

6. Existence of W_c geodesics

We will use the splitting result Theorem 2, and discuss optimal curves for $W_{0, \text{id}}$ and D_R separately.

6.1 Optimal Interpolation of Moments

Recall the moment part of the kinetic energy

$$I(m, C) := \int_0^1 \frac{1}{2} \|\dot{m}_t\|_{C_t}^2 + \frac{1}{8} \text{Tr}[\dot{C}_t \bar{C}_t^{-1} \dot{C}_t C_t^{-1}] dt$$

and the definitions

$$D_R(\mu_0, \mu_1)^2 = \inf_{(m, C)} \left\{ I(m, C) : \begin{array}{l} (m_0, C_0) = (M(\mu_0), C(\mu_0)), \\ (m_1, C_1) = (M(\mu_1), C(\mu_1)), \\ \bar{C}_1^{-1/2} A_1 = R \end{array} \right\} \text{ and}$$

$$D(\mu_0, \mu_1)^2 = \inf_{(m, C)} \left\{ I(m, C) : \begin{array}{l} (m_0, C_0) = (M(\mu_0), C(\mu_0)), \\ (m_1, C_1) = (M(\mu_1), C(\mu_1)) \end{array} \right\} = \inf_{R \in \text{SO}(d)} D_R(\mu_0, \mu_1)^2$$

Proposition 6 | (Existence)

For any $C_0, C_1 \in S_{\geq 0}^d$ and $m_0, m_1 \in \mathbb{R}^d$ with $m_1 - m_0 \in \text{Im } C_0 = \text{Im } C_1$, there exists an optimal interpolation $(m_t, C_t)_{t \in [0,1]}$ for both D and D_R .

Proof

Existence for D (Sketch): $I(m, C) = \int_0^1 L[\dot{m}_t, \dot{C}_t] dt$ with L quadratic in $(\dot{m}, \dot{C})$, continuous and bounded below. Lagrangian hence I is weakly lsc

→ use Direct Method. Only need to show (C_t^{-1}) stays in compact set: follows from Lemma 2.10 in [Burger et al.]

Existence for D_R (Sketch): Rewrite as an optimization problem for (m_t, A_t) where $A_t A_t^T = C_t$ and A_t solves (3). Interpret it as existence of geodesics on $M := \mathbb{R}^d \times GL_+(d)$ with the sub-Riemannian metric $\langle (m_0, A_0), (m_1, A_1) \rangle = m_0 \cdot m_1 + A_0 \cdot A_1$ with a suitable definition of admissible tangential vectors. M is non-compact. Relative compactness follows from a stability estimate for curves (m_t, A_t) with finite action, and so we obtain a limit for minimizing sequences.

Proposition 7 | (Optimality Conditions for D)

Optimal curves for D solve

$$(4a) \quad C_t^{-1} \dot{m}_t = \alpha$$

$$(4b) \quad \ddot{C}_t = \dot{C}_t C_t^{-1} \dot{C}_t - 2 C_t (\alpha \otimes \alpha) C_t$$

for some constant $\alpha \in \mathbb{R}^d$.

Proof: For any $n \in AC([0,1], \mathbb{R}^d)$ with $n(0)=n(1)=0$ and any $K \in AC([0,1], \mathbb{R}_{sym}^{d \times d})$ with $K(0)=K(1)=0$, we have for any (m, C) minimizing curve for $I(m, C)$ (dropping t from the notation):

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon} I(m + \varepsilon n, C + \varepsilon K) \\ &= \int_0^1 \left(\langle \dot{m}, \bar{C} \dot{m} \rangle - \frac{1}{2} \langle \dot{m}, \bar{C}' K \bar{C}' \dot{m} \rangle + \frac{1}{4} \text{Tr} [\dot{K} \bar{C}' \dot{C} \bar{C}'] - \frac{1}{4} \text{Tr} [K \bar{C}' \dot{C} \bar{C}' \dot{C} \bar{C}'] \right) dt \\ &= \int_0^1 \left(\langle -\frac{d}{dt}(\bar{C} \dot{m}), n \rangle - \frac{1}{2} \bar{C}'(\dot{m} \otimes \dot{m}) \bar{C}' : K - \frac{1}{4} \frac{d}{dt}[\bar{C} \dot{C} \bar{C}'] : K - \frac{1}{4} \bar{C} \dot{C} \bar{C}' \dot{C} \bar{C}' : K \right) dt \\ &= \int_0^1 \langle -\frac{d}{dt}(\bar{C} \dot{m}), n \rangle dt + \int_0^1 \frac{1}{4} [\bar{C}' \dot{C} \bar{C}' \dot{C} \bar{C}' - \bar{C}' \ddot{C} \bar{C}' - 2 \bar{C}'(\dot{m} \otimes \dot{m}) \bar{C}'] : K dt \end{aligned}$$

hence (m, C) solve (4). $\square$

type in Trop 2.14 in [Burger et al]

