EVI flows via Nonnegative Cross-Curvature

Gabriele Todeschi

LIGM, Université Gustave Eiffel

in collaboration with Flavien Léger and François-Xavier Vialard

FVOT Workshop, Orsay, 20/11/2024

Wasserstein distance

$\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ probability measures with finite second moment

Pushforward measure: $\mu = T_\# \nu \iff \mu(A) = \nu(T^{-1}(A))$

$\Gamma(\mu, \nu) = \{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : (\mathrm{p}_1)_\# \gamma = \mu,\ (\mathrm{p}_2)_\# \gamma = \nu\}$

Quadratic optimal transport problem: for $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ solve

$$W_2^2(\mu, \nu) := \min_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, d\gamma(x, y) \qquad (\mathcal{W})$$

T such that $\mu = T_\# \nu$ is an optimal transport map if

$$W_2^2(\mu, \nu) = \int_{\mathbb{R}^d} |T(y) - y|^2 \, d\nu(y)$$

$W_2 : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}_{\geq 0}$ is a distance and $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a geodesic space

Wasserstein gradient flows

Euclidean gradient flow: $\dot{x}(t) = -\nabla F(x(t))$ for $F : \mathbb{R}^d \to \mathbb{R}$

Wasserstein gradient flow: for $\mathcal{E} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, a curve $\rho : [0, T] \to \mathcal{P}_2(\mathbb{R}^d)$ solution to

$$\partial_t \rho - \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{E}}{\delta \rho}(\rho)\right) = 0 \tag{1}$$

Example: the linear/nonlinear Fokker-Planck equation

One needs to couple solutions to (1) with a "dissipation principle"

- Generalized Minimizing Movement (GMM)
- Energy Dissipation Equality (EDE)
- Evolution Variational Inequality (EVI)
- Characterization of Wasserstein subdifferential

Generalized Minimizing Movement (GMM)

JKO scheme[1]: compute ρ_n recursively as

$$\rho_n \in \underset{\rho \in \mathcal{P}_2(\mathbb{R}^d)}{\text{argmin}} \frac{1}{2\tau} W_2^2(\rho, \rho_{n-1}) + \mathcal{E}(\rho)$$

Generalization of Implicit Euler Scheme: find x_n iteratively as

$$x_n \in \text{argmin} \frac{1}{2\tau}|x - x_{n-1}|^2 + F(x) \longrightarrow \frac{x_n - x_{n-1}}{\tau} = -\tau \nabla F(x_n)$$

At each step

$$\frac{1}{2\tau} W_2^2(\rho_n, \rho_{n-1}) + \mathcal{E}(\rho_n) \leq \mathcal{E}(\rho_{n-1})$$

GMM[2]: find (weak) limits of this discrete process for $\tau \to 0$

- connection with PDE by showing this is a solution
- general approach: one can replace W_2^2 with any distance squared
- "not quantitative": how much is the energy decreasing?

[1]After Jordan, Kinderlehrer, Otto (1998), [2]De Giorgi

Evolution Variational Inequality (EVI)

EVI flow[1]: for $\lambda \in \mathbb{R}$, a curve $\rho : [0, T] \to \mathcal{P}_2(\mathbb{R}^d)$ such that for (almost every) $t \in (0, T)$

$$\frac{1}{2}\frac{d}{dt}W_2^2(\sigma, \rho(t)) \leq \mathcal{E}(\sigma) - \mathcal{E}(\rho(t)) - \frac{\lambda}{2}d^2(\sigma, \rho(t)), \quad \forall \sigma \in \mathcal{P}_2(\mathbb{R}^d)$$

Generalization of the Euclidean setting: if $F : \mathbb{R}^d \to \mathbb{R}$ λ-convex then

$$\dot{x}(t) = -\nabla F(x(t)) \iff \frac{1}{2}\frac{d}{dt}|x(t) - z|^2 \leq F(z) - F(x(t)) - \frac{\lambda}{2}|x(t) - z|^2, \ \forall z \in \mathbb{R}^d$$

Again, one can replace W_2^2 with any distance squared

Favorable properties:

- uniqueness of solutions
- stability
- exponential decay to the equilibrium ($\lambda > 0$)

Example: the linear/nonlinear Fokker-Planck equation is an EVI flow in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

[1] Ambrosio, Gigli, Savaré (2005)

Outline

- Ambrosio, Gigli, Savaré *original* construction of EVI flows
- EVI flows via Nonnegative Cross-Curvature (NNCC)
- The LJKO scheme

Outline

- Ambrosio, Gigli, Savaré *original* construction of EVI flows
- EVI flows via Nonnegative Cross-Curvature (NNCC)
- The LJKO scheme

Evolution Variational Inequality (EVI)

Let (X, d) be a geodesic space, $f : X \to \mathbb{R}$

EVI flow[1]: for $\lambda \in \mathbb{R}$, a curve $x : [0, T] \to X$ such that for (almost every) $t \in (0, T)$

$$\frac{1}{2}\frac{d}{dt}d^2(z, x(t)) \leq f(z) - f(x(t)) - \frac{\lambda}{2}d^2(z, x(t)), \quad \forall z \in X$$

Two key ingredients to prove existence of EVI flows:

- JKO scheme in metric setting

$$x_n \in \operatorname{argmin} \frac{1}{2\tau}d^2(x, x_{n-1}) + f(x) \tag{2}$$

- Convexity

The first one is not necessary

On the other hand, some convexity is needed[2]

[1] Ambrosio, Gigli, Savaré (2005), [2] Daneri, Savaré (2008)

