Particle Methods for Nonlinear Diffusion Equations

Claire Murphy

UC Santa Barbara

April 7, 2025

Outline

Motivation

The Generalized Fokker-Planck Equation

Suppose that

- $V : \mathbb{R}^d \to \mathbb{R}$ is differentiable, λ-convex for some $\lambda > 0$ and $\min V(x) = 0$.
- $f : [0, \infty) \to \mathbb{R}$ is l.s.c., coercive, and convex, $f(0) = 0$.

[Craig, Jacobs, Turanova '24] showed that there exist weak solutions to the **Generalized Fokker-Planck equation**

$$\begin{cases} \partial_t \rho = \nabla \cdot (\nabla f^*(p) + \rho \nabla V) \\ p \in \partial f(\rho) \\ \rho(0, x) = \rho_0(x). \end{cases} \tag{PDE}$$

Notice that in the case $f(s) = s \ln(s)$, this equation simplifies to the Fokker-Planck equation.

Application 1: Difficult Diffusion Cases

For f sufficiently regular, PDE becomes $\partial_t \rho - \nabla \cdot (\rho \nabla V) = \Delta F(\rho)$, where $F'(s) = sf''(s)$.

This equation describes difficult cases of nonlinear diffusion, such as:

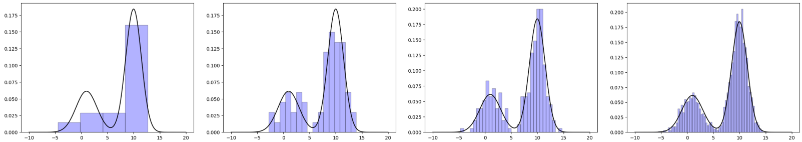
1. Fast diffusion, $f(s) = \frac{s^m}{m-1}$, $m < 1$.
2. Avalanche dynamics, $f(s) = \begin{cases} 0 & s \in [0, r_c] \\ s \log\left(\frac{s}{r_c}\right) & s > r_c. \end{cases}$
3. Height-constrained transport, formally modeled by $f(s) = \begin{cases} 0 & s \in [0, 1] \\ \infty & \text{otherwise.} \end{cases}$

Application 2: Sampling

Setup: Let $\tilde{\rho}$ be a probability measure on Euclidean space $\mathbb{R}^d$.

Goal: We seek $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ such that the empirical measure $\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ converges to $\tilde{\rho}$ as $n \to \infty$.

Our definition of “convergence" depends on the context of the problem. For example, we may define convergence in terms of the 2-Wasserstein metric.



Classical Sampling: Langevin Dynamics

Assumption: The target measure $\tilde{\rho}$ is strongly log-concave, i.e. $\tilde{\rho}(x) = e^{-V(x)}dx$ for a λ-convex, continuously differentiable function $V : \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$.

PDE perspective: If $\rho(t, x)$ is a solution to the **Fokker-Planck equation**, then

$$W_2(\rho(t), \tilde{\rho}) \leq e^{-\lambda t} W_2(\rho_0, \tilde{\rho}).$$

Particle method: Let $\{x_{i,0}\}_{i=1}^n \subseteq \mathbb{R}^d$ be iid samples from a measure with finite second moment. If we evolve our particles by the stochastic differential equation

$$\begin{cases} dx_i(t) = \nabla \log(\tilde{\rho}(x_i))dt + dW_i \\ x_i(0) = x_{i,0} \end{cases}$$

then $\lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i(t)}(x) = \tilde{\rho}(x)$.

Sampling and the Generalized Fokker-Planck equation

Assumption: The target measure $\tilde{\rho}$ is of the form

$$\tilde{\rho}(x) = \max((f')^{-1}(Z - V(x)), 0).$$

Z is a normalization constant chosen so that $\int \tilde{\rho} = 1$.

Remark

If $f(s) = s\ln(s)$, then $\tilde{\rho}$ is strongly log-concave.