$\frac{d}{dt}(\bar{C} \dot{C} \bar{C}') = -2 \bar{C}' \dot{C} \bar{C}' \dot{C} \bar{C}' + \bar{C}' \ddot{C} \bar{C}'$

Exercise The action density $t \mapsto \frac{1}{2} \|\dot{m}_t\|_{C_t}^2 + \frac{1}{8} \text{Tr} [\dot{C}_t \bar{C}_t' \dot{C}_t C_t^{-1}]$ is constant on $[0,1]$ for any $(m_t, C_t)_{t \in [0,1]}$ solving the optimality conditions (4).

Remark For $\alpha = 0$, the geodesic equations (4) provide a Riemannian distance on $S_{>0}^d$.
The case $\alpha \neq 0$ gives rise to new effects.

Proposition 8 | (Optimality Conditions for $\mathcal{D}_R$)

For any optimal curve (m_t, C_t) , define A_t as the solution to (3).

Then there exists $\alpha \in \mathbb{R}^d$ and $Q \in \mathbb{R}^{d \times d}$ skew-symmetric, such that

$$A_t^{-T} A_t^{-1} \dot{m}_t = \alpha \quad (5a)$$

$$\frac{d}{dt} (A_t^{-1} \dot{A}_t) = [A_t^{-1} \dot{A}_t, Q] - (A_t^T \alpha) \otimes (A_t^T \alpha) \quad (5b)$$

The solution satisfies $A_0 = C_0^{\vee/2}$, $A_1 = C_1^{\vee/2} R$.

Proof See Section 2.3 in [Burger et al]

Remark: If $m_1 = m_0$, then

• optimizers for $\mathcal{D}_R$: $m_t = m_0$, $A_t = A_0 e^{tB} e^{-tB_{asym}} \forall t \in [0,1]$, with B an optimizer of $\frac{1}{2} \inf_{B \in \mathbb{R}^{d \times d}} \left\{ \|B_{sym}\|_{HS}^2 : A_0^{-1} A_1 = e^B e^{-B_{asym}} \right\}$ with $B_{sym} := \frac{B+B^T}{2}$, $B_{asym} := \frac{B-B^T}{2}$.

• optimizers for $\mathcal{D}$: $m_t = m_0$, $C_t = C_0^{\vee/2} (C_0^{-\vee/2} C_1 C_0^{-\vee/2})^t C_0^{\vee/2} \forall t \in [0,1]$

See Lemma 2.13 and Corollary 2.15 in [Burger et al]

Remark For Gaussians $\mu_0 = N(m_0, C)$ and $\mu_1 = N(m_1, C)$ with the same covariance, denote $\xi = m_0 - m_1$. Then

$$W_c(\mu_0, \mu_1) \sim \log \|\xi\| \text{ for } \|\xi\| \gg 1$$

whereas

$$W_2(\mu_0, \mu_1) = \|\xi\|$$

6.2 Existence of Shape Geodesics

We show two partial results: (1) existence under small distance

(2) existence under symmetry

Theorem 9 | (Existence at small distance)

Any $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ with $M(\mu_0) = M(\mu_1) = 0$ and $C(\mu_0) = C(\mu_1) = Id$ with $W_{0, Id}(\mu_0, \mu_1)^2 < \frac{1}{8}$ are connected via a $W_{0, Id}$ -geodesic.

Proof Strategy:

- Consider a minimizing sequence (μ^n, V^n) for $W_{0, Id}(\mu_0, \mu_1)$.
- Show relative compactness of (μ^n, V^n) in weak topologies by the same arguments as for W_2 .
- Extract and identify a limiting pair (μ, V) solving $\partial_t \mu + \operatorname{div}(\mu V) = 0$.
- Challenge: Are the moment constraints preserved in the limit?

$$M(\mu^n) = 0, C(\mu^n) = Id \Rightarrow M(\mu) = 0, C(\mu) = Id ??$$

For this, tightness of second moments is sufficient. But this doesn't come for free.