Discrete EVI

Suppose x_n exists, let $z \in X$ and $\gamma : [0,1] \to X$ geodesic, $\gamma(0) = x_n$ and $\gamma(1) = z$

Suppose $\psi(x) := \frac{1}{2\tau} d^2(x, x_{n-1}) + f(x)$ is geodesically $(\frac{1}{\tau} + \lambda)$-convex

$$\psi(\gamma(s)) \le (1-s)\psi(x_n) + s\psi(z) - \left(\frac{1}{\tau} + \lambda\right)\frac{s(1-s)}{2} d^2(z, x_n)$$

Discrete EVI

Suppose x_n exists, let $z \in X$ and $\gamma : [0,1] \to X$ geodesic, $\gamma(0) = x_n$ and $\gamma(1) = z$

Suppose $\psi(x) := \frac{1}{2\tau} d^2(x, x_{n-1}) + f(x)$ is geodesically $(\frac{1}{\tau} + \lambda)$-convex

$$\psi(\gamma(s)) \leq (1-s)\psi(x_n) + s\psi(z) - \left(\frac{1}{\tau} + \lambda\right)\frac{s(1-s)}{2} d^2(z, x_n)$$

Since x_n minimizes ψ and using convexity

$$0 \leq \psi(\gamma(s)) - \psi(x_n) \leq s\Big(\psi(z) - \psi(x_n)\Big) - \left(\frac{1}{\tau} + \lambda\right)\frac{s(1-s)}{2} d^2(z, x_n)$$

Discrete EVI

Suppose x_n exists, let $z \in X$ and $\gamma : [0,1] \to X$ geodesic, $\gamma(0) = x_n$ and $\gamma(1) = z$

Suppose $\psi(x) := \frac{1}{2\tau} d^2(x, x_{n-1}) + f(x)$ is geodesically $(\frac{1}{\tau} + \lambda)$-convex

$$\psi(\gamma(s)) \leq (1-s)\psi(x_n) + s\psi(z) - \left(\frac{1}{\tau} + \lambda\right)\frac{s(1-s)}{2} d^2(z, x_n)$$

Since x_n minimizes ψ and using convexity

$$0 \leq \psi(\gamma(s)) - \psi(x_n) \leq s\Big(\psi(z) - \psi(x_n)\Big) - \left(\frac{1}{\tau} + \lambda\right)\frac{s(1-s)}{2} d^2(z, x_n)$$

Dividing by s and passing to the limit $s \to 0$

$$0 \leq \frac{1}{2\tau} d^2(z, x_{n-1}) + f(z) - f(x_n) - \frac{1}{2\tau} d^2(x_n, x_{n-1}) - \left(\frac{1}{\tau} + \lambda\right)\frac{1}{2} d^2(z, x_n)$$

Discrete EVI

Discrete EVI inequality: $\forall z \in X$

$$\frac{1}{2\tau}d^2(z,x_n) - \frac{1}{2\tau}d^2(z,x_{n-1}) \le f(z) - f(x_n) - \frac{\lambda}{2}d^2(z,x_n) - \frac{1}{2\tau}d^2(x_n,x_{n-1})$$

Fine characterization of the minimizer x_n ($\frac{1}{\tau} + \lambda \ge 0$)

Discrete EVI

Discrete EVI inequality: $\forall z \in X$

$$\frac{1}{2\tau} d^2(z, x_n) - \frac{1}{2\tau} d^2(z, x_{n-1}) \leq f(z) - f(x_n) - \frac{\lambda}{2} d^2(z, x_n) - \frac{1}{2\tau} d^2(x_n, x_{n-1})$$

Fine characterization of the minimizer x_n ($\frac{1}{\tau} + \lambda \geq 0$)

Note that through the Discrete EVI one has

- sublinear ($\lambda = 0$) / linear ($\lambda > 0$) convergence to the global minimizer of f
- uniqueness of solution ($\frac{1}{\tau} + \lambda \geq 0$)

Passing to the limit one obtains (formally) the continuous EVI

$$\frac{1}{\tau} d^2(z, x_n) - \frac{1}{\tau} d^2(z, x_{n-1}) \longrightarrow \frac{d}{dt} d^2(z, x(t))$$

Example: NPC spaces

Nonpositively Curved (NPC) space: for any $\gamma : [0,1] \to X$ geodesic, $\forall y \in X$

$$d^2(\gamma(t), y) \leq (1-s)d^2(\gamma(0), y) + sd^2(\gamma(1), y) - s(1-s)d^2(\gamma(0), \gamma(1)) \quad (3)$$

The distance function is 1-convex along geodesics

If f is geodesically λ-convex, we have for any $z \in X$,

$$\psi(x) := \frac{1}{2\tau} d^2(x, y) + f(x) \text{ is geodesically } \left(\frac{1}{\tau} + \lambda\right)\text{-convex, } \forall y \in X$$

$\longrightarrow$ the discrete EVI follows taking $y = x_{n-1}$

Example: NPC spaces

Nonpositively Curved (NPC) space: for any $\gamma : [0,1] \to X$ geodesic, $\forall y \in X$

$$d^2(\gamma(t), y) \leq (1-s)d^2(\gamma(0), y) + sd^2(\gamma(1), y) - s(1-s)d^2(\gamma(0), \gamma(1)) \quad (3)$$

The distance function is 1-convex along geodesics

If f is geodesically λ-convex, we have for any $z \in X$,

$$\psi(x) := \frac{1}{2\tau} d^2(x, y) + f(x) \text{ is geodesically } \left(\frac{1}{\tau} + \lambda\right)\text{-convex, } \forall y \in X$$

$\longrightarrow$ the discrete EVI follows taking $y = x_{n-1}$

- for Hilbert spaces equality holds in (3)
- Positively Curved (PC) space: the inequality (3) holds with reverse sign
- if (X, d) is PC $\longrightarrow$ ψ is not geodesically convex

Unfortunately, the Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is PC!

