A bound between the JKO and entropic JKO schemes

Sofiane Cherf, Aymeric Baradat

Institut Camille Jordan

Funded by the European Union

European Research Council
Established by the European Commission

EYAWKAJKOS
Everything You Always Wanted to Know About the JKO Scheme

Optimal transport and entropic optimal transport

Definition

The Benamou Brenier formula for the Wasserstein distance is:

$$\frac{W_2^2(\mu,\nu)}{2} = \min\left\{ \frac{1}{2}\int_0^1 \int |v|^2 \, \mathrm{d}\rho \mathrm{d}t \;\middle|\; \begin{array}{l} \rho(0) = \mu, \quad \rho(1) = \nu, \\ \partial_t \rho + \mathrm{div}(\rho v) = 0 \end{array} \right\},$$

Optimal transport and entropic optimal transport

Definition

The Benamou Brenier formula for the Wasserstein distance reparametrize in time is:

$$\frac{W_2^2(\mu,\nu)}{2\tau} = \min\left\{ \frac{1}{2}\int_0^{\tau}\int |v|^2 \,\mathrm{d}\rho \mathrm{d}t \;\middle|\; \begin{array}{l} \rho(0)=\mu, \quad \rho(\tau)=\nu, \\ \partial_t\rho + \mathrm{div}(\rho v) = 0 \end{array} \right\},$$

Optimal transport and entropic optimal transport

Definition

The Benamou Brenier formula for the Wasserstein distance reparametrize in time is:

$$\frac{W_2^2(\mu,\nu)}{2\tau} = \min\left\{ \frac{1}{2}\int_0^\tau \int |v|^2 \,\mathrm{d}\rho \mathrm{d}t \;\middle|\; \begin{array}{l} \rho(0)=\mu, \quad \rho(\tau)=\nu, \\ \partial_t\rho + \mathrm{div}(\rho v) = 0 \end{array} \right\},$$

Schrödinger cost with regularized-parameter $\varepsilon := \alpha\tau$ can be defined as:

$$\frac{Sch^{\alpha\tau}(\mu,\nu)}{\tau} = \frac{\alpha}{2}(H(\mu)+H(\nu)) + \min\left\{ \frac{1}{2}\int_0^\tau \int |c|^2 + \left|\frac{\alpha}{2}\nabla \ln(\rho)\right|^2 \mathrm{d}\rho \mathrm{d}t \;\middle|\; \begin{array}{l} \rho(0)=\mu, \quad \rho(\tau)=\nu, \\ \partial_t\rho + \mathrm{div}(\rho c) = 0 \end{array} \right\}.$$

Where H is the Boltzmann entropy.

JKO scheme

Let $V, W \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}_+)$ with W symmetric and $f \in \mathcal{C}^3(\mathbb{R}_+, \mathbb{R}_+)$

$$\mathcal{F}(\rho) := \begin{cases} \int_{\mathbb{R}^d} V(x)\,\rho(x)\,dx + \int_{\mathbb{R}^d} (W * \rho)(x)\,\rho(x)\,dx + \int_{\mathbb{R}^d} f(\rho(x))\,dx, & \text{if } \rho \ll \text{Leb} \\ +\infty & \text{otherwise} \end{cases}$$

Then we can define the one step JKO associated to $\mathcal{F}$

$$J_\tau^0(\mu) := \underset{\rho \in \mathcal{P}_2(\mathbb{R}^d)}{\arg\min} \left\{ \frac{W_2^2(\mu, \rho)}{2\tau} + \mathcal{F}(\rho) \right\},$$

where W_2 is the Wasserstein distance; the iterate scheme is:
$J_{k\tau}^0(\mu) := (J_\tau^0)^{\circ k}(\mu) = \underbrace{J_\tau^0 \circ \cdots \circ J_\tau^0}_{k \text{ times}}(\mu).$

Entropic JKO

We can define the one-step entropic JKO associated to $\mathcal{F}$

$$J_\tau^\alpha(\mu) \in \underset{\rho \in \mathcal{P}_2(\mathbb{R}^d)}{\arg\min} \left\{ \frac{Sch^{\alpha\tau}(\mu,\rho)}{\tau} + \mathcal{F}(\rho) \right\},$$

and the iterate scheme: $J_{k\tau}^\alpha(\mu) := (J_\tau^\alpha)^{\circ k}(\mu) = \underbrace{J_\tau^\alpha \circ \cdots \circ J_\tau^\alpha}_{k \text{ times}}(\mu)$.

When it's possible let's define $\bar{\rho}^{\alpha,\tau}(t) := J^\alpha_{\lceil \frac{t}{\tau} \rceil \tau}(\mu_0)$.