- Show tightness of $\{|x|^2 d\mu_t^n\}$ via contradiction, using $W_{0, Id}(\mu_0, \mu_1)^2 < \frac{1}{8}$.

(Key challenge! See Lemma 10 below)

- Tightness of $\{|x|^2 d\mu_t^n\}_{n \in \mathbb{N}}, \forall t \in [0, 1] \Rightarrow \mu_t \in \mathcal{P}_{0, Id}(\mathbb{R}^d) \forall t \in [0, 1]$
- Show that limit (μ, V) achieves minimal action using joint l.s.c. of $(\mu, V) \mapsto \int |V|^2 d\mu$

Example ($d=1$): $\mu^n = (1 - \frac{1}{n}) \delta_0 + \frac{1}{n} \delta_{\frac{1}{n}}$. Then $M(\mu^n) = \frac{1}{n^2} \rightarrow 0$, $C(\mu^n) = \int |x - \frac{1}{n}|^2 d\mu^n = (1 - \frac{1}{n}) \frac{1}{n} + \frac{1}{n} \left| \frac{n-1}{n} \right|^2 = \frac{1}{n} - \frac{1}{n^2} + \frac{1}{n^2} (n^2 - 2n + 1) = 1 - \frac{1}{n} \rightarrow 1$
But $\mu^n \rightarrow \delta_0$ narrowly and $C(\delta_0) = 0$.

Lemma 10 (Key ingredient for tightness) For $(\mu, \nu) \in CE_{0, Id}(\mu_0, \mu_1) = \left\{ (\mu, \nu) : \begin{array}{l} \partial_t \mu + \operatorname{div}(\mu, \nu) = 0, \quad M(\mu_t) = 0 \\ \mu|_{t=0} = \mu_0, \mu|_{t=1} = \mu_1, \quad C(\mu_t) = Id \end{array} \right\}$
 write $A(\mu, \nu) := \frac{1}{2} \int \int \|V_t\|^2 d\mu_t dt$.

Assume $\mu_0^n, \mu_1^n \in \mathcal{P}_{0, Id}(\mathbb{R}^d)$. Let $(\mu^n, \nu^n) \in CE_{0, Id}(\mu_0^n, \mu_1^n)$ s.t. $\lim_n A(\mu^n, \nu^n) < \frac{1}{8}$
 and $\mu_t^n \rightarrow \mu_t$ weakly $\forall t \in [0, 1]$ for a curve $(\mu_t)_{t \in [0, 1]} \in \mathcal{P}_*(\mathbb{R}^d)$ with $\mu_0, \mu_1 \in \mathcal{P}_{0, Id}(\mathbb{R}^d)$.
 If $\exists t_0 \in (0, 1)$ s.t. $\mu_{t_0} \notin \mathcal{P}_{0, Id}(\mathbb{R}^d)$, then $\exists (\tilde{\mu}^n, \tilde{\nu}^n) \in CE_{0, Id}(\mu_0^n, \mu_1^n)$ s.t.

$$\liminf_{n \rightarrow \infty} A(\tilde{\mu}^n, \tilde{\nu}^n) < \lim_{n \rightarrow \infty} A(\mu^n, \nu^n)$$

Proof Strategy: • $M(\mu_{t_0}) = \lim_{n \rightarrow \infty} M(\mu_{t_0}^n) = 0$. Hence $C(\mu_{t_0}) \neq Id = \lim_{n \rightarrow \infty} C(\mu_{t_0}^n)$.

- This means tightness of $\{ |x|^2 d\mu_{t_0}^n \}_{n \geq 1}$ cannot hold (else $C(\mu_{t_0}) = Id$).
- Construct competitor $(\tilde{\mu}^n, \tilde{\nu}^n)$ by rerouting mass that leaves a large ball and normalizing the resulting measures
- Rerouting decreases the length for a fraction of curves, but potentially decreases the covariance.
- Carefully estimate those competing effects.
- $W_{0, I}(\mu_0, \mu_1) < \frac{1}{8} \Rightarrow$ competitor has smaller action.

Thanks to Theorem 9, we can construct geodesics of W_c at small distance.

Theorem 11 Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ with $C(\mu_0), C(\mu_1) > 0$ and $W_c(\mu_0, \mu_1)^2 < \frac{1}{8} + D(\mu_0, \mu_1)$.
 Then μ_0, μ_1 are connected by a W_c -geodesic.

