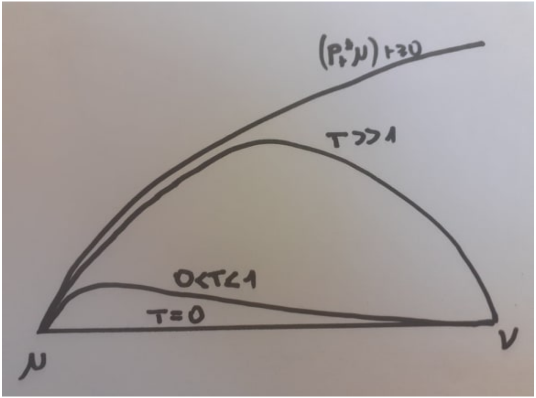
Long-time entropic interpolations

A work in collaboration with Giovanni Conforti and Ivan Gentil

Gauthier Clerc

Institut Camille Jordan

June 22, 2021



T-entropic interpolations from μ to ν for different values of T

Table of contents

- Let $(N, \mathfrak{g})$ be a smooth, complete and connected Riemannian manifold.
- We consider the generator $L = \Delta_{\mathfrak{g}}$ where $\Delta_{\mathfrak{g}}$ is the Laplace-Beltrami operator of $(N, \mathfrak{g})$.
- dx is the Riemannian measure on $(N, \mathfrak{g})$).
- $(P_t)_{t\geq 0}$ denotes the semigroup of L, that is for every smooth function f: $(P_t(f))_{t\geq 0}$ is the unique solution of

$$\begin{cases} \partial_t P_t(f) = LP_t(f), \; t > 0 \\ P_0(f) = f. \end{cases}$$

- For the rest of the presentation I will make the notation abuse: $\mu = \frac{d\mu}{dx}$.

Kantorovitch problem and Wasserstein distance

- We define

$$\mathcal{P}_2(N) := \{\mu \in \mathcal{P}(N) : \ \forall x_0 \in N \ \int d^2(x_0, x) d\mu(x) < \infty\}$$

- For $\mu, \nu \in P_2(N)$ let's define

$$\Pi(\mu, \nu) := \{\gamma \in \mathcal{P}(N \times N) : \text{ with } \mu \text{ and } \nu \text{ as marginals}\}.$$

and

$$W_2^2(\mu, \nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \int d(x, y)^2 d\gamma(x, y).$$

Benamou-Brenier formula

Theorem (Benamou-Brenier 2000)

Let $\mu, \nu \in \mathcal{P}_2(N)$ *then*

$$W_2^2(\mu, \nu) = \inf \int_0^1 \|v_t\|^2_{L^2(\mu_t)} \, dt,$$

where the infimum is taken over all smooth couples $(\mu_t, v_t)_{t\geq 0}$ *such that* $\forall t \in [0,1]$

$$\begin{cases} v_t \in L^2(\mu_t), \\ \partial_t \mu_t = -\nabla \cdot (\mu_t v_t), \\ \mu_0 = \mu, \; \mu_1 = \nu. \end{cases}$$

Existence of the velocity

Theorem (Ambrosio-Gigli-Savaré 2005)

Let $(\mu_t)_{t\in I} \subset \mathcal{P}_2(N)$ be a suitable curve in $\mathcal{P}_2(N)$. Then for every $t \in I$ there exists a unique vector field $V_t \in \overline{\{\nabla\phi : \ \phi \in C_c^\infty(N)\}}^{L^2(\mu_t)}$ such that

$$\partial_t \mu_t = -\nabla \cdot (\mu_t V_t).$$

Furthermore if v_t is another solution of the previous equation then

$$\|V_t\|_{L^2(\mu_t)} \leq \|v_t\|_{L^2(\mu_t)}.$$

Weak Riemannian structure of $\mathcal{P}_2(N)$

According to the previous formula, we can define

- For $\mu \in \mathcal{P}_2(N)$, $T_\mu \mathcal{P}_2(N) := \overline{\{\nabla\phi : \phi \in C_c^\infty(N)\}}^{L^2(\mu)}$.

