Variational formulation for degenerate and non-local PDEs

Hong Duong

University of Birmingham

Nonlinear Processes and their Applications, Université Jean Monnet, 02-05 July 2019

Wasserstein gradient flows

Jordan-Kinderlehrer-Otto, 1998: the Fokker-Planck equation

$$\partial_t \rho = \mathrm{div}(\nabla V \rho) + \Delta\rho, \quad \rho(0) = \rho_0,$$

is a gradient flow of the free energy

$$\mathcal{F}(\rho) = \int \rho \log \rho + V\rho$$

with respect to the Wasserstein distance on the probability measures with finite second moments

$$W_2^2(\mu,\nu) = \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \gamma(dxdy)$$

($\Gamma(\mu,\nu)$ is the set of couplings between μ and ν)

JKO-scheme

JKO-scheme: Given a time-step $h > 0$;

- $\rho_0^h := \rho_0$,
- for $n \geq 1$, ρ_n^h is defined to be the unique minimizer of

$$\mathcal{K}(\rho) := \frac{1}{2h} W_2^2(\rho, \rho_{n-1}^h) + \mathcal{F}(\rho).$$

Then the sequence $\{\rho_n^h\}_{n\geq 0}$, after an appropriate interpolation, converges to the solution to the Fokker-Planck equation as $h \to 0$.

Key ingredients in the JKO-proof

- Derive the Euler-Langrange equation (optimality conditions) for the minimizer: consider the flow

$$\partial_\tau \Phi_\tau(x) = \xi(\Phi_\tau(x)), \quad \Phi_0 = id$$

  then define $\rho_\tau := (\Phi_\tau)_\sharp \rho_h^k$ and compute $\frac{d}{d\tau}\mathcal{K}(\rho_\tau)|_{\tau=0} = 0$ to obtain the E-L equation.
- Establish time-discretization of the weak formulation for the Fokker-Planck equation (tested again a smooth function φ).
- To match two equations: choose $\xi = \nabla\varphi$.

The theory of Wasserstein gradient flows

The theory of Wasserstein gradient flow has developed tremendously over the last 20 years

- Many PDEs are Wasserstein gradient flows: porous medium equation, McKean-Vlasov equations, finite Markov chain, etc;
- Link different areas of Mathematics together (optimal transport, measure geometry & probability)

see e.g., monographs

- Ambrosio L., Gigli N., Savaré G (2008),
- Villani C. (2003 & 2009).

or the most recent survey by Santambrogio (2017).

Questions

- Can one establish JKO-type scheme for degenerate and nonlocal PDEs, e.g., the kinetic Fokker-Planck equation?
- Why should one minimize Wasserstein distance and entropy?
- Can one exploit variational formulation for multi-scale analysis?

Questions

- Can one establish JKO-type scheme for degenerate and nonlocal PDEs, e.g., the kinetic Fokker-Planck equation?
- Why should one minimize Wasserstein distance and entropy?
- Can one exploit variational formulation for multi-scale analysis?

Variational formulation for degenerate and nonlocal PDEs

Our aim is to construct JKO-type schemes for the following degenerate and non-local PDEs:

- the kinetic Fokker-Planck equation,
- a degenerate diffusion of Kolmogorov-type equation,
- the fractional kinetic Fokker-Planck equation.

Main difficulty: the Wasserstein metric does not work. Need to introduce new suitable optimal transportation cost functionals.

THE KINETIC FOKKER PLANCK EQUATION

Langevin dynamics and the kinetic Fokker-Planck equation

The kinetic Fokker-Planck equation

$$\partial_t \rho = -\operatorname{div}_q \left(\frac{p}{m}\rho\right) + \operatorname{div}_p \left(\nabla V \rho\right) + \gamma \operatorname{div}_p \left(\frac{p}{m}\rho\right) + \gamma\beta^{-1}\Delta_p \rho.$$

The Langevin dynamics

$$\begin{aligned} dQ &= \frac{P}{m}\,dt, \\ dP &= -\nabla V(Q)\,dt - \gamma\frac{P}{m}\,dt + \sqrt{2\gamma\beta^{-1}}\,dW_t. \end{aligned}$$

Neither a gradient flow nor a Hamiltonian flow, but a combination of them.

