in chaotic and random flows.
- We develop interacting particle methods (IPMs) to compute principal eigenvalues of non-self-adjoint elliptic operators, which can compute KPP front speeds of reaction-diffusion-advection equations and entropy production rates in diffusion processes.
- Our particle methods are mesh-free and self-adaptive.
- We developed DeepParticle methods to learn and generate distributions of solutions under variations of physical parameters.
- We present numerical results to demonstrate the accuracy and efficiency of the proposed methods.

Thank you, Prof. Hou

- I would like to take this opportunity to express my deepest gratitude for your exceptional guidance and mentorship.
- Your unwavering support and invaluable assistance throughout my academic journey have been truly remarkable. I very much appreciate it.
- I want to extend my heartfelt wishes for your good health and continued success in all your future endeavors.

Thank the support from RGC and NSFC!

Thank you very much for your attention!