Proof To show this, use the splitting result Theorem 2

$$W_c(\mu_0, \mu_1)^2 = \inf_{R \in SO(d)} \left\{ W_{0, Id}(R_{\#} \bar{\mu}_0, \bar{\mu}_1)^2 + D_R(\mu_0, \mu_1) \right\}. \quad (6)$$

Take a minimizing sequence $R^n \in SO(d)$.

For each R^n , by Prop 6 and Thm 9, there exist optimal $(\bar{\mu}^n, \bar{\nu}^n) \in CE_{0, Id}(R_{\#} \bar{\mu}_0, \bar{\mu}_1)$
 and $(m^n, C^n) \in AC([0, 1], \mathbb{R}^d \times S_{\geq 0}^d)$ with $m_i^n = M(\mu_i)$, $C_i^n = C(\mu_i)$ for $i=0, 1$, and
 $R^n = (C_1^n)^{-1/2} A_1^n$ such that $W_c(\mu_0, \mu_1)^2 = \lim_{n \rightarrow \infty} A(\bar{\mu}^n, \bar{\nu}^n) + I(m^n, C^n)$

By compactness of $SO(d)$, $\exists$ limit $R \in SO(d)$ s.t. $R^n \rightarrow R$. Using similar arguments to the proofs of Prop 6 and Thm 9, $(m^n, C^n) \rightarrow (m, C)$ with $R = C_1^{-1/2} A_1$,
 $(\bar{\mu}^n, \bar{\nu}^n) \rightarrow (\bar{\mu}, \bar{\nu})$ with $(\bar{\mu}, \bar{\nu}) \in CE(R_{\#} \bar{\mu}_0, \bar{\mu}_1)$

For a subsequence, $\lim_{n \rightarrow \infty} A(\bar{\mu}^n, \bar{\nu}^n) = \liminf_{n \rightarrow \infty} A(\bar{\mu}^n, \bar{\nu}^n)$

Then since $I(m^n, C^n) \geq D(\mu_0, \mu_1)$, we have $\lim_n A(\bar{\mu}^n, \bar{\nu}^n) \leq \frac{1}{8}$,

we can apply Lemma 10 to conclude $(\bar{\mu}, \bar{\nu}) \in (E_{0, \text{Id}}(R_{\#} \bar{\mu}_0, \bar{\mu}_1))$.

By l.s.c. of $A(\cdot, \cdot)$ and $I(\cdot, \cdot)$, $A(\bar{\mu}, \bar{\nu}) + I(m, C) \leq \liminf_{n \rightarrow \infty} A(\bar{\mu}^n, \bar{\nu}^n) + I(m^n, C^n) = W_c(\mu_0, \mu_1)^2$

Hence, $(R, \bar{\mu}, \bar{\nu}, m, C)$ is a minimizer for the optimization problem (6).

Define A_t as a solution to (3). Then $\mu_t := (A_t \cdot + m_t)_{\#} \bar{\mu}_t$

is a W_c -geodesic between μ_0 and μ_1 . $\square$

Theorem 12 (Existence under symmetry) Let $\bar{\mu}_0, \bar{\mu}_1 \in \mathcal{P}_{0, \text{Id}}(\mathbb{R}^d)$, $\bar{\mu}_i \ll \mathcal{L}^d$ and with d -fold reflection symmetry. Then $\bar{\mu}_0, \bar{\mu}_1$ are connected by a $W_{0, \text{Id}}$ -geodesic.

The proof uses the dual formulation of $W_{0, \text{Id}}$, which we discuss in the next section.

7. Dual Formulation of W_c

12:00

Theorem 13 (Dual formulation)

$$W_{0, \text{Id}}(\mu_0, \mu_1)^2 = \sup_{\psi, \alpha, \Lambda} \int \psi_1 d\mu_1 - \int \psi_0 d\mu_0 + \int_0^1 \text{Tr} \Lambda_t dt$$

where the sup is taken over functions $\psi: [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$, $\Lambda: [0, 1] \rightarrow S^d$, $\alpha: [0, 1] \rightarrow \mathbb{R}^d$ subject to the modified Hamilton-Jacobi subsolution constraint

$$\partial_t \psi + \frac{1}{2} |\nabla \psi|^2 + \text{Tr}[\Lambda(x \otimes x)] - \langle \alpha, x \rangle \leq 0.$$