Variational formulation for the kinetic Fokker-Planck equation

New optimal transportation cost functional

$$C_h(q,p;q',p') := h\inf\Big\{\int_0^h \big|m\ddot{\xi}(t)+\nabla V(\xi(t))\big|^2\,dt : \xi\in C^1([0,h],\mathbb{R}^d)$$

such that

$$(\xi,m\dot{\xi})(0)=(q,p),\ (\xi,m\dot{\xi})(h)=(q',p')\Big\}.$$

Given $\mu(dqdp),\nu(dq'dp')$, define

$$\widetilde{W}_h(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)}\int_{\mathbb{R}^{2d}\times\mathbb{R}^{2d}} C_h(q,p;q',p')\gamma(dqdpdq'dp')$$

Explanation of the cost

Wasserstein cost

$$C(x,y) = |x-y|^2 = h \inf \Big\{ \int_0^h \big|\dot{\xi}(t)\big|^2 \, dt : \xi \in C^1([0,h],\mathbb{R}^d) \text{ such that } \xi(0)=x, \xi(h)=y \Big\}.$$

Freidlin-Wentzell small-noise large deviation

$$dX_t^\varepsilon = \sqrt{2\varepsilon} dW_t.$$

then $\{X^\varepsilon : X^\varepsilon(0)=x, X^\varepsilon(h)=y\}$ satisfies a LDP with rate function $C(x,y)$.

Our cost: small-noise perturbation of the Hamiltonian system

$$dQ = \frac{P}{m} \, dt$$
$$dP = -\nabla V(Q) \, dt + \sqrt{2\varepsilon} dW_t$$

Variational formulation for the kinetic Fokker-Planck equation

JKO-type scheme for the kinetic Fokker-Planck equation: $\rho_0^h := \rho_0$, for $n \geq 1$, ρ_n^h as the solution of the minimization problem

$$\min_{\rho} \frac{1}{2h}\frac{1}{\gamma}\widetilde{W}_h(\rho_{k-1}^h, \rho) + \mathcal{F}(\rho),$$

Theorem (D.-Peletier-Zimmer, M2AS 2014)

Under the piece-wise constant interpolation, the sequence $\{\rho_n^h\}_n$ converges, as $h \to 0$, to the solution of the kinetic Fokker-Planck equation.

Huang 2000 introduced a different scheme where the external force is not included in the cost functional.

A DEGENERATE DIFFUSION OF KOLMOGOROV-TYPE EQUATION

A degenerate diffusion of Kolmogorov-type equation

A degenerate diffusion of Kolmogorov-type equation:

$$\partial_t \rho(t, x_1, \ldots, x_n) = -\sum_{i=2}^{n} x_i \cdot \nabla_{x_{i-1}} \rho + \operatorname{div}_{x_n}(\nabla V(x_n)\rho) + \Delta_{x_n} \rho$$

the corresponding SDEs:

$$\begin{aligned} dX_1 &= X_2\, dt, \\ dX_2 &= X_3\, dt, \\ &\vdots \\ dX_{n-1} &= X_n\, dt, \\ dX_n &= -\nabla V(X_n)\, dt + \sqrt{2}\, dW_t. \end{aligned}$$

This is a (very) simple example of a chain of differential equations studied by Eckmann-Hairer 2000 and Delarue-Menozzi 2010

A degenerate diffusion of Kolmogorov-type equation

A rich history ($V = 0$):

- $n = 2$: Kolmogorov 1934,
- Hörmander 1967: hypoelliptic PDEs,
- many papers by Polidoro, Pascucci and co-authors (fundamental sol., Harnack inequality, Schauder estimates, etc),
- recent interests: Chen-Zhang 2019 (propagation of regularity).

Optimal transport cost functional

Let $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^{dn}, \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^{dn}$. We define the following cost between $\mathbf{x}$ and $\mathbf{y}$:

$$\mathcal{C}_t(\mathbf{x}, \mathbf{y}) := t \inf_{\xi} \int_0^t |\xi^{(n)}(s)|^2 \, ds,$$

where the infimum is taken over all curves $\xi \in C^n([0, T], \mathbb{R}^d)$ that satisfy the boundary conditions

$$\begin{aligned} &(\xi, \dot{\xi}, \ldots, \xi^{(n-1)})(0) = (x_1, x_2, \ldots, x_n) \quad \text{and} \\ &(\xi, \dot{\xi}, \ldots, \xi^{(n-1)})(t) = (y_1, y_2, \ldots, y_n). \end{aligned}$$

NB: $\mathcal{C}_t(\mathbf{x}, \mathbf{y})$ is called the mean squared derivative cost function and has many applications in motor control, biometrics and online-signatures and robotics, etc.