PDE perspective: We formally expect that if $\rho(t,x)$ is a solution to the **Generalized Fokker-Planck equation**, then

$$W_2(\rho(t), \tilde{\rho}) \leq e^{-\lambda t} W_2(\rho_0, \tilde{\rho}).$$

Remark

If $f(s) = s\ln(s)$, the Generalized Fokker-Planck equation reduces to the Fokker-Planck equation.

Particle method: Our goal!

Our Goal

Our goal is to develop a new numerical particle method to solve the Generalized Fokker-Planck equation.

Why a particle method?

1. Intrinsically adaptive, i.e. better resolution in areas with higher density.
2. In principle, can be scaled to higher dimensions (unlike grid-based methods).
3. Directly connected to sampling application.

Our Method

Previous Work

Consider the function

$$f_\varepsilon(s) = \begin{cases} \frac{\delta(\varepsilon)}{2}|s|^2 + {}^{\delta(\varepsilon)}f(s) - {}^{\delta(\varepsilon)}f(0) & s \geq 0 \\ +\infty & s < 0, \end{cases}$$

where ${}^{\varepsilon}f$ is the Moreau-Yosida regularization of f and $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Let φ_ε be a mollifier.

We obtain the following nonlocal approximation of PDE:

$$\begin{cases} \partial_t \rho_\varepsilon = \nabla \cdot (\rho_\varepsilon \nabla (p_\varepsilon + V)) \\ p_\varepsilon = \varphi_\varepsilon * f'_\varepsilon(\varphi_\varepsilon * \rho_\varepsilon) \\ \rho_\varepsilon(0,x) = \rho^0_\varepsilon(x). \end{cases} \qquad (\text{PDE}_{\varepsilon,\text{moll}})$$

Theorem (Craig, Jacobs, Turanova '24)

Let $\rho_\varepsilon \in AC^2_{loc}([0,\infty]; \mathcal{P}_2(\mathbb{R}^d))$ be the solution of $PDE_{\varepsilon,moll}$ with initial data ρ^0_ε and velocity V. For some $\rho^0 \in P_2(\mathbb{R}^d)$, suppose that $\rho^0_\varepsilon \to \rho^0$ in W_2. We assume sufficient conditions on ρ^0_ε, ρ^0, f, V, δ, and φ_ε. Then, there exists $\rho \in AC^2_{loc}([0,\infty); \mathcal{P}_2(\mathbb{R}^d))$ so that, up to a subsequence, $\rho_\epsilon(t) \to \rho(t)$ in weak $L^1_{loc}(\mathbb{R}^d)$ and W_1 for all $t \geq 0$, and ρ is a solution of PDE with initial data ρ^0.

Our Contribution: A Particle Method

The double convolution structure of $\text{PDE}_{\varepsilon,\text{moll}}$ is important in convergence analysis, but can be dropped in the numerical scheme:

$$\begin{cases} \partial_t \rho_\varepsilon = \nabla \cdot (\rho_\varepsilon \nabla (p_\varepsilon + V)) \\ p_\varepsilon = f_\varepsilon'(\varphi_\varepsilon * \rho_\varepsilon) \\ \rho_\varepsilon(0, x) = \rho_\varepsilon^0(x). \end{cases} \tag{PDE_ε}$$

Suppose that $\rho_\varepsilon^0(x) = \sum_{i=1}^{n} m_i \delta_{x_i^0}(x)$. Then, there exists a unique solution of PDE_ε of the form $\rho_\varepsilon(t, x) = \sum_{i=1}^{n} m_i \delta_{x_i(t)}(x)$, where

$$\begin{cases} \dot{x}_j(t) = -\nabla p_\varepsilon(x_j(t)) - \nabla V(x_j(t)) \\ x_j(0) = x_j^0. \end{cases}$$

Elementary calculations show that the trajectory of the jth particle at time t is given by

$$\begin{cases} \dot{x}_j(t) = -f_\varepsilon'' \left(\sum_{i=1}^{n} m_i \varphi_\varepsilon(x_j(t) - x_i(t)) \right) \sum_{i=1}^{n} m_i \nabla \varphi_\varepsilon(x_j(t) - x_i(t)) - \nabla V(x_j(t)) \\ x_j(0) = x_j^0. \end{cases}$$

We investigate multiple (implicit and explicit) ways to numerically solve the ODE system.