Remark: The functions α and Λ act as Lagrange multipliers for the constraints $M(\mu_t) = 0$ and $C(\mu_t) = \text{Id}$

The mean constraint is not active at the optimum since $m_0 = m_1 = 0 \Rightarrow m_t = 0 \forall t \in [0, 1]$

Theorem 14 (Optimality Conditions for $W_{0, \text{Id}}$)

$$\partial_t \mu + \text{div}(\mu \nabla \psi) = 0,$$

$$\alpha = 0,$$

$$\partial_t \psi + \frac{1}{2} |\nabla \psi|^2 + \text{Tr}[\Lambda x \otimes x] = 0, \quad \Lambda = \frac{1}{2} \int \nabla \psi \otimes \nabla \psi d\mu.$$

Remark In particular, $\text{Tr } \Lambda_t = \frac{1}{2} \int |\nabla \psi_t|^2 d\mu_t = W_{0, \text{Id}}(\mu_0, \mu_1)^2$ is constant in time.

For the proof of Theorem 13 and 14, see Section 5.1 in [Burger et al]

Using duality for the proof of Theorem 12:

let $\gamma: [0,1] \rightarrow \mathbb{R}^d$ be the mass particle trajectories. Assume $\gamma \sim M \in \mathcal{M}_+(H^1([0,1], \mathbb{R}^d))$

Rewrite covariance-constrained OT as

$$\min_M \mathbb{E}_M \left[\int |\dot{\gamma}_t|^2 dt \right] \text{ subject to } \begin{aligned} \text{Law}_M(\gamma_0) &= \bar{\mu}_0, & \mathbb{E}_M[\dot{\gamma}_t] &= 0, & \text{for a.e. } t \in [0,1] \\ \text{Law}_M(\gamma_1) &= \bar{\mu}_1, & \mathbb{E}_M[\dot{\gamma}_t \otimes \dot{\gamma}_t] &= \text{Id}. \end{aligned} \quad (7)$$

By explicit construction, one can show existence of $\bar{M}$ solving (7) s.t.

$\bar{M}$ -a.e. trajectory $\gamma(t)$ emerging from $\gamma(0) = x \in \mathbb{R}^d$ is given by

$$\gamma_k(t) = \frac{\sin(w_k(1-t))}{\sin w_k} x_k + \frac{\sin(w_k t)}{\sin w_k} S_k(x) \quad \text{for } k=1, \dots, d$$

← mass transport by componentwise dilation

for some parameters $w_1, \dots, w_d \in [0, \frac{\pi}{2}]$ and map $S: \mathbb{R}^d \rightarrow \mathbb{R}^d$.

$$\text{Then } W_{0, \text{Id}}(\bar{\mu}_0, \bar{\mu}_1)^2 = W_2(G_{\#}^{\omega} \bar{\mu}_0, G_{\#}^{\omega} \bar{\mu}_1) - 2 \sum_{k=1}^d \frac{w_k}{\sin w_k} \left(1 - \cos w_k - \frac{1}{2} w_k \sin w_k\right)$$

$$\text{with } [G^{\omega}(x)]_k = \left(\frac{w_k}{\sin w_k}\right)^{\frac{1}{2}} x_k \quad \leftarrow \text{componentwise linear dilation}$$

(see Theorem 3.2 in [Burger et al] for details)

Remarks

- Thanks to the symmetry assumption of the d^2 constraints in $C(\mu_t) = \text{Id}$, only d survive.
- Idea of proof for Theorem 12: write dual formulation for (7), and notice that the optimal Λ_t is diagonal thanks to the symmetry assumption.

Thus, the optimality constraint involving Λ_t decouples into d 1-dimensional problems that can each be solved explicitly

Existence of suitable $w_1, \dots, w_d$ s.t. $\mathbb{E}_{\bar{M}}[\gamma_k(t)^2] = 1$ is obtained via a fixed point argument

Define $S = (G^{\omega})^{-1} \circ T^{\omega} \circ G^{\omega}$, where T^{ω} is the W_2 optimal OT map

from $G_{\#}^{\omega} \bar{\mu}_0$ to $G_{\#}^{\omega} \bar{\mu}_1$. ← $\bar{\mu}_i \ll \mathcal{L}^d$ guarantees T^{ω} exists

Open Problem: Is it possible to remove the symmetry assumptions in Theorem 12 and still obtain a $W_{0,Id}$ -geodesic via a duality formulation?