Assumption 4.0.1[1]: $\forall x_0, x_1, y \in X$, $\exists \omega : [0,1] \to X$, $\omega(0) = x_0$, $\omega(1) = x_1$ and

$$\psi(\omega(s)) \leq (1-s)\psi(\omega(0)) + s\psi(\omega(1)) - \frac{1}{2}(\frac{1}{\tau} + \lambda)s(1-s)d^2(\omega(0), \omega(1)) \quad (4)$$

$\longrightarrow$ there exists a curve along which ϕ is "$\left(\frac{1}{\tau} + \lambda\right)$-convex"

Equivalently, $s \mapsto \psi(\omega(s)) - \frac{s^2}{2}(\frac{1}{\tau} + \lambda)d^2(\omega(0), \omega(1))$ is convex

Then discrete EVI follows taking $y = x_{n-1}$

[1] Ambrosio, Gigli, Savaré (2005)

AGS hypothesis

Assumption 4.0.1[1]: $\forall x_0, x_1, y \in X$, $\exists \omega : [0,1] \to X$, $\omega(0) = x_0$, $\omega(1) = x_1$ and

$$\psi(\omega(s)) \le (1-s)\psi(\omega(0)) + s\psi(\omega(1)) - \frac{1}{2}\left(\frac{1}{\tau} + \lambda\right)s(1-s)d^2(\omega(0),\omega(1)) \quad (4)$$

$\longrightarrow$ there exists a curve along which ϕ is "$\left(\frac{1}{\tau} + \lambda\right)$-convex"

Equivalently, $s \mapsto \psi(\omega(s)) - \frac{s^2}{2}\left(\frac{1}{\tau} + \lambda\right)d^2(\omega(0),\omega(1))$ is convex

Then discrete EVI follows taking $y = x_{n-1}$

Condition (4) can be split naturally into: $\forall y \in X$, $\exists \omega$ such that

$d^2(\omega(s), y) - \frac{s^2}{2\tau} d^2(\omega(0),\omega(1))$ is convex $\longrightarrow$ "structural" condition on (X,d)

$f(\omega(s)) - \frac{s^2}{2}\lambda d^2(\omega(0),\omega(1))$ is convex $\longrightarrow$ condition on f

Additional hypothesis on the structure of (X,d) are needed...

[1] Ambrosio, Gigli, Savaré (2005)

Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

For $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ Ambrosio, Gigli and Savaré introduced the generalized geodesics

Let $\mu_0, \mu_1, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and assume (for simplicity) $\exists T_0$, T_1 optimal transport maps, i.e.

$$W_2^2(\mu_0, \nu) = \int_\Omega |T_0(y) - y|^2 \nu(y)\,, \quad W_2^2(\mu_1, \nu) = \int_\Omega |T_1(y) - y|^2 \nu(y)$$

Generalized geodesic: $\omega : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ defined as

$$\omega(s) = ((1-s)T_0 + sT_1)_{\#}\nu\,, \quad s \in [0,1]$$

Generalization of a geodesic:

$$\nu = \mu_0 \text{ and } T_0 = \text{Id} \implies \omega \text{ geodesic}$$

Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

For $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ Ambrosio, Gigli and Savaré introduced the generalized geodesics

Let $\mu_0, \mu_1, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and assume (for simplicity) $\exists T_0$, T_1 optimal transport maps, i.e.

$$W_2^2(\mu_0, \nu) = \int_\Omega |T_0(y) - y|^2 \nu(y)\,, \quad W_2^2(\mu_1, \nu) = \int_\Omega |T_1(y) - y|^2 \nu(y)$$

Generalized geodesic: $\omega : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ defined as

$$\omega(s) = ((1-s)T_0 + sT_1)_{\#}\nu\,, \quad s \in [0,1]$$

Generalization of a geodesic:

$$\nu = \mu_0 \text{ and } T_0 = \mathrm{Id} \implies \omega \text{ geodesic}$$

Then $W_2^2(\omega(s), \nu) - \frac{s^2}{2\tau} W_2^2(\omega(0), \omega(1))$ is convex

If $s \mapsto f(\omega(s)) - \frac{s^2}{2}\lambda d^2(\omega(0), \omega(1))$ is convex, take $\nu = \rho_{n-1} \implies$ discrete EVI

Outline

- Ambrosio, Gigli, Savaré *original* construction of EVI flows
- EVI flows via Nonnegative Cross-Curvature (NNCC)
- The LJKO scheme

Another route to discrete EVI

Denote $g(x) := f(x) + \frac{1}{2\tau} d^2(x, x_{n-1}) - \left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} d^2(x, x_n)$ then discrete EVI is

$$g(x_n) \leq g(z), \quad \forall z \in X$$

Another route to discrete EVI

Denote $g(x) := f(x) + \frac{1}{2\tau} d^2(x, x_{n-1}) - \left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} d^2(x, x_n)$ then discrete EVI is

$$g(x_n) \leq g(z)\,, \quad \forall z \in X$$

If we find $\mathrm{x}(s)$ with $\mathrm{x}(0) = x_n$ and $\mathrm{x}(1) = z$, $g(\mathrm{x}(s))$ is convex and $\frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$ then

$$g(z) - g(x_n) \geq \frac{g(\mathrm{x}(s)) - g(x_n)}{s} \geq -\left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} \frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$$