1-Dimensional Numerical Setup

1. Assume that the support of $\rho_0(x)$ is contained in some interval $[a, b]$. Initialize equidistant particles $x_1^0 < x_2^0 < \ldots < x_n^0 \subseteq \mathbb{R}$ so that

$$[a, b] \subseteq [x_1^0, x_n^0].$$

2. Define m_i as the integral of $\rho_0(x)$ on the interval of length $h := x_2^0 - x_1^0$ centered at x_i.
3. Evolve particles according to the ODE system described above.
4. Our particle solution $\sum_{i=1}^{n} m_i \delta_{x_i(t)}(x)$ is impossible to visualize graphically, so we mollify by the Gaussian function

$$\xi_\varepsilon(x) = \frac{1}{(4\pi\varepsilon^2)^{1/2}} e^{-x^2/4\varepsilon^2}.$$

In other words, in all graphs appearing in this presentation, the smooth solution $\mu_\varepsilon(t, x)$ has the form

$$\mu_\varepsilon(t, x) = \xi_\varepsilon * \sum_{i=1}^{n} m_i \delta_{x_i(t)}(x) = \sum_{i=1}^{n} m_i \xi_\varepsilon(x - x_i(t)).$$

Parameters

1. h: Initial distance between adjacent particles.
2. ε_φ: Choice of ε used by the mollifier φ_ε in our ODE system.
3. ε_{MY}: Choice of ε used in the Moreau-Yosida approximation of f.
4. ε_ξ: Choice of ε used in the construction of our visualization mollifier ξ_ε. **Does not influence our underlying particle method.**

Performance

Nonlinear Diffusion

Consider the function

$$f(s) = \begin{cases} \frac{s^m}{m-1} & m \neq 1 \\ s\ln(s) & m = 1 \end{cases}$$

and the external potential $V(x) = \frac{x^2}{2(m+1)}$.

PDE becomes $\partial_t \rho + \nabla \cdot (\rho V) = \Delta \rho^m$.

1. $m > 1$: Porous medium equation.
2. $m = 1$: Heat equation.
3. $m < 1$: Fast diffusion equation.

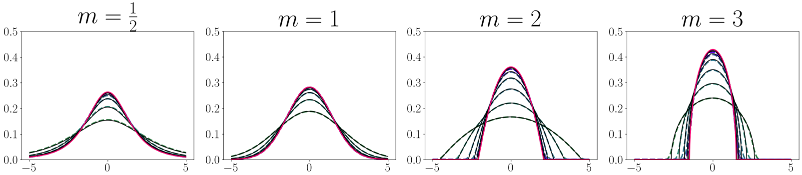


Figure: Numerical solutions are plotted with dashed lines, exact solutions are plotted with solid lines. Solutions are plotted in increments of one unit of time, up until $t_{max} = 6$. $h = .01$, $\varepsilon_{\varphi} = 5h^{.9}$, $\varepsilon_{MY} = 0$, $\varepsilon_{\xi} = 7h^{.9}$.

Evaluating Accuracy

For a given distance h between initial particles, define

$$\varepsilon_\varphi = a_\varphi h^{.9}.$$

Assessing accuracy:

1. For $a_\varphi \in \{1, 5, 10\}$, and for $h \in \{.01, .02, .04, .08\}$, compute the particle solution up to $t_{max} = 1.5$.
2. Compare the numerical and exact solutions using the 2-Wasserstein metric, which can be approximated numerically.

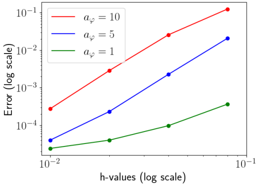


Figure: $m = 1$, $V(x) = \frac{x^2}{2(m+1)}$, $\varepsilon_{MY} = 0$.