Optimal transport cost functional

Let $h > 0$ be given and $\mathcal{C}_h(\mathbf{x}, \mathbf{y})$ be the mean square derivative cost function. Let μ and ν be in $\mathcal{P}_2(\mathbb{R}^{dn})$. The Monge-Kantorovich optimal transport cost $\mathcal{W}_h(\mu, \nu)$ between μ and ν is defined by

$$\mathcal{W}_h(\mu, \nu)^2 = \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^{dn} \times \mathbb{R}^{dn}} \mathcal{C}_h(\mathbf{x}, \mathbf{y})\, \gamma(d\mathbf{x} d\mathbf{y}).$$

Approximation scheme: let $\rho_0^h := \rho_0$. For $k \geq 1$, define ρ_k^h as the solution of the minimization problem

$$\min_{\rho \in \mathcal{P}_2(\mathbb{R}^{dn})} \frac{1}{2h} \mathcal{W}_h(\rho_{k-1}^h, \rho) + \int_{\mathbb{R}^{dn}} \big(V(x_n) + \log \rho\big) \rho\, d\mathbf{x}.$$

Convergence of the approximation scheme

Theorem (D.-Tran, DCDS-A 2018)

Under the piece-wise constant interpolation, the sequence $\{\rho_n^h\}_n$ converges, as $h \to 0$, to the solution of the degenerate Kolmogorov equation

Key ingredients in the proofs

Follows JKO procedure:

- Derive the Euler-Langrange equation (optimality conditions) for the minimizer by perturbing the optimizer under a flow

$$\partial_\tau \Phi_\tau(x) = \xi(\Phi_\tau(x)), \quad \Phi_0 = id.$$

- Establish time-discretization of the weak formulation (tested again a smooth function φ).
- To match two equations: choose ξ appropriately (in terms of φ),
- Estimates of the new cost functions via the Wasserstein distance.

THE FRACTIONAL KINETIC FOKKER PLANCK EQUATION

The fractional kinetic Fokker-Planck equation

The fractional kinetic Fokker-Planck equation:

$$\partial_t \rho + v \cdot \nabla_x \rho = \mathrm{div}_v(\nabla \Psi(v)\rho) - (-\Delta_v)^s \rho,$$

Here $-(-\Delta_v)^s$ is the fractional Laplacian operator on the variable v, where the fractional Laplacian $-(-\Delta)^s$, is defined by

$$\begin{aligned} -(-\Delta)^s f(x) &:= -\mathcal{F}^{-1}(|\xi|^{2s}\mathcal{F}[f](\xi))(x) \\ &= -C_{d,s} \int_{\mathbb{R}^d} \frac{f(x) - f(y)}{|x-y|^{d+2s}} dy. \end{aligned}$$

Gradient flow formulation of the fractional heat equation

Erbar (2014) proved that the fractional heat equation

$$\partial_t \rho = -(-\Delta)^s \rho$$

is a gradient flow of the entropy w.r.t. a new metric defined via a non-local variant of the dynamical characterization of the Wasserstein distance by Benamou and Brenier.

Open problem: Is the fractional Fokker Planck equation

$$\partial_t \rho = -(-\Delta)^s \rho + \mathrm{div}(\nabla \Psi \rho)$$

a gradient flow of the free energy w.r.t. some distance?

The distance introduced by Erbar only works for the entropy!

Operator splitting scheme for the fractional Fokker Planck equation

Instead, Agueh-Bowles (2015) developed a splitting scheme for the fractional Fokker Planck equation:

1. transport equation, $\partial_t \rho = \mathrm{div}(\nabla \Psi \rho)$, as a Wasserstein gradient flow of the potential energy $\int \Psi \rho$ (Kinderlehrer-Tudorascu 2006),
2. fractional heat equation, $\partial_t \rho = -(-\Delta)^s \rho$, exactly solvable by convolution with the fractional heat kernel.

Question: can we develop an operator splitting scheme for the fractional kinetic Fokker Planck equation?

Operator splitting scheme for the FKFPE

1. Kinetic transport phase, $\partial_t \rho + v \cdot \nabla_x \rho = \mathrm{div}_v(\nabla\Psi(v)\rho)$, using a JKO-type variational formulation.
2. Fractional diffusion phase, $\partial_t \rho = -(-\Delta_v)^s \rho$, exactly solvable by convolution with the fractional heat kernel (in v-variable only).