Another route to discrete EVI

Denote $g(x) := f(x) + \frac{1}{2\tau} d^2(x, x_{n-1}) - \left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} d^2(x, x_n)$ then discrete EVI is

$$g(x_n) \leq g(z)\,, \quad \forall z \in X$$

If we find $\mathrm{x}(s)$ with $\mathrm{x}(0) = x_n$ and $\mathrm{x}(1) = z$, $g(\mathrm{x}(s))$ is convex and $\frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$ then

$$g(z) - g(x_n) \geq \frac{g(\mathrm{x}(s)) - g(x_n)}{s} \geq -\left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} \frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$$

The convexity condition on g can be split into

$$s \mapsto \frac{1}{2\tau} d^2(\mathrm{x}(s), x_{n-1}) - \frac{1}{2\tau} d^2(\mathrm{x}(s), x_n) \qquad \text{convex}$$

$$s \mapsto f(\mathrm{x}(s)) - \frac{\lambda}{2} d^2(\mathrm{x}(s), x_n) \qquad \text{convex}$$

Another route to discrete EVI

Denote $g(x) := f(x) + \frac{1}{2\tau} d^2(x, x_{n-1}) - \left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} d^2(x, x_n)$ then discrete EVI is

$$g(x_n) \leq g(z) \,, \quad \forall z \in X$$

If we find $\mathrm{x}(s)$ with $\mathrm{x}(0) = x_n$ and $\mathrm{x}(1) = z$, $g(\mathrm{x}(s))$ is convex and $\frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$ then

$$g(z) - g(x_n) \geq \frac{g(\mathrm{x}(s)) - g(x_n)}{s} \geq -\left(\frac{1}{\tau} + \lambda\right) \frac{1}{2} \frac{d^2(\mathrm{x}(s), x_n)}{s} \to 0$$

The convexity condition on g can be split into

$$s \mapsto \frac{1}{2\tau} d^2(\mathrm{x}(s), x_{n-1}) - \frac{1}{2\tau} d^2(\mathrm{x}(s), x_n) \qquad \text{convex}$$
$$s \mapsto f(\mathrm{x}(s)) - \frac{\lambda}{2} d^2(\mathrm{x}(s), x_n) \qquad \text{convex}$$

Remark: the condition "below the chord" is sufficient instead of convexity

How to ensure the first condition?

Smooth NNCC spaces

Let X, Y smooth manifolds, $c \in C^4(X \times Y)$ (plus some other regularity hypotheses)

Kim and McCann[1] introduced the pseudo-Riemannian metric

$$g_{KM}(x,y) = -\frac{1}{2}\begin{bmatrix} 0 & \nabla_{xy}c(x,y) \\ \nabla_{xy}c(x,y) & 0 \end{bmatrix}$$

c-segments: (x, y) with $\mathrm{x} : [0,1] \to X$:

$$\nabla_y c(\mathrm{x}(s), y) = (1-s)\nabla_y c(\mathrm{x}(0), y) + s\nabla_y c(\mathrm{x}(1), y)$$

$\longrightarrow$ particular geodesics where one variable is kept fix

Nonnegative cross-curvature: g_{KM} has nonnegative sectional curvature

[1] Kim, McCann (2010)

Smooth NNCC spaces

Let X, Y smooth manifolds, $c \in C^4(X \times Y)$ (plus some other regularity hypotheses)

Kim and McCann[1] introduced the pseudo-Riemannian metric

$$g_{KM}(x,y) = -\frac{1}{2}\begin{bmatrix} 0 & \nabla_{xy}c(x,y) \\ \nabla_{xy}c(x,y) & 0 \end{bmatrix}$$

c-segments: (x, y) with $\mathrm{x} : [0,1] \to X$:

$$\nabla_y c(\mathrm{x}(s), y) = (1-s)\nabla_y c(\mathrm{x}(0), y) + s\nabla_y c(\mathrm{x}(1), y)$$

$\longrightarrow$ particular geodesics where one variable is kept fix

Nonnegative cross-curvature: g_{KM} has nonnegative sectional curvature

Theorem (Kim-McCann[1]): $(X \times Y, c)$ has nonnegative cross-curvature $\iff$

$$s \mapsto c(\mathrm{x}(s), y) - c(\mathrm{x}(s), z) \text{ is convex } \forall z \in Y \text{ along c-segments } (\mathrm{x}, y)$$

[1] Kim, McCann (2010)

Nonsmooth NNCC spaces[1]

Key observation:

Theorem: Let X, Y and c be smooth, $\mathrm{x} : [0,1] \to X$ a curve such that, $\forall z \in Y$,

$$c(\mathrm{x}(s), y) - c(\mathrm{x}(s), z) \leq (1-s)\big[c(\mathrm{x}(0), y) - c(\mathrm{x}(0), z)\big] + s\big[c(\mathrm{x}(1), y) - c(\mathrm{x}(1), z)\big]$$

then (x, y) is a c-segment.

[1] Léger, Todeschi, Vialard (2024)

Nonsmooth NNCC spaces[1]

Key observation:

Theorem: Let X, Y and c be smooth, $\mathrm{x} : [0,1] \to X$ a curve such that, $\forall z \in Y$,

$$c(\mathrm{x}(s), y) - c(\mathrm{x}(s), z) \leq (1-s)\big[c(\mathrm{x}(0), y) - c(\mathrm{x}(0), z)\big] + s\big[c(\mathrm{x}(1), y) - c(\mathrm{x}(1), z)\big]$$

then (x, y) is a c-segment.