→ in principle should be possible

→ the optimal Λ_t would be symmetric positive semi-definite matrix, but not necessarily diagonal. In particular, if such Λ_t exists, it solves $\ddot{\gamma}(t) + \Lambda_t \dot{\gamma}(t) = 0$, for given $\gamma(0), \gamma(1)$.

8. Convergence to Equilibrium in W_c gradient flows

12:20

Let $\pi = e^{-J(x)}$, $J(x) = \frac{1}{2} \|x - x_0\|_B^2$, $B \in \mathbb{S}_{>0}^d$

$$E(\rho | \pi) = \int \rho \log \frac{\rho}{\pi} = H(\rho) + \int J(x) d\rho(x), \quad H(\rho) = \int \rho \log \rho$$

Gradient flow of $E(\rho | \pi)$ in W_c :

$$\begin{aligned} \partial_t \rho &= \operatorname{div}(\rho C(\rho) \nabla \log \frac{\rho}{\pi}) \\ &= \operatorname{div}(\rho C(\rho) \nabla \rho + \rho C(\rho) \bar{B}'(x - x_0)) \end{aligned} \quad (8)$$

For ρ_t solving (8),

let $m_t = M(\rho_t)$, $C_t = C(\rho_t)$,

and define A_t as the solution of:

$$\dot{A}_t = \frac{1}{2} \dot{C}_t A_t^{-T}, \quad A_0 A_0^T = C_0 \quad (9)$$

degree of freedom
how to choose
initial A_0

Exercise: (m_t, C_t) solve $\dot{m}_t = -C_t \bar{B}'(m_t - x_0)$

$$\dot{C}_t = 2C_t - 2C_t \bar{B}' C_t$$

Recall $T_{m_t, A_t}(x) = A_t^{-1}(x - m_t)$.

The solutions are given by $m_t = x_0 + e^{-t} A_t A_0^{-1} (m_0 - x_0)$

$$C_t^{-1} = (1 - e^{-2t}) \bar{B}' + e^{-2t} C_0^{-1}$$

Lemma 15

The normalized solution with $\eta_t := T_{m_t, A_t} \# \rho_t$

solves $\partial_t \eta_t = \Delta \eta_t + \operatorname{div}(x \eta_t)$ (10)

In other words, $\eta_t = \operatorname{Law}(X_t)$, where X_t is the Ornstein-Uhlenbeck process

$$dX_t = -X_t dt + \sqrt{2} dB_t$$

↳ This is a key insight: (10) is a Fokker-Planck equation with quadratic potential, an extremely well-studied equation. Using T_{m_t, A_t} , we can translate all the well-known properties of (10) to the W_c -GF (8).

Proof Lemma 15: Let $|A| = \det A$. Then $\eta_t(x) = |A_t| \rho_t(A_t x + m_t)$

$$\Rightarrow \dot{\eta}(x) = |A| \left[\dot{\rho}(Ax+m) + \langle \nabla \rho(Ax+m), \dot{A}x + \dot{m} \rangle + \rho(Ax+m) \frac{d}{dt} \log |A| \right]$$

Compute $\nabla \eta(x) = A^T \nabla \rho(Ax+m) |A|$,

$$\Delta \eta(x) = (A^T)_i^j \partial_{j_i} \rho(Ax+m) |A|$$

$$= \operatorname{div}(C \nabla \rho)(Ax+m) |A|,$$

$$\operatorname{div}[\eta(x) \nabla \frac{1}{|x|^2}] = \langle \nabla \eta(x), x \rangle + d \eta(x)$$

First term above: sub into (8),

$$|A| \dot{\rho}(Ax+m) = |A| \operatorname{div}(\rho C \nabla \rho)(Ax+m) + |A| \operatorname{div}(\rho C \nabla J)(Ax+m)$$

$$= |A| \operatorname{div}(\rho C \nabla \rho)(Ax+m) + |A| \langle \nabla \rho(Ax+m), C \bar{B}'(Ax+m-x_0) \rangle + |A| \rho(Ax+m) \operatorname{Tr}[C \bar{B}']$$

$$= \Delta \eta(x) + \langle \nabla \eta(x), A^T \bar{B}'(Ax+m-x_0) \rangle + \eta(x) \operatorname{Tr}[C \bar{B}']$$