Entropic JKO

We can define the one-step entropic JKO associated to $\mathcal{F}$

$$J_\tau^\alpha(\mu) \in \underset{\rho \in \mathcal{P}_2(\mathbb{R}^d)}{\arg\min} \left\{ \frac{Sch^{\alpha\tau}(\mu,\rho)}{\tau} + \mathcal{F}(\rho) \right\},$$

and the iterate scheme: $J_{k\tau}^\alpha(\mu) := (J_\tau^\alpha)^{\circ k}(\mu) = \underbrace{J_\tau^\alpha \circ \cdots \circ J_\tau^\alpha}_{k \text{ times}}(\mu)$.

When it's possible let's define $\bar{\rho}^{\alpha,\tau}(t) := J^\alpha_{\lceil \frac{t}{\tau} \rceil \tau}(\mu_0)$.

Question

Is $\bar{\rho}^{\alpha,\tau}$ a good approximation of the solution of:

$$\begin{cases} \partial_t \rho - \operatorname{div}\left(\rho(\nabla V + \nabla W * \rho + \nabla(f'(\rho)))\right) = 0 \\ \rho(0,\cdot) = \mu_0 \end{cases}$$

A prior convergence result

Let's recall that $\bar{\rho}^{\alpha,\tau}(t) := J^{\alpha}_{\lceil \frac{t}{\tau} \rceil \tau}(\mu_0)$

Baradat, Hraivoronska, Santambrogio (25)

For $\mathcal{F}(\rho) := \int_{\mathbb{R}^d} V(x)\,\rho(x)\,dx + \int_{\mathbb{R}^d} (W * \rho)(x)\,\rho(x)\,dx + \int_{\mathbb{R}^d} f(\rho(x))\,dx, \quad$ if $\rho \ll$ Leb

with

- $(\Delta V)_+ \leq K, \quad (\Delta W)_+ \leq K$
- f is convex and verify some growth assumptions

then $\bar{\rho}^{\alpha,\tau}$ is well-defined and converge when $\tau \to 0$ in the distributional sense to a distributional solution of:

$$\begin{cases} \partial_t \rho - \operatorname{div}\left(\rho(\nabla V + \nabla W * \rho + \nabla(f'(\rho)))\right) = \frac{\alpha}{2}\Delta\rho \\ \rho(0,\cdot) = \mu_0 \end{cases}$$

The continuous case

Let's first take a look at the continuous level. For this let's introduce the solution of the limiting PDE's:
Let's call ρ^0 the solution of:

$$\begin{cases} \partial_t \rho - \operatorname{div}\left(\rho(\nabla V + \nabla W * \rho + \nabla(f'(\rho)))\right) = 0 \\ \rho(0, \cdot) = \mu_0 \end{cases}$$

Let's call ρ^α the solution of:

$$\begin{cases} \partial_t \rho - \operatorname{div}\left(\rho(\nabla V + \nabla W * \rho + \nabla(f'(\rho)))\right) = \frac{\alpha}{2}\Delta\rho \\ \rho(0, \cdot) = \mu_0 \end{cases}$$

At the continuous level, we have a bound under extra assumption on $\mathcal{F}$:

Theorem ('Fanch's calculus')

- V *convex,*
- W *symmetric and convex,*
- $s \mapsto s^d f(s^{-d})$ *is convex and non increasing on* $(0, +\infty)$ *(McCann condition),*

then $\mathcal{F}$ *is convex along generalized geodesics. Moreover, if:*

- $\Delta V \leq \tilde{K}$,
- $\Delta W \leq \bar{K}$,
- f *is convex,*

then for every $t > 0$ *and* $K = \tilde{K} + \bar{K}$ *the following inequality holds:*

$$W_2(\rho^0(t), \rho^\alpha(t)) \leq \sqrt{\frac{\alpha}{2} t(H(\mu_0) - H(\rho^\alpha(t)) + Kt)}$$

Main results

At the discrete level we have some similar inequlity:

C. Baradat (in prep)

Under the same assumption. Then the iterates $J^0_{k\tau}(\mu_0)$ and $J^\alpha_{k\tau}(\mu_0)$ are well-defined and satisfy:

$$W_2(J^0_{n\tau}(\mu_0), J^\alpha_{n\tau}(\mu_0)) \leq \sqrt{2\tau(\mathcal{F}(\mu_0) - \mathcal{F}(J^0_{n\tau}(\mu_0))} + \sqrt{n\tau\alpha\,(H(\mu_0) - H(J^\alpha_{n\tau}(\mu_0)) + Kn\tau)}$$