	$h = .01$	.02	.04	.08
$a_\varphi = 10$	34.599	7.832	3.922	0.898
$a_\varphi = 5$	53.523	10.047	2.641	1.012
$a_\varphi = 1$	36255	1782	11.542	1.521

Table: Runtime, in seconds.

$m =$ slope of blue line $\approx 3.00181 \Rightarrow \ln(E(h)) \approx m \ln(h) + C_0 \Rightarrow E(h) \approx C_1 h^m$.

The Moreau-Yosida Regularization and Accuracy

For a given distance h between initial particles, define

$$\varepsilon_{MY} = a_{MY} h^{.9}.$$

Assessing accuracy:

1. For $a_{MY} \in \{0, .1, 1\}$, and for $h \in \{.01, .02, .04, .08\}$, compute the particle solution up to $t_{max} = 1.5$.
2. Compare the numerical and exact solutions using the 2-Wasserstein metric, which can be approximated numerically.

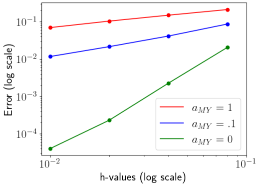


Figure: $m = 1$, $V(x) = \frac{x^2}{2(m+1)}$, $\varepsilon_\varphi = 5h^{.9}$.

	$h = .01$	.02	.04	.08
$a_{MY} = 1$	17.397	4.154	0.980	0.224
$a_{MY} = .1$	19.918	4.740	1.096	0.231
$a_{MY} = 0$	18.507	4.064	0.745	0.190

Table: Runtime, in seconds.

The Moreau-Yosida Regularization and Step Size

Avalanche dynamics: Particles above a critical height r_c diffuse according to the heat equation.

$$f(s) = \begin{cases} 0 & s \in [0, r_c] \\ s \log\left(\frac{s}{r_c}\right) & s > r_c. \end{cases}$$

We solve our ODE system using the backward Euler method with step size Δt.

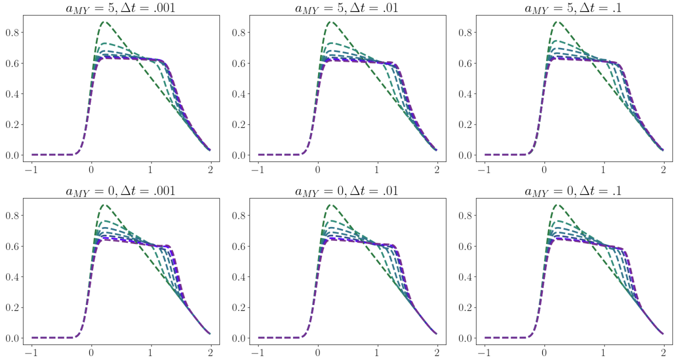


Figure: $h = .005$, $\varepsilon_\varphi = 5h^{.9}$, $\varepsilon_\xi = 9h^{.9}$.

The Visualization Mollifier

In all figures, we plot

$$\mu_\varepsilon(t,x) = \xi_\varepsilon * \sum_{i=1}^{n} m_i \delta_{x_i(t)}(x) = \sum_{i=1}^{n} m_i \xi_\varepsilon(x - x_i(t)).$$

where ξ_ε is a mollifier.

How does the strength of ε influence $\mu_\varepsilon(t,x)$?

Example

- $f(s) = s\ln(s)$.
- The double well $V(x) = (1-x)^2(1+x)^2$.

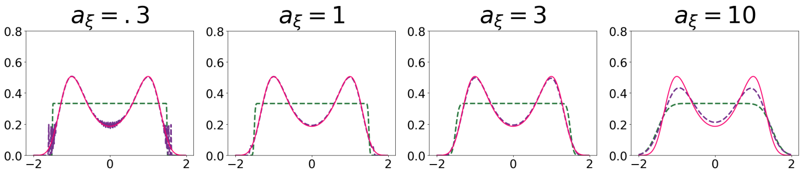


Figure: Initial solution and solution at $t_{max} = .45$. $h = .01$, $\varepsilon_\varphi = h^{.9}$, $\varepsilon_{MY} = 0$.