Let X, Y be any space and $c : X \times Y \to \mathbb{R} \cup \pm\infty$ any function

Variational c-segment: a curve (x, y) such that $\forall z \in Y$

$$c(\mathrm{x}(s), y) - c(\mathrm{x}(s), z) \leq (1-s)\big[c(\mathrm{x}(0), y) - c(\mathrm{x}(0), z)\big] + s\big[c(\mathrm{x}(1), y) - c(\mathrm{x}(1), z)\big]$$

Nonnegatively cross-curved (NNCC) space: $(X \times Y, c)$ is NNCC if $\forall x_0, x_1 \in X$ and $\forall y \in Y$, there exists a variational c-segment from (x_0, y) to (x_1, y).

[1] Léger, Todeschi, Vialard (2024)

Discrete EVI via NNCC

Let $(X \times X, c)$ be a cost space, $c : X \times X \to \mathbb{R}_{\geq 0}$, $c(x, x) = 0$

JKO with general cost: compute x_n iteratively as

$$x_n \in \operatorname{argmin} \frac{1}{\tau} c(x, x_{n-1}) + f(x)$$

Suppose x_n exists. We want x such that

$$s \mapsto c(\mathrm{x}(s), x_{n-1}) - c(\mathrm{x}(s), x_n) \quad \text{is "below the chord"}$$

Discrete EVI via NNCC

Let $(X \times X, c)$ be a cost space, $c : X \times X \to \mathbb{R}_{\geq 0}$, $c(x,x) = 0$

JKO with general cost: compute x_n iteratively as

$$x_n \in \operatorname{argmin} \frac{1}{\tau} c(x, x_{n-1}) + f(x)$$

Suppose x_n exists. We want x such that

$$s \mapsto c(\mathrm{x}(s), x_{n-1}) - c(\mathrm{x}(s), x_n) \quad \text{is "below the chord"}$$

If $(X \times X, c)$ is NNCC, take a variational c-segment (x, x_{n-1})

Discrete EVI via NNCC

Let $(X \times X, c)$ be a cost space, $c : X \times X \to \mathbb{R}_{\geq 0}$, $c(x, x) = 0$

JKO with general cost: compute x_n iteratively as

$$x_n \in \operatorname{argmin} \frac{1}{\tau} c(x, x_{n-1}) + f(x)$$

Suppose x_n exists. We want x such that

$$s \mapsto c(\mathrm{x}(s), x_{n-1}) - c(\mathrm{x}(s), x_n) \quad \text{is "below the chord"}$$

If $(X \times X, c)$ is NNCC, take a variational c-segment (x, x_{n-1})

If $(X \times X, -c)$ is NNCC, take a variational c-segment (x, x_n)

Discrete EVI via NNCC

Let $(X \times X, c)$ be a cost space, $c : X \times X \to \mathbb{R}_{\geq 0}$, $c(x, x) = 0$

JKO with general cost: compute x_n iteratively as

$$x_n \in \operatorname{argmin} \frac{1}{\tau} c(x, x_{n-1}) + f(x)$$

Suppose x_n exists. We want x such that

$$s \mapsto c(\mathrm{x}(s), x_{n-1}) - c(\mathrm{x}(s), x_n) \quad \text{is "below the chord"}$$

If $(X \times X, c)$ is NNCC, take a variational c-segment (x, x_{n-1})

If $(X \times X, -c)$ is NNCC, take a variational c-segment (x, x_n)

Then, if $s \mapsto f(\mathrm{x}(s)) - \lambda c(\mathrm{x}(s), y)$ is "below the chord", we obtain the discrete EVI

$$\frac{1}{\tau} c(z, x_n) - \frac{1}{\tau} c(z, x_{n-1}) \leq f(z) - f(x_n) - \lambda c(z, x_n) - \frac{1}{\tau} c(x_n, x_{n-1}), \quad \forall z \in X$$

Examples of NNCC spaces

- Hilbert space with $c = |\cdot - \cdot|^2$

  $\longrightarrow$ variational c-segments are simply $x(s) = (1-s)x_0 + sx_1$
- Bregman divergences for $u : X \to \mathbb{R}$

$$c(x,y) = u(x) - u(y) - \nabla u(y)(x-y)$$

- The Bures-Wasserstein distance squared on the space of symmetric positive semi-definite matrices

$$\mathrm{BW}^2(S_1, S_2) = \mathrm{tr}(S_1) + \mathrm{tr}(S_2) - 2\,\mathrm{tr}((S_1^{1/2} S_2 S_1^{1/2})^{1/2})\,.$$

  (a.k.a. the quadratic Wasserstein distance between gaussian measures)
- The Hellinger and the Fisher-Rao distances squared between probability measures
- The Kullback–Leibler divergence between probability measures
- The Gromov-Wasserstein distance squared between measure metric spaces

The Wasserstein space

Let X, Y be Polish spaces

$c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ lower semi-continuous and bounded from below

The optimal transport problem is

$$\mathcal{T}_c(\mu, \nu) := \min_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} c(x, y)\, d\gamma$$

Theorem: $(X \times Y, c)$ is NNCC $\iff$ $(\mathcal{P}(X) \times \mathcal{P}(Y), \mathcal{T}_c)$ is NNCC.

In particular:

The Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d), W_2^2)$ is NNCC

A similar result holds also for unbalanced optimal transport

Stability of NNCC

Stability by products (works also for infinite products):

If $(X_1 \times Y_1, c_1)$ and $(X_2 \times Y_2, c_2)$ are NNCC then so is $\big((X_1 \times X_2) \times (Y_1 \times Y_2), c_1 + c_2\big)$.