Main results

At the discrete level we have some similar inequlity:

C. Baradat (in prep)

Under the same assumption. Then the iterates $J^0_{k\tau}(\mu_0)$ and $J^\alpha_{k\tau}(\mu_0)$ are well-defined and satisfy:

$$W_2(J^0_{n\tau}(\mu_0), J^\alpha_{n\tau}(\mu_0)) \leq \sqrt{2\tau(\mathcal{F}(\mu_0) - \mathcal{F}(J^0_{n\tau}(\mu_0))} + \sqrt{n\tau\alpha\left(H(\mu_0) - H(J^\alpha_{n\tau}(\mu_0)) + Kn\tau\right)}$$

We can give a first answer to our result

Corollary

$$W_2(\bar{\rho}^{\alpha,\tau}(t), \rho^0(t)) \leq \tilde{C}(\sqrt{\alpha} + \sqrt{\tau})$$

where $\tilde{C}$ depends only on the second moment of μ_0, t, $\mathcal{F}(\mu_0)$, $H(\mu_0)$, and K.

At the continuous level, we have a bound under extra assumption on $\mathcal{F}$:

Theorem

- V λ_1-*convex,*
- W *symmetric and* λ_2-*convex,*
- $s \mapsto s^d f(s^{-d})$ *is convex and non increasing on* $(0, +\infty)$ *(McCann condition),*

then $\mathcal{F}$ *is* $\lambda := \lambda_1 + \lambda_2$-*convex along generalized geodesics. Moreover, if:*

- $\Delta V \leq \tilde{K}$,
- $\Delta W \leq \bar{K}$,
- f *is convex,*

then for every $t > 0$ *and* $K = \tilde{K} + \bar{K}$ *the following inequality holds:*

$$W_2(\rho^0(t), \rho^\alpha(t)) \leq \sqrt{\frac{1 - e^{-2\lambda t}}{2\lambda} \frac{\alpha}{2} (H(\mu_0) - H(\rho^\alpha(t)) + Kt)}$$

Main results

At the discrete level we have some similar inequlity:

C. Baradat (in prep)

Under the same assumption. For every $\tau < \frac{1}{2\lambda_-}$

$$W_2(J^0_{n\tau}(\mu_0), J^\alpha_{n\tau}(\mu_0)) \leq 4(1-\lambda_-\tau)^{-n}\sqrt{\tau}\sqrt{(\mathcal{F}(\mu_0) - \mathcal{F}(J^0_{n\tau}(\mu_0))}$$
$$+(1+3\lambda_-\tau)\sqrt{\frac{1-(1+\lambda\tau)^{-2n}}{2\lambda}}\sqrt{\alpha\left(H(\mu_0) - H(J^\alpha_{n\tau}(\mu_0)) + Kn\tau\right)}$$

We can give a first answer to our result

Corollary

$$W_2(\bar{\rho}^{\alpha,\tau}(t), \rho^0(t)) \leq \tilde{C}(\sqrt{\alpha} + \sqrt{\tau})$$

where $\tilde{C}$ depends only on the second moment of μ_0, λ, t, $\mathcal{F}(\mu_0)$, $H(\mu_0)$, and K.

Idea behind the proof

Let's look at the difference step by step:

$$W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) = W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))$$

We dealing with one step of two different schemes starting from two different measures, let's rewrite the inequality in two sub-problem:

$$\begin{aligned} &W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) \\ &\leq \underbrace{W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^0_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(I)}} + \underbrace{W_2(J^0_\tau(J^\alpha_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(II)}}. \end{aligned}$$

Idea behind the proof

Let's look at the difference step by step:

$$W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) = W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))$$

We dealing with one step of two different schemes starting from two different measures, let's rewrite the inequality in two sub-problem:

$$\begin{aligned} &W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) \\ &\leq \underbrace{W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^0_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(I)}} + \underbrace{W_2(J^0_\tau(J^\alpha_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(II)}}. \end{aligned}$$

- (I) is the distance between the measure obtain from the same scheme but starting from different measures.

Idea behind the proof

Let's look at the difference step by step:

$$W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) = W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))$$

We dealing with one step of two different schemes starting from two different measures, let's rewrite the inequality in two sub-problem:

$$\begin{aligned} &W_2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) \\ &\leq \underbrace{W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^0_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(I)}} + \underbrace{W_2(J^0_\tau(J^\alpha_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(II)}}. \end{aligned}$$

- (I) is the distance between the measure obtain from the same scheme but starting from different measures.
- (II) is the distance between the image from the same measure by the JKO scheme and the entropic JKO.