Variational formulation for the kinetic transport equation

The kinetic transport equation

$$\partial_t \rho + v \cdot \nabla_x \rho = \mathrm{div}_v(\nabla\Psi(v)\rho)$$

The minimum acceleration cost functional: given $(x, v), (x', v') \in \mathbb{R}^{2d}$

$$C_h(x, v; x', v') = h \min_{\xi} \int_0^h |\ddot{\xi}(t)|^2 \, dt$$

where the minimum is taken over all curves $\xi \in C^2([0, h], \mathbb{R}^d)$ such that

$$(\xi, \dot{\xi})(0) = (x, v), \ (\xi, \dot{\xi})(h) = (x', v').$$

This cost has been studied by Huang 200, Gangbo-Westdickenberg 2009, Westdickenberg 2010, and Cavalletti-Sedjro-Westdickenberg 2019 for other PDEs.

Variational formulation for the kinetic transport equation

Explicit expression:

$$C_h(x,v;x',v') = |v'-v|^2 + 12\left|\frac{x'-x}{h} - \frac{v'+v}{2}\right|^2.$$

Given $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^{2d})$, define

$$\mathcal{W}_h(\mu,\nu)^2 = \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{R^{4d}} C_h(x,v;x',v')\gamma(dxdvdx'dv'),$$

then the kinetic transport equation can be approximated using the JKO-type scheme: ρ_h^k minimizes

$$\frac{1}{2h} W_h(\rho_h^{k-1},\rho)^2 + \int \Psi\rho.$$

NB: note that there is no entropy term which has a super-linear growth making the analysis harder. This generalizes Kinderlehrer-Tudorascu's result to the kinetic case.

The fractional diffusion equation

The fractional diffusion equation

$$\partial_t \rho = -(-\Delta_v \rho)^s \rho, \quad \rho(0, x, v) = \rho_0(x, v)$$

We solve this equation exactly

$$f(x, v, t) = \Phi_s(\cdot, t) *_v f_0(x, v)$$

where $*_v$ is the convolution operator in v-variable, where Φ_s is the fractional heat kernel

$$\Phi_s(v, t) := \mathcal{F}^{-1}(e^{-t|\cdot|^{2s}})(v).$$

Technical difficulty: infinite second moment

$$\int |v|^2 \Phi_s(v,t)\, dv = \infty \quad \forall s \in (0,1), t > 0.$$

Need to renormalize the convolution by introducing

$$\Phi^h_{s,R}(v) := \Phi^v_s(h) \mathbf{1}_{B_R}(v), \quad \Phi^h_s(v) := \Phi_s(v,h).$$

Operator splitting scheme

With an initial condition $f_h^0 = f_0$, for $n = 1, \cdots, N$ we iteratively compute the following:

- Given a trunction parameter $R > 0$, compute the renormalised convolution
$$\bar{f}_{h,R}^n := \frac{\Phi_{s,R}^h *_v f_{h,R}^{n-1}}{\|\Phi_{s,R}^h\|_{L^1(\mathbb{R}^d)}}.$$
- Solve for the minimizer $f_{h,R}^n$ of the problem

$$f_{h,R}^n := \operatorname{argmin}_{f \in \mathcal{P}_a^2(\mathbb{R}^d)} \Big\{ \frac{1}{2h} \mathcal{W}_h(\bar{f}_{h,R}^n, f)^2 + \int_{\mathbb{R}^{2d}} \Psi(v) f(x,v)\, dx dv \Big\}.$$

Time-interpolation: We define $f_{h,R}$ by setting

$$f_{h,R}(t) := \Phi_s(t - t_n) *_v f_{h,R}^n \text{ for } t \in [t_n, t_{n+1}).$$

Convergence of the scheme

Theorem (D.-Lu, DCDS-A 2019)

The time-interpolation process converges, as $h \downarrow 0$ and $R = h^{-1/2}$, converges to a weak solution of the FKFPE.

Only existence, no uniqueness! (difficulty: lack of product rule for the fractional operator)

Summary

We have developed variational formulation, by introducing new optimal transportation cost functionals, for some degenerate and non-local PDEs

- the kinetic Fokker-Planck equation,
- a degenerate diffusion of Kolmogorov-type equation,
- the fractional kinetic Fokker-Planck equation.

Future work: extensions to other degenerate and non-local PDEs, develop a unified framework.

Acknowledgment

This talk is based on collaborative works over the last few years with: M. Peletier (TU Eindhoven), J. Zimmer (Bath), H. Tran (Esmart, Norway), Y. Lu (Duke).

THANK YOU FOR YOUR ATTENTION!