Compactly Supported Mollifers

In all other simulations, the mollifier in our ODE system is a Gaussian with standard deviation ε_φ. What happens if we use a mollifier with compact support?

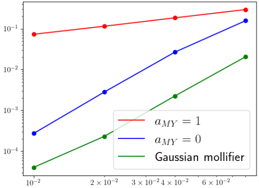


Figure: $m = 1$, $V(x) = \frac{x^2}{2(m+1)}$.

	$h = .01$	.02	.04	.08
$a_{MY} = 1$	11.909	3.045	0.755	0.247
$a_{MY} = 0$	27.614	2.667	0.631	0.195
Gaussian	18.370	3.551	0.646	0.156

Table: Runtime, in seconds.

The Random Batch Method

Example: Height Constraint

Height constraint: Our particle solution cannot rise above height 1.

$$f(s) = \begin{cases} 0 & s \in [0,1] \\ \infty & \text{otherwise.} \end{cases}$$

We approximate our energy via $f(s) = \frac{s^{100}}{99}$. $V(x) = \frac{x^2}{2}$. Up to time $\ln(3)$, particles are attracted to $x_0 = 0$. After the particle solution hits the “ceiling" 1, particles remain steady.

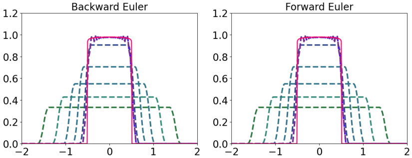


Figure: Solutions are plotted in increments of .25 units of time, up until $t_{max} = 1.5$. $h = .005$, $\varepsilon_\varphi = 2h^{.9}$, $\varepsilon_{MY} = 0$, $\varepsilon_\xi = 15h^{.9}$.

	Backward Euler	Forward Euler
Error	$1.461 * 10^{-4}$	$1.461 * 10^{-4}$
Runtime (s)	8351	3132

Previous Work: The Batch Method

The Batch Method: At each time step, randomly select p particles. Evolve ALL particles according to the approximate velocity field based on the location of the p particles [Jin, Li, and Liu, '20].

In practice, we lose accuracy and obtain a stiffer ODE system. Consider slow diffusion with $m = 2$ and no external potential.

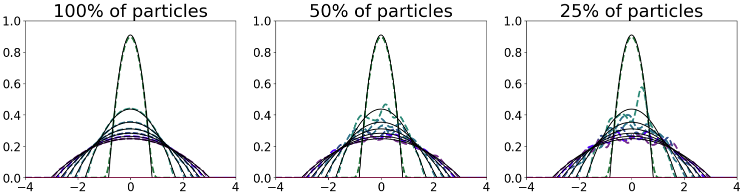


Figure: Numerical solutions are plotted with dashed lines, exact solutions are plotted with solid lines. Solutions are plotted in increments of .5 units of time, up until $t_{max} = 3$. $h = .01$, $\varepsilon_\varphi = 5h^{.9}$, $\varepsilon_{MY} = 0$, $\varepsilon_\xi = 15h^{.9}$.

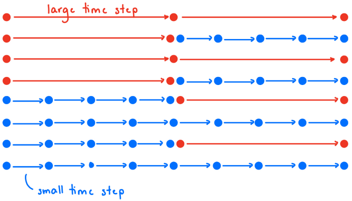
	100% of particles	50% of particles	25% of particles
Error at t_{max}	$3.078 * 10^{-4}$	$3.966 * 10^{-4}$	$8.995 * 10^{-4}$
Runtime (s)	13.12809	17.25817	19.03422

The Batch Forward Euler Method

Notice that we can successfully use explicit (forward Euler) methods with even our stiffest simulations!