Upper Bound on Term (I)

Discrete Evolutionary variational inequality (Discrete EVI)

For every $\rho, \upsilon \in \mathcal{P}_2(\mathbb{R}^d)$.

$$\frac{1}{2\tau}\left(W_2^2(\rho, J_\tau^0(\upsilon)) - W_2^2(\rho, \upsilon)\right) \leq \mathcal{F}(\rho) - \mathcal{F}(J_\tau^0(\upsilon)) - \frac{1}{2\tau} W_2^2(J_\tau^0(\upsilon), \upsilon)$$

Ambrosio, Gigli, Savaré (2005)

$$(\mathrm{I})^2 \leq W_2^2(J_{k\tau}^0(\mu_0), J_{k\tau}^\alpha(\mu_0)) + 2\tau\left(\mathcal{F}(J_{k\tau}^0(\mu_0)) - \mathcal{F}(J_{(k+1)\tau}^0(\mu_0))\right)$$

Upper Bound on Term (II)

Discrete Evolutionary variational inequality (Discrete EVI)

For every $\rho, \upsilon \in \mathcal{P}_2(\mathbb{R}^d)$.

$$\frac{1}{2\tau} W_2^2(\rho, J_\tau^0(\upsilon)) \leq \mathcal{F}(\rho) + \frac{1}{2\tau} W_2^2(\rho, \upsilon) - \left(\mathcal{F}(J_\tau^0(\upsilon)) + \frac{1}{2\tau} W_2^2(J_\tau^0(\upsilon), \upsilon) \right)$$

Theorem

For all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*:*

$$\frac{1}{2\tau} W_2^2(J_\tau^0(\mu), J_\tau^\alpha(\mu)) \leq \frac{1}{1+\lambda\tau} \left(\frac{W_2^2(\mu, J_\tau^\alpha(\mu))}{2\tau} + \mathcal{F}(J_\tau^\alpha(\mu)) - \left(\frac{W_2^2(\mu, J_\tau^0(\mu))}{2\tau} + \mathcal{F}(J_\tau^0(\mu)) \right) \right)$$

Theorem

For all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$\frac{1}{2\tau} W_2^2(J_\tau^0(\mu), J_\tau^\alpha(\mu)) \leq \frac{1}{1+\lambda\tau} \left(\frac{W_2^2(\mu, J_\tau^\alpha(\mu))}{2\tau} + \mathcal{F}(J_\tau^\alpha(\mu)) - \left(\frac{W_2^2(\mu, J_\tau^0(\mu))}{2\tau} + \mathcal{F}(J_\tau^0(\mu)) \right) \right)$$

Lemma 1

For all ρ, ν such that $H(\rho) < +\infty$ and $H(\nu) < +\infty$, then the following inequality holds:

$$\frac{Sch^{\alpha\tau}(\rho, \nu)}{\tau} \geq \alpha \frac{H(\rho) + H(\nu)}{2} + \frac{W_2^2(\rho, \nu)}{2\tau}$$

Theorem

For all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*:*

$$\frac{1}{2\tau} W_2^2(J_\tau^0(\mu), J_\tau^\alpha(\mu)) \le \frac{1}{1+\lambda\tau}\left(\frac{W_2^2(\mu, J_\tau^\alpha(\mu))}{2\tau} + \mathcal{F}(J_\tau^\alpha(\mu)) - \left(\frac{W_2^2(\mu, J_\tau^0(\mu))}{2\tau} + \mathcal{F}(J_\tau^0(\mu))\right)\right)$$

Lemma 1

For all ρ, ν such that $H(\rho) < +\infty$ and $H(\nu) < +\infty$, then the following inequality holds:

$$\frac{Sch^{\alpha\tau}(\rho, \nu)}{\tau} \ge \alpha \frac{H(\rho) + H(\nu)}{2} + \frac{W_2^2(\rho, \nu)}{2\tau}$$

Lemma 2

$$\frac{\mathrm{Sch}^{\alpha\tau}(\mu, J_\tau^\alpha(\mu))}{\tau} + \mathcal{F}(J_\tau^\alpha(\mu)) - \left(\frac{W_2^2(\mu, J_\tau^0(\mu))}{2\tau} + \mathcal{F}(J_\tau^0(\mu))\right) \le \alpha H(\mu) + K\frac{\alpha}{2}\tau$$

Comparing iterates at step $(k+1)$

$$W_2^2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) \leq \Big(\underbrace{W_2(J^0_\tau(J^0_{k\tau}(\mu_0)), J^0_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(I)}} + \underbrace{W_2(J^0_\tau(J^\alpha_{k\tau}(\mu_0)), J^\alpha_\tau(J^\alpha_{k\tau}(\mu_0)))}_{\text{(II)}}\Big)^2$$