By evolution $\dot{A}$, we have $\dot{A} \bar{A} = \frac{1}{2} \dot{A} \bar{A} \bar{A}^T = \operatorname{Id} - A^T \bar{B}' A \Rightarrow \frac{d}{dt} \log |A| = \operatorname{Tr}[\dot{A} \bar{A}] = d - \operatorname{Tr}[A^T \bar{B}' A] = d - \operatorname{Tr}[C \bar{B}']$

And $\dot{A} \bar{m} = -A^T \bar{B}'(m-x_0) \Rightarrow |A| \langle \nabla \rho(Ax+m), \dot{A}x + \dot{m} \rangle = \langle \nabla \eta(x), x - A^T \bar{B}'(Ax+m-x_0) \rangle$

Combining everything: $\dot{\eta}(x) = \Delta \eta(x) + \langle \nabla \eta(x), x \rangle + d \eta(x) = \Delta \eta(x) + \operatorname{div}(\eta(x)x)$. $\square$

Proposition 16 (Entropy Splitting) For any $\rho \in \mathcal{P}_2$, A with $AA^T = C = C(\rho)$ and $m = M(\rho)$, for any $x_0 \in \mathbb{R}^d$ and $B \in \mathcal{S}_{>0}^d$,

we have $E(\rho | N(x_0, B)) = E(\eta | N(0, \operatorname{Id})) + E(N(m, C) | N(x_0, B))$ where $\eta = T_{m,A} \# \rho$ and

$$I_{\operatorname{cov}}(\rho | N(x_0, B)) = \underbrace{I_{\operatorname{cov}}(\eta | N(0, \operatorname{Id}))}_{\text{shape (non-Gaussian part)}} + \underbrace{I_{\operatorname{cov}}(N(m, C) | N(x_0, B))}_{\text{moments (Gaussian part)}} \quad I_{\operatorname{cov}}(\rho | \pi) := \int \langle \nabla \log \frac{\rho}{\pi}, C(\rho) \nabla \log \frac{\rho}{\pi} \rangle d\rho$$

Proof See Lemma 4.3 in [Burger et al]

Follows by explicit computation: $\log \frac{\rho}{N(x_0, B)} = \log \frac{\rho}{N(m, C)} + \log \frac{N(m, C)}{N(x_0, B)}$ and $\eta = T_{m,A} \# \rho$, $N(0, \operatorname{Id}) = T_{m,A} \# N(m, C)$
 $\Rightarrow \int \log \frac{\rho}{N(m, C)} d\rho = \int \log \left(\left(\frac{\eta}{N(0, \operatorname{Id})} \right) \circ T_{m,A} \right) d\rho = \int \log \left(\frac{\eta}{N(0, \operatorname{Id})} \right) d\eta = E(\eta | N(0, \operatorname{Id}))$ etc $\square$

If A_t solves (9), then η_t solves (10) by Lemma 15, and so $E(\eta_t | N(0, \operatorname{Id})) \leq e^{-2t} E(\eta_0 | N(0, \operatorname{Id}))$

by classical results.

$$I(\eta_t | N(0, \operatorname{Id})) \leq e^{-2t} I(\eta_0 | N(0, \operatorname{Id})) \leftarrow \text{note that } C(\eta_t) = \operatorname{Id} \forall t.$$

So it remains to control the relaxation of Gaussians.

This can be done explicitly using solution (m_t, C_t) to the moment equations above, and noting that

$$E(N(m, C) | N(x_0, B)) = -\frac{1}{2} \left(\log \det(\bar{B}'C) + \operatorname{Tr}[\operatorname{Id} - \bar{B}'C] - |\bar{B}^{-\frac{1}{2}}(m-x_0)|^2 \right)$$

$$I_{\operatorname{cov}}(N(m, C) | N(x_0, B)) = \|\operatorname{Id} - \bar{B}'C\|_{\operatorname{HS}}^2 + |C^{\frac{1}{2}} \bar{B}'(m-x_0)|^2$$

where $\|M\|_{\operatorname{HS}}^2 := \operatorname{Tr}[MM^T]$ is the Hilbert-Schmidt norm (=Frobenius norm)

Theorem 17 (Entropy and Fisher information decay)