- For $(\mu_t)_{t\geq 0} \subset \mathcal{P}_2(N)$ smooth enough, for all $t > 0$ $\dot{\mu_t} = V_t$ is the unique solution of $\partial_t \mu_t = -\nabla \cdot (\mu_t v_t)$ in $T_{\mu_t} M$.

- The Riemannian scalar product on $T_\mu \mathcal{P}_2(N)$ is the scalar product induced by the scalar product of $L^2(\mu)$ that is for $v, w \in T_\mu \mathcal{P}_2(N)$

$$\langle v, w \rangle_\mu = \int v\, w\, d\mu.$$

The fundamental functional $\mathcal{E}nt$

- For (suitable) $\mu \in \mathcal{P}_2(N)$ lets define

$$\mathcal{E}nt(\mu) = \int \log(\mu)\, d\mu.$$

- Let's $(\mu_t)_{t \geq 0}$ be a smooth curve in $\mathcal{P}_2(N)$. Then we can easily compute

$$\partial_t \mathcal{E}nt(\mu_t) = \langle \dot{\mu}_t, \nabla \log(\mu_t) \rangle_{\mu_t}.$$

- Hence we can define for $\mu \in \mathcal{P}_2(N)$

$$grad_\mu \mathcal{E}nt = \nabla \log(\mu).$$

Gradient flow equation of $\mathcal{E}nt$

- The gradient flow equation of $\mathcal{E}nt$ is

$$\dot{\mu}_t = -grad_{\mu_t}\mathcal{E}nt = -\nabla \log(\mu_t).$$

- Putting together the previous equation and the continuity equation $\partial_t \mu_t = -\nabla \cdot (\dot{\mu}_t \mu_t)$ we find that the gradient flow equation of $\mathcal{E}nt$ can be rewritten as

$$\partial_t \mu_t = L\mu_t.$$

- Hence the gradient flow semigroup of $\mathcal{E}nt$ is $(P_t)_{t\geq 0}$ the dual semigroup of L.

Gamma operator

- The Carré du champ operator associated with L is given by for all smooth function f and g

$$\Gamma(f,g) = \frac{1}{2}\left(L(fg) - fLg - gLf\right) = \nabla f \cdot \nabla g,$$

and the iterated Carré du champ is given by

$$\begin{aligned}\Gamma_2(f,f) &= \frac{1}{2}\left(L\Gamma(f,f) - 2\Gamma(Lf,f)\right) \\ &= |\nabla^2 f|_{HS}^2 + Ricci(\nabla f, \nabla f).\end{aligned}$$

Curvature dimension conditions

Definition

For $\rho \geq 0$ and $n \in (0, \infty]$ we say that L verify a $CD(\rho, n)$ condition if

$$\Gamma_2(f) \geq \frac{1}{n}(Lf)^2 + \rho\Gamma(f).$$

- $(\mathbb{R}^n, \Delta)$ verify the $CD(0, n)$ CDC.
- $(\mathbb{R}^n, \Delta - x \cdot \nabla)$ verify the $CD(1, \infty)$ CDC.
- $(S^{d-1}, \Delta_{\mathfrak{g}})$ verify the $CD(d-1, d)$ CDC.

Hessian of $\mathcal{E}nt$

- We can compute the Hessian of $\mathcal{E}nt$ by computing $\partial_t^2 \mathcal{E}nt(\mu_t)$ along a geodesic (μ_t) of speed $\dot{\mu}_t = \nabla\phi_t$, and we find

$$Hess_{\mu_t}\mathcal{E}nt(\nabla\phi_t, \nabla\phi_t) := \partial_t^2 \mathcal{E}nt(\mu_t) = \int \Gamma_2(\phi_t, \phi_t) d\mu_t.$$

Curvature dimension condition and Otto calculus

Definition

Let $\rho \geq 0$ and $n \in (0, +\infty]$. A function $\mathcal{F} : \mathcal{P}_2(N) \to \mathbb{R}$ is (ρ, n)-convex if for every $\mu \in \mathcal{P}_2(N)$ and $\dot{\mu} \in T_\mu \mathcal{P}_2(N)$,

$$Hess_\mu \mathcal{F}(\dot{\mu}, \dot{\mu}) \geq \rho |\dot{\mu}|^2_\mu + \frac{1}{n} \langle grad_\mu \mathcal{F}, \dot{\mu} \rangle^2_\mu.$$