Bounds on (I) and (II)

$$\begin{aligned}\text{(I)}^2 \ &\leq \ W_2^2(J^0_{k\tau}(\mu_0), J^\alpha_{k\tau}(\mu_0)) + 2\tau\Big(\mathcal{F}(J^0_{k\tau}(\mu_0)) - \mathcal{F}(J^0_{(k+1)\tau}(\mu_0))\Big),\\ \text{(II)}^2 \ &\leq \ K\alpha\tau + \alpha\Big(H(J^\alpha_{k\tau}(\mu_0)) - H(J^\alpha_{(k+1)\tau}(\mu_0))\Big).\end{aligned}$$

$$\begin{aligned}W_2^2(J^0_{(k+1)\tau}(\mu_0), J^\alpha_{(k+1)\tau}(\mu_0)) \ \leq \ \Bigg(&\sqrt{W_2^2(J^0_{k\tau}(\mu_0), J^\alpha_{k\tau}(\mu_0)) + 2\tau\Big(\mathcal{F}(J^0_{k\tau}(\mu_0)) - \mathcal{F}(J^0_{(k+1)\tau}(\mu_0))\Big)}\\ &+ \sqrt{\tau}\sqrt{K\alpha\tau + \alpha\Big(H(J^\alpha_{k\tau}(\mu_0)) - H(J^\alpha_{(k+1)\tau}(\mu_0))\Big)}\Bigg)^2\end{aligned}$$

We conclude thanks to

Lemma (Squared discrete Gronwall lemma)

Let $(a_k), (b_k) \geq 0$*. If* (u_k) *satisfies*

$$u_{k+1}^2 \leq \left(\sqrt{u_k^2 + a_{k+1}} + b_{k+1} \right)^2, \quad u_0 = 0,$$

then

$$u_n \leq \sqrt{\sum_{k=1}^{n} a_k} + \sum_{k=1}^{n} b_k.$$

$$\text{for } u_k = W_2(J^0_{k\tau}(\mu_0), J^\alpha_{k\tau}(\mu_0)), \quad a_{k+1} = 2\tau \Big(\mathcal{F}(J^0_{k\tau}(\mu_0)) - \mathcal{F}(J^0_{(k+1)\tau}(\mu_0)) \Big),$$

$$b_{k+1} = \sqrt{\tau} \sqrt{K\alpha\tau + \alpha \Big(H(J^\alpha_{k\tau}(\mu_0)) - H(J^\alpha_{(k+1)\tau}(\mu_0)) \Big)}.$$

and the Cauchy-Schwartz inequality.

C. Baradat (in prep)

Under the same assumption. Then the iterates $J^0_{k\tau}(\mu_0)$ and $J^\alpha_{k\tau}(\mu_0)$ are well-defined and satisfy:

$$W_2(J^0_{n\tau}(\mu_0), J^\alpha_{n\tau}(\mu_0)) \leq \sqrt{2\tau(\mathcal{F}(\mu_0) - \mathcal{F}(J^0_{n\tau}(\mu_0))} + \sqrt{n\tau\alpha\left(H(\mu_0) - H(J^\alpha_{n\tau}(\mu_0)) + Kn\tau\right)}$$

C. Baradat (in prep)

Under the same assumption. Then the iterates $J^0_{k\tau}(\mu_0)$ and $J^\alpha_{k\tau}(\mu_0)$ are well-defined and satisfy:

$$W_2(J^0_{n\tau}(\mu_0), J^\alpha_{n\tau}(\mu_0)) \leq \sqrt{2\tau(\mathcal{F}(\mu_0) - \mathcal{F}(J^0_{n\tau}(\mu_0))} + \sqrt{n\tau\alpha\left(H(\mu_0) - H(J^\alpha_{n\tau}(\mu_0)) + Kn\tau\right)}$$

From the classic result

Theorem (Ambrosio, Gigli, Savaré (2005))

$$W_2(\rho^\tau(t), \rho^0(t)) \leq \sqrt{2\tau\mathcal{F}(\mu_0)}$$

we easily deduce

Corollary

$$W_2(\bar{\rho}^{\alpha,\tau}(t), \rho^0(t)) \leq \tilde{C}(\sqrt{\alpha} + \sqrt{\tau})$$

where $\tilde{C}$ depends only on the second moment of μ_0, t, $\mathcal{F}(\mu_0)$, $H(\mu_0)$, and K.

Thank you for listening