Solutions ρ_t to (8) satisfy $E(\rho_t | \pi) \leq k e^{-2t} E(\rho_0 | \pi)$

$$I_{\operatorname{cov}}(\rho_t | \pi) \leq k^2 e^{-2t} I_{\operatorname{cov}}(\rho_0 | \pi)$$

where $k = k(B, C_0) = \max\{1, \|\bar{B}^{\frac{1}{2}} C_0^{-1} \bar{B}^{\frac{1}{2}}\|_2\} \max\{1, \|\bar{B}^{\frac{1}{2}} C_0 \bar{B}^{\frac{1}{2}}\|_2\}$

$\leftarrow$ multiplicative relative condition number

Remark If $M(\rho_0) = x_0$ and $C(\rho_0) \geq B$, then

$$E(\rho_t | \pi) \leq e^{-2t} E(\rho_0 | \pi)$$

$$I_{\text{cov}}(\rho_t | \pi) \leq e^{-2t} I_{\text{cov}}(\rho_0 | \pi)$$

No multiplicative constant!

Corollary 18 $W_2(\rho_t, \pi) \leq e^{-t} \tilde{k} \left(\inf_{R \in \text{SOLA}} W_2(R_{\#} \bar{\rho}_0, N(0, Id))^2 + \underbrace{|m_0 - x_0|_{C_0}^2 + \|Id - B^{1/2} C_0^{-1} B^{1/2}\|_{HS}^2}_{= W_{2, C_0}^2(N(m_0, C_0), N(x_0, B))} \right)^{1/2}$
 where $\tilde{k} = \|B\|_2 \max\{1, \|B^{-1/2} C_0 B^{-1/2}\|_2\}$

Remarks

- replacing W_2 with W_{2, C_0} on the LHS, then get the same estimate with k instead of $\tilde{k}$
 $\Rightarrow k$ is a universal constant for $E(\cdot | \pi)$, $I_{\text{cov}}(\cdot | \pi)$ and $W_{2, C_0}(\cdot, \pi)$.
- Rate 1 was already shown for W_2 in [Carrillo, Vaes 2021], but with a different factor k that is very non-optimal and can become quite large.
- Rate 1 is optimal [Carrillo, Vaes 2021]

Proof (Sketch): (builds on [Carrillo, Vaes 2021])

- Splitting of W_2 : $W_2(\rho_t, \pi) \leq \|C_t\|_2^{1/2} W_2(\eta_t, N(0, Id)) + W_2(N(m_t, C_t), N(x_0, B))$
 by triangle inequality and $W_2(T_{m, A}^{-1} \# \eta, T_{m, A}^{-1} \# N(0, Id)) \leq \|C_t\|_2^{1/2} W_2(\eta, N(0, Id))$ by Lemma 3.1 in [Carrillo, Vaes '21]
- $\|C_t\|_2 \leq \tilde{k}$ using explicit solution for C_t
- $W_2(N(m_t, C_t), N(x_0, B))^2 \leq e^{-2t} \tilde{k} \left(|m_0 - x_0|_{C_0}^2 + \|Id - B^{1/2} C_0^{-1} B^{1/2}\|_{HS}^2 \right)$
 using the explicit representation of W_2 for Gaussians and explicit solutions m_t, C_t :
 $W_2(N(m_t, C_t), N(x_0, B))^2 = |m_t - x_0|^2 + \text{Tr}[C_t + B - 2(B^{1/2} C_t B^{1/2})^{1/2}]$ □

Comparison to variance-modulated optimal transport

Variance-modulated GF

$$\partial_t \rho = \text{var}(\rho) \nabla \cdot (\rho \nabla \mathcal{E}'(\rho | \pi))$$

$$\mathcal{E}(\rho_t | \pi) \leq e^{-2t\lambda} \mathcal{E}(\rho_0 | \pi)$$

with

$$\lambda = \min \left\{ \frac{d}{\|B\|_2 \|B^{-1}\|_2}, \frac{d}{\|B\|_2 \|C_0^{-1}\|_2} \right\}$$

► Exponential rate depends on EV of B !

Covariance-modulated GF

$$\partial_t \rho = \nabla \cdot (\rho C(\rho) \nabla \mathcal{E}'(\rho | \pi))$$

$$\mathcal{E}(\rho_t | \pi) \leq \kappa e^{-2t} \mathcal{E}(\rho_0 | \pi)$$

with

$$\kappa = \left(1 \vee \|B^{1/2} C_0^{-1} B^{1/2}\|_2\right) \left(1 \vee \|B^{-1/2} C_0 B^{-1/2}\|_2\right)$$

► Only prefactor depends on EV of B !