Stability by a certain type of projections:

Let $P_1 : X \to \underline{X}$ and $P_2 : Y \to \underline{Y}$ and

$$\underline{c}(\underline{x}, \underline{y}) = \inf \left\{ c(x, y) \,,\, P_1(x) = \underline{x}, P_2(y) = \underline{y} \right\}$$

If $(X \times Y, c)$ is NNCC then (under some hypotheses) $(\underline{X} \times \underline{Y}, \underline{c})$ is NNCC.

Stability by Gromov-Hausdorff convergence of compact metric spaces:

If $(X_k \times X_k, d_k^2)$ is a sequence of NNCC spaces which converges in the Gromov-Hausdorff sense to $(X \times X, d^2)$, the limit is NNCC.

Wasserstein space

Let $(X \times Y, c)$ be an NNCC space, $\mu_0, \mu_1 \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$

Assume (for simplicity) $\exists T_0, T_1$ optimal transport maps for $(\mu_0, \nu), (\mu_1, \nu)$, i.e.

$$\mathcal{T}_c(\mu_0, \nu) = \int_Y c(T_0(y), y)\nu(y)\,, \quad \mathcal{T}_c(\mu_1, \nu) = \int_Y c(T_1(y), y)\nu(y)$$

Let $\Lambda_s : X \times X \times Y$ maps triplets (x_0, x_1, y) to evaluation at time s of a corresponding variational c-segment

$$\Lambda_s(x_0, x_1, y) \mapsto \mathrm{x}(s)$$

Wasserstein space

Let $(X \times Y, c)$ be an NNCC space, $\mu_0, \mu_1 \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$

Assume (for simplicity) $\exists T_0, T_1$ optimal transport maps for $(\mu_0, \nu), (\mu_1, \nu)$, i.e.

$$\mathcal{T}_c(\mu_0, \nu) = \int_Y c(T_0(y), y)\nu(y)\,, \quad \mathcal{T}_c(\mu_1, \nu) = \int_Y c(T_1(y), y)\nu(y)$$

Let $\Lambda_s : X \times X \times Y$ maps triplets (x_0, x_1, y) to evaluation at time s of a corresponding variational c-segment

$$\Lambda_s(x_0, x_1, y) \mapsto \mathrm{x}(s)$$

Lifted c-segments: $\mu : [0,1] \to \mathcal{P}(X)$ defined as

$$\mu(s) = (\Lambda_s(T_0(y), T_1(y), y))_{\#}\nu\,, \quad s \in [0,1]$$

If $X = Y = \mathbb{R}^d$ and $c = |\cdot - \cdot|^2$ lifted c-segments are generalized geodesics!

Wasserstein space

Let $(X \times Y, c)$ be an NNCC space, $\mu_0, \mu_1 \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$

Assume (for simplicity) $\exists T_0$, T_1 optimal transport maps for $(\mu_0, \nu), (\mu_1, \nu)$, i.e.

$$\mathcal{T}_c(\mu_0, \nu) = \int_Y c(T_0(y), y)\nu(y)\,, \quad \mathcal{T}_c(\mu_1, \nu) = \int_Y c(T_1(y), y)\nu(y)$$

Let $\Lambda_s : X \times X \times Y$ maps triplets (x_0, x_1, y) to evaluation at time s of a corresponding variational c-segment

$$\Lambda_s(x_0, x_1, y) \mapsto \mathrm{x}(s)$$

Lifted c-segments: $\mu : [0, 1] \to \mathcal{P}(X)$ defined as

$$\mu(s) = (\Lambda_s(T_0(y), T_1(y), y))_{\#}\nu\,, \quad s \in [0, 1]$$

If $X = Y = \mathbb{R}^d$ and $c = |\cdot - \cdot|^2$ lifted c-segments are generalized geodesics!

Theorem: Lifted c-segments are variational c-segments on $(\mathcal{P}(X) \times \mathcal{P}(Y), \mathcal{T}_c)$.

Relation with Assumption 4.0.1

Let (X, d) be geodesic space

Theorem: If $(X \times X, d^2)$ is NNCC then (X, d) is a PC space.

(extension of a known result for classical NNCC)

Relation with Assumption 4.0.1

Let (X, d) be geodesic space

Theorem: If $(X \times X, d^2)$ is NNCC then (X, d) is a PC space.

(extension of a known result for classical NNCC)

Let $y \in Y$ and (x, y) be a variational c-segment then

Theorem: $d^2(\mathrm{x}(s), y)$ is 1-convex in the sense that

$$d^2(\mathrm{x}(s), y) \leq (1-s)d^2(\mathrm{x}(0), y)) + sd^2(\mathrm{x}(1), y) - s(1-s)d^2(\mathrm{x}(0), \mathrm{x}(1))$$

$\implies$ variational c-segments satisfy the assumption of Ambrosio, Gigli, Savaré

On the other hand, we do not restrict to metric spaces

Outline

- Ambrosio, Gigli, Savaré *original* construction of EVI flows
- EVI flows via Nonnegative Cross-Curvature (NNCC)
- The LJKO scheme

From JKO to LJKO

JKO scheme: compute ρ_n recursively as

$$\rho_n \in \operatorname{argmin} \frac{1}{2\tau} W_2^2(\rho, \rho_{n-1}) + \mathcal{E}(\rho)$$

[1]Benamou, Brenier (2000), [2]Villani (2003)