Theorem

For $\rho \geq 0$ and $n \in (0, \infty]$ the following are equivalent

- *L verify a $CD(\rho, n)$ dimension condition.*
- *$dim(N) \leq n$ and $Ricci \geq \rho Id$.*
- *$\mathcal{E}nt$ is a (ρ, n)-convex functional.*

The Schrödinger problem

- For every $T > 0$ we define $\Omega_T = C\left([0,T], N\right)$ and R^T is the unique diffusion measure with generator L and dx as initial value.

- For every $T > 0$ and $\mu, \nu \in \mathcal{P}(N)$ we define the T-Schrödinger cost between μ and ν by

$$Sch_T(\mu,\nu) = \inf\left\{H(Q|R^T):\ Q \in \mathcal{P}(\Omega_T),\ P_0 = \mu\, P_T = \nu\right\}.$$

Benamou-Brenier-Schrödinger formula

Theorem (Gentil-Leonard-Ripani, Gigli-Tamanini, Chen-Georgiou-Pavon)

Let $\mu, \nu \in \mathcal{P}_2(N)$ be two compactly supported measures with bounded densities, then for any $T > 0$

$$Sch_T(\mu, \nu) = \frac{T}{4}\mathcal{C}_T(\mu, \nu) + \frac{T}{2}\left(\mathcal{E}nt(\mu) + \mathcal{E}nt(\nu)\right).$$

Where

$$\mathcal{C}_T(\mu, \nu) = \inf \int_0^T |\dot{\mu}_t|^2_{\mu_t} + |grad_{\mu_t}\mathcal{E}nt|^2_{\mu_t} dt,$$

and the infimum is taken over all smooth paths connecting μ to ν in $\mathcal{P}_2(N)$. Furthermore $\mathcal{C}_T(\mu, \nu)$ admit a unique minimiser called entropic interpolation from μ to ν.

Finite dimensional toy model

Definition

Let $F : \mathbb{R}^n \to \mathbb{R}$ be a smooth convex function. Then for all $T > 0$ we can define the F-cost between x and y by

$$C_T(x, y) = \inf \int_0^T |\dot{X}_s|^2 + |\nabla F(X_s)|^2 ds,$$

where the infimum is taken over all smooth paths from x to y.

The previous problem has exactly one minimizer denoted by $(X_t^T)_{t\in[0,T]}$ and called F-interpolation.

Gradient flow

Definition

We denote $(S_t)_{t\geq 0}$ the gradient flow semigroup of F, that is for every $x \in \mathbb{R}^n$: $(S_t(x))_{t\geq 0}$ is the unique solution of

$$\begin{cases} \dot{S}_t(x) = -\nabla F(S_t(x)),\ t > 0, \\ S_0(x) = x. \end{cases}$$

Convergence of F-interpolation

Theorem (Clerc-Conforti-Gentil)

- *If there exists $\rho > 0$ such that F is ρ-convex. Then there exists a constant $C > 0$ (depending only on x, y and ρ) such that for every $t \in [0,1]$ and $T > 1$*

$$|X_t^T - S_t(x)| \leq e^{-\rho T} C.$$

- *If there exist $n > 0$ such that F is $(0,n)$-convex. Then there exists a constant $C > 0$ (depending only on x and y) such that for every $t \in [0,1]$ and $T > 1$*

$$|X_t^T - S_t(x)| \leq C \sqrt{\frac{n \log(T)}{T}}.$$

Sketch of proof for the ρ-convex case

- The F-interpolations satisfy the Newton equation

$$\ddot{X}_t^T = \nabla^2 F(X_t^T)\nabla F(X_t^T).$$

- Recall that Giovanni Conforti and Luca Tamanini had shown that for all $x, y \in \mathbb{R}^n$ there exists a constant $C > 0$ which does not depend on T such that $C_T(x, y) \leq C$.
- For every $0 \leq t \leq T$

$$\begin{aligned}
\partial_t |X_t^T - S_t(x)|^2 &= 2\langle \dot{X}_t^T + \nabla F(S_t(x)), X_t^T - S_t(x)\rangle \\
&= -\langle \nabla F(X_t^T) - \nabla F(S_t(x)), X_t^T - S_t(x)\rangle \\
&+ \langle \dot{X}_t^T + \nabla F(X_t^T), X_t^T - S_t(x)\rangle \\
&\leq |X_t^T - S_t(x)||\dot{X}_t^T + \nabla F(X_t^T)|.
\end{aligned}$$