Is there a way to retool the batch method to avoid adding stiffness to our system?

The Batch Forward Euler Method: At each time step, randomly select p particles. These p particles make a “fast" forward Euler jump with a large time step. All other particles evolve according to a forward Euler method with a slower / smaller time step.



Pseudocode

Input: fast_t_step, slow_t_step, p, N, T, f, x_0
$k \leftarrow$ fast_t_step / slow_t_step
for $i = 1$ to $T/$ slow_t_step **do**
    if i mod k = 0 **then**
        Let I_{fast} equal p randomly selected indices of $\{1, \ldots, N\}$
        $I_{\text{slow}} \leftarrow \{1, \ldots, N\} \setminus I_{\text{fast}}$
        $x_{i+1}[I_{\text{fast}}] \leftarrow x_i[I_{\text{fast}}] + \text{fast\_t\_step} \cdot f(i \cdot \text{slow\_t\_step}, x_i)[I_{\text{fast}}]$
    else
        $x_{i+1}[I_{\text{fast}}] = x_i[I_{\text{fast}}]$
    end if
    $x_{i+1}[I_{\text{slow}}] \leftarrow x_i[I_{\text{slow}}] + \text{slow\_t\_step} \cdot f(i \cdot \text{slow\_t\_step}, x_i)[I_{\text{slow}}]$
end for

Example: Height Constraint

$$f(s) = \begin{cases} 0 & s \in [0,1] \\ \infty & \text{otherwise} \end{cases} \quad \text{and} \quad V(x) = \frac{x^2}{2}.$$

We approximate our energy via $f(s) = \frac{s^{100}}{99}$.

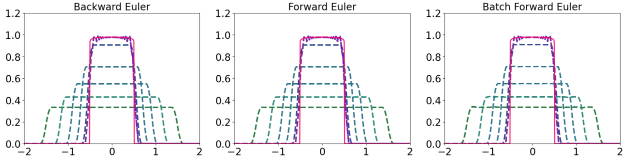


Figure: Solutions are plotted in increments of .25 units of time, up until $t_{max} = 1.5$. $h = .005$, $\varepsilon_\varphi = 2h^{.9}$, $\varepsilon_{MY} = 0$, $\varepsilon_\xi = 15h^{.9}$.

	Backward Euler	Forward Euler	Batch Forward Euler
Error	$1.461 * 10^{-4}$	$1.461 * 10^{-4}$	$1.461 * 10^{-4}$
Runtime (s)	8351	3132	886

Future Work

2D simulations: Currently programming a 2-dimensional simulation of the vorticity formulation of the Navier-Stokes equation.

Variational Inference: Let $C = \{\mu_1 \otimes \ldots \otimes \mu_d : \mu_i \in P(\mathbb{R})\}$. Approximate a probability measure $\pi \propto \exp(-V)$ over $\mathbb{R}^d$ via

$$\pi_C^* = \underset{\mu \in C}{\operatorname{argmin}} \, KL(\mu || \pi).$$

- [Jiang, Chewi, Pooladian '24] If π is λ-strongly log-concave, then $KL(\cdot || \pi)$ is λ-strongly convex over the geodesically convex set C.
- Avoid curse of dimensionality on measures on $\mathbb{R}^d$.

References

References

CEHT Craig, Katy, Karthik Elamvazhuthi, Matt Haberland, and Olga Turanova. “A blob method for inhomogeneous diffusion with applications to multi-agent control and sampling.” Mathematics of Computation (2023).

CJT Craig, Katy, Matt Jacobs, and Olga Turanova. “Nonlocal approximation of slow and fast diffusion.” Journal of Differential Equations (2025).

JCP Jiang, Yiheng, Sinho Chewi, and Aram-Alexandre Pooladian. “Algorithms for mean-field variational inference via polyhedral optimization in the Wasserstein space." (2024).

JLL Jin, Shi, Lei Li, and Jian-Guo Liu. “Random batch methods (RBM) for interacting particle systems.” Journal of Computational Physics (2020).