From JKO to LJKO

JKO scheme: compute ρ_n recursively as

$$\rho_n \in \operatorname{argmin} \frac{1}{2\tau} W_2^2(\rho, \rho_{n-1}) + \mathcal{E}(\rho)$$

Dynamical form of the squared Wasserstein distance[1]:

$$W_2^2(\mu, \nu) = \inf_{\rho \geq 0, m} \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|m|^2}{\rho}, \ \partial_t \rho + \operatorname{div}(m) = 0, \ \rho(0, \cdot) = \mu, \ \rho(1, \cdot) = \nu \right\}$$

Suited for Eulerian discretization

$W_2^2(\rho, \rho_{n-1})$ expensive optimization problem

[1]Benamou, Brenier (2000), [2]Villani (2003)

JKO scheme: compute ρ_n recursively as

$$\rho_n \in \operatorname{argmin} \frac{1}{2\tau} W_2^2(\rho, \rho_{n-1}) + \mathcal{E}(\rho)$$

Dynamical form of the squared Wasserstein distance[1]:

$$W_2^2(\mu,\nu) = \inf_{\rho \geq 0, m} \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|m|^2}{\rho},\ \partial_t \rho + \operatorname{div}(m) = 0,\ \rho(0,\cdot) = \mu,\ \rho(1,\cdot) = \nu \right\}$$

Suited for Eulerian discretization

$W_2^2(\rho, \rho_{n-1})$ expensive optimization problem

It holds[2] $\|\rho - \mu\|_{\dot{H}_\mu^{-1}} = W_2(\rho,\mu) + o(W_2(\rho,\mu)),\ \forall \rho,\mu \in \mathcal{P}_2(\mathbb{R}^d)$, where

$$\|\rho - \mu\|^2_{\dot{H}_\mu^{-1}} = \sup_\phi \left\{ \int_{\mathbb{R}^d} \phi(\rho - \mu),\ ||\phi||_{H_\mu^1} \leq 1 \right\}$$

[1]Benamou, Brenier (2000), [2]Villani (2003)

From JKO to LJKO

The H^{-1} norm can be rewritten as

$$\|\rho - \mu\|_{\dot{H}_\mu^{-1}}^2 = \inf_m \left\{ \int_{\mathbb{R}^d} \frac{|m|^2}{\mu}, \ \mu - \rho + \operatorname{div}(m) = 0 \right\}$$

$\longrightarrow$ one-step time approximation of the dynamical Wasserstein distance

[1] Cancès, Gallouët, Todeschi (2019)

From JKO to LJKO

The H^{-1} norm can be rewritten as

$$\left\|\rho-\mu\right\|^2_{\dot{H}^{-1}_\mu} = \inf_m \left\{ \int_{\mathbb{R}^d} \frac{|m|^2}{\mu},\ \mu-\rho+\operatorname{div}(m)=0 \right\}$$

$\longrightarrow$ one-step time approximation of the dynamical Wasserstein distance

Linearized JKO (LJKO) scheme[1]: compute ρ_n recursively as

$$\rho_n \in \operatorname*{arginf}_{\rho\in\mathcal{P}_2(\mathbb{R}^d)} \frac{1}{2\tau}\left\|\rho^{n-1}-\rho\right\|^2_{\dot{H}^{-1}_\rho} + \mathcal{E}(\rho)$$

- simpler convex optimization problem
- the metric structure (of the discrete scheme) is lost
- does $c(\mu,\nu) := \left\|\mu-\nu\right\|^2_{\dot{H}^{-1}_\mu}$ satisfy NNCC?

[1] Cancès, Gallouët, Todeschi (2019)

LJKO scheme

Original PDE formulation

LJKO scheme

$$\partial_t \rho - \mathrm{div}\big(\rho \nabla \frac{\delta \mathcal{E}}{\delta \rho}(\rho)\big) = 0$$

LJKO scheme

Original PDE formulation

$$\begin{cases} \partial_t \rho - \mathrm{div}(\rho \nabla \phi) = 0 \\ \phi = \dfrac{\delta \mathcal{E}}{\delta \rho}(\rho) \quad \rho\text{-a.e.} \end{cases}$$

LJKO scheme

LJKO scheme

Original PDE formulation

$$\begin{cases} \partial_t \rho - \mathrm{div}(\rho \nabla \phi) = 0 \\ \phi = \dfrac{\delta \mathcal{E}}{\delta \rho}(\rho) \quad \rho\text{-a.e.} \end{cases}$$

LJKO scheme

$$\begin{cases} \dfrac{1}{\tau}(\rho_n - \rho_{n-1}) - \mathrm{div}(\rho_n \nabla \phi_n) = 0 \\ \phi_n + \dfrac{\tau}{2}|\nabla \phi_n|^2 = \dfrac{\delta \mathcal{E}}{\delta \rho}(\rho_n) \quad \rho_n\text{-a.e.} \end{cases}$$

LJKO scheme

Original PDE formulation

$$\begin{cases} \partial_t \rho - \mathrm{div}(\rho \nabla \phi) = 0 \\ \phi = \dfrac{\delta \mathcal{E}}{\delta \rho}(\rho) \quad \rho\text{-a.e.} \end{cases}$$

LJKO scheme

$$\begin{cases} \dfrac{1}{\tau}(\rho_n - \rho_{n-1}) - \mathrm{div}(\rho_n \nabla \phi_n) = 0 \\ \phi_n + \dfrac{\tau}{2}|\nabla \phi_n|^2 = \dfrac{\delta \mathcal{E}}{\delta \rho}(\rho_n) \quad \rho_n\text{-a.e.} \end{cases}$$