Sketch of proof for the ρ-convex case

- We define $\varphi_t^T := \dot{X}_t^T + \nabla F(X_t^T)$. Then

$$\begin{aligned}\partial_t|\varphi_t^T|^2 &= 2\langle \partial_t \varphi_t^T, \varphi_t^T\rangle \\ &= 2\nabla^2 F(X_t^T)(\varphi_t^T, \varphi_t^T) \\ &\geq 2\rho|\varphi_t^T|^2.\end{aligned}$$

Whence for every $t \leq s \leq T$ we have $\exp(2\rho(s-t))|\varphi_t^T|^2 \leq |\varphi_s^T|^2$ and we have

$$\begin{aligned}\frac{\exp(2\rho(T-t))-1}{2\rho}|\varphi_t^T|^2 &\leq \int_t^T |\varphi_s^T|^2 ds \\ &\leq \int_0^T \left(|\dot{X}_s^T|^2 + |\nabla F(X_s^T)|^2 + 2\langle \dot{X}_s^T, \nabla F(X_s^T)\rangle\right) ds \\ &= C_T(x,y) + 2(F(y) - F(x)).\end{aligned}$$

Sketch of proof for the $(0, n)$-convex case

- We introduce the conserved quantity $E_T(x, y) = |\dot{X}_t^T|^2 - |\nabla F(X_t^T)|^2$. That is well defined thanks to the Newton equation

$$\ddot{X}_t^T = \nabla^2 F(X_t^T) \nabla F(X_t^T)$$

.

- It can be shown that the map

$$[0, \infty) \ni t \to \exp\left(-\frac{1}{n}\left[F(X_t^T) + \int_0^t |\nabla F(X_s^T)|^2 ds\right]\right)$$

is concave. From this concavity property and an "inégalité des pentes", we can deduce that

$$-E_T(x, y) \leq \frac{2n}{T}.$$

Sketch of proof for the $(0, n)$-convex case

- Giovanni Conforti and Luca Tamanini had shown that $T \mapsto C_T(x, y)$ is derivable and for all $T > 0$

$$\frac{d}{dT} C_T(x, y) = -E_T(x, y).$$

Whence by integration of this equality and the previous inequality

$$C_T(x, y) \leq C_1(x, y) + 2n \log(T).$$

- Thanks to the convexity of F the map $t \mapsto |\varphi_t^T|^2$ is non decreasing and

$$(T - t)|\varphi_t^T|^2 \leq \int_t^T |\varphi_s^T|^2 ds \leq C_T(x, y) + 2(F(y) - F(x)).$$

Sketch of proof for the $(0, n)$-convex case

Whence

$$|\varphi_t^T|^2 \leq \frac{2(F(y) - F(x)) + C_1(x, y) + 2n\log(T)}{T - t}.$$

The final result can be obtain from this bound as in the ρ-convex case.

Convergence of entropic interpolations

Theorem (Clerc-Conforti-Gentil)

Let $\mu, \nu \in \mathcal{P}_2(N)$ be two compactly supported absolutely continuous measures with smooth densities against m.

- *If L verify a $CD(\rho, \infty)$ curvature dimension condition with $\rho > 0$ then there exists a constant $C > 0$ such that for every $t \in [0,1]$ and $T > 1$,*

$$W_2^2(\mu_t^T, P_t^*(\mu)) \leq Ce^{-\rho T}.$$

- *If L verify a $CD(0, n)$ curvature dimension condition with $n > 0$ then there exists a constant $C > 0$ such that for every $t \in [0,1]$ and $T > 1$,*

$$W_2^2(\mu_t^T, P_t^*(\mu)) \leq C\frac{n\log(T)}{T}.$$