The HJ equation can be saturated (equality holds almost everywhere)

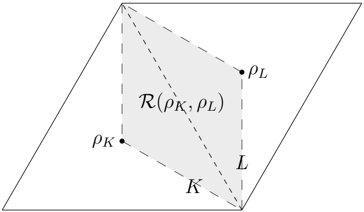
Convergence towards the PDE for $\tau \to 0$

- requires some regularity for the potential ϕ
- example: linear Fokker-Planck equation

TPFA discretization

TPFA approximation of the flux

$$m = \rho \nabla \phi \ \approx \ \mathcal{R}(\rho_K, \rho_L) \frac{(\phi_K - \phi_L)}{|x_K - x_L|}$$



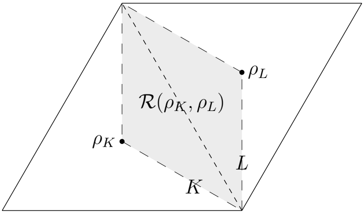
TPFA discretization

TPFA approximation of the flux

$$m = \rho\nabla\phi \;\approx\; \mathcal{R}(\rho_K, \rho_L)\frac{(\phi_K - \phi_L)}{|x_K - x_L|}$$

Upwind choice:

$$\mathcal{R}(\rho_K, \rho_L) = \begin{cases} \rho_K & \text{if } \phi_K \geq \phi_L \\ \rho_L & \text{else} \end{cases}$$

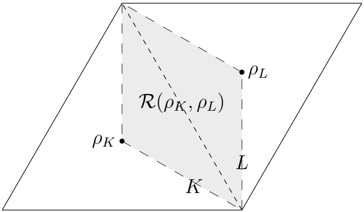


- preserve monotonicity $\implies$ HJ equation can be saturated at the discrete level
- very efficient, first order accurate

TPFA discretization

TPFA approximation of the flux

$$m = \rho \nabla \phi \;\approx\; \mathcal{R}(\rho_K, \rho_L) \frac{(\phi_K - \phi_L)}{|x_K - x_L|}$$



Upwind choice:

$$\mathcal{R}(\rho_K, \rho_L) = \begin{cases} \rho_K & \text{if } \phi_K \geq \phi_L \\ \rho_L & \text{else} \end{cases}$$

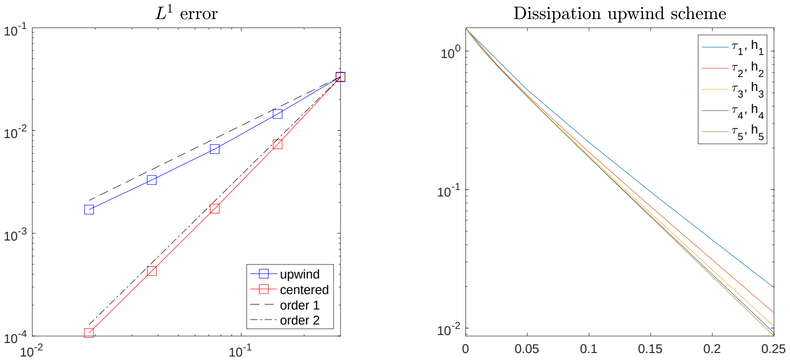
- preserve monotonicity $\implies$ HJ equation can be saturated at the discrete level
- very efficient, first order accurate

Centered choice:

$$\mathcal{R}(\rho_K, \rho_L) = (1 - \theta)\rho_K + \theta \rho_L \,, \quad \text{for } \theta \in (0, 1)$$

- does not preserve monotonicity $\implies$ HJ equation cannot be saturated
- less efficient but second order accurate

2D convergence test



Linear Fokker-Planck equation: W_2 gradient flow of $\mathcal{E}(\rho) = \int_\Omega \rho \log(\rho) + \rho V$

- the LJKO scheme is one order accurate in time
- energy decreases exponentially fast

Does $c(\mu,\nu) := \|\mu-\nu\|^2_{\dot{H}^{-1}_\mu}$ satisfy NNCC?

A variational c-segment $\mu(s)$ between (μ_0,ν) and (μ_1,ν) must satisfy

$$\phi(s) = (1-s)\phi_0 + s\phi_1$$

where ϕ_0 optimal for (μ_0,ν) and ϕ_1 optimal for (μ_1,ν)

Then the "candidate" variational c-segment must solve

$$\mu(s) - \nu - \operatorname{div}(\mu(s)\nabla\phi(s)) = 0$$

Does $c(\mu,\nu) := \|\mu-\nu\|^2_{\dot{H}^{-1}_\mu}$ satisfy NNCC?

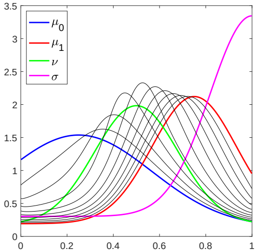
A variational c-segment $\mu(s)$ between (μ_0,ν) and (μ_1,ν) must satisfy

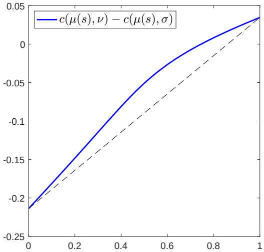
$$\phi(s) = (1-s)\phi_0 + s\phi_1$$

where ϕ_0 optimal for (μ_0,ν) and ϕ_1 optimal for (μ_1,ν)

Then the "candidate" variational c-segment must solve

$$\mu(s) - \nu - \mathrm{div}(\mu(s)\nabla\phi(s)) = 0$$





Thank you for your attention!