Gradient Flows in the Geometry of the Sinkhorn Divergence

Mathis Hardion

10/10/2024

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \end{cases}$

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V) \end{cases}$

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive . . .

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive ...

↳ Use its entropic regularization $\mathrm{OT}_{\varepsilon}$ instead (Peyré, 2015).

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \end{cases}$ Our focus ($V \in \mathcal{C}(\mathcal{X})$).

But W_2^2 is computationally expensive . . .

↳ Use its entropic regularization $\mathrm{OT}_{\varepsilon}$ instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

Introduction

Wasserstein JKO: $\mu_{k+1}^\tau \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^\tau)^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive . . .

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \mathrm{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive ...

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

$\mathrm{OT}_\varepsilon(\mu, \mu) > 0$ ...

Introduction

Wasserstein JKO: $\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^{\tau})^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive ...

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

$\mathrm{OT}_\varepsilon(\mu, \mu) > 0$ ... $\to$ use the debiased Sinkhorn Divergence.

Introduction

Wasserstein JKO: $\mu^\tau_{k+1} \in \underset{\mu\in\mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu^\tau_k)^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V\rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive ...

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

$\mathrm{OT}_\varepsilon(\mu, \mu) > 0$ ... $\to$ use the debiased Sinkhorn Divergence.

Sinkhorn JKO: $\mu^\tau_{k+1} \in \underset{\mu\in\mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} S_\varepsilon(\mu, \mu^\tau_k)$.

Introduction

Wasserstein JKO: $\mu_{k+1}^\tau \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu_k^\tau)^2$

$$\text{as } \tau \to 0: \begin{cases} E(\mu) = \int_{\mathcal{X}} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$$

But W_2^2 is computationally expensive . . .

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

$\mathrm{OT}_\varepsilon(\mu, \mu) > 0$ . . . $\to$ use the debiased Sinkhorn Divergence.

Sinkhorn JKO: $\mu_{k+1}^\tau \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} E(\mu) + \frac{1}{2\tau} S_\varepsilon(\mu, \mu_k^\tau)$.

as $\tau \to 0$: **??**

Introduction

Wasserstein JKO: $\mu^\tau_{k+1} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} W_2(\mu, \mu^\tau_k)^2$

as $\tau \to 0$: $\begin{cases} E(\mu) = \int_\mathcal{X} \log(\mu)\, d\mu \to \dot{\mu}_t = \Delta \mu_t \\ \boxed{E(\mu) = \langle \mu, V \rangle \to \dot{\mu}_t = \operatorname{div}(\mu_t \nabla V)} \text{ Our focus } (V \in \mathcal{C}(\mathcal{X})). \end{cases}$

But W_2^2 is computationally expensive ...

↳ Use its entropic regularization OT_ε instead (Peyré, 2015).

Above PDEs are recovered when $\varepsilon \ll \tau$ (Carlier et al., 2017), but it slows down convergence.

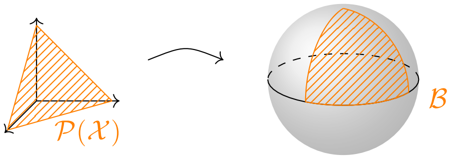
For fixed $\varepsilon > 0$: OT_ε numerically accessible, smooth, better statistical properties (Genevay et al., 2018, 2019).

$\mathrm{OT}_\varepsilon(\mu, \mu) > 0$ ... $\to$ use the debiased Sinkhorn Divergence.

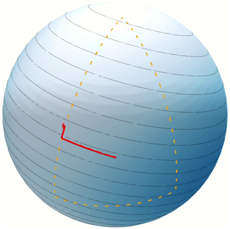
Sinkhorn JKO: $\mu^\tau_{k+1} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min}\, E(\mu) + \frac{1}{2\tau} S_\varepsilon(\mu, \mu^\tau_k)$.

as $\tau \to 0$: ?? $\to$ **We derive the equation, analyze its structure and properties.**

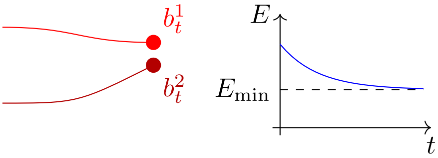
Plan



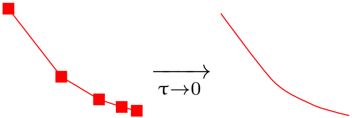
1. The Riemannian Geometry of S_ε



2. The equation and its structure

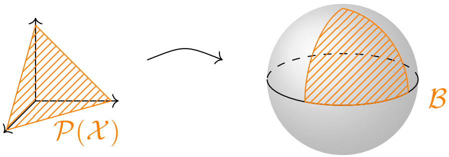


3. Well posedness and properties

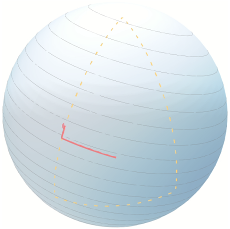


4. Convergence of the SJKO scheme

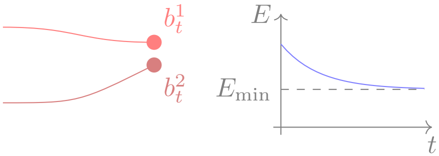
Plan



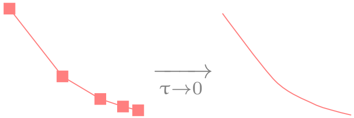
1. The Riemannian Geometry of S_ε



2. The equation and its structure



3. Well posedness and properties



4. Convergence of the SJKO scheme

The Riemannian Geometry of S_ε

Notation.

- $k_c \coloneqq \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),

The Riemannian Geometry of S_ε

Notation.

- $k_c := \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,

The Riemannian Geometry of S_ε

Notation.

- $k_c \coloneqq \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,
- $k_\mu \coloneqq \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Notation.

- $k_c \coloneqq \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,
- $k_\mu \coloneqq \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Theorem. *Let $\mu_t = \mu + t\dot{\mu}$ with $\dot{\mu} \in \mathcal{M}_0(\mathcal{X})$ a signed balanced measure. Then*

$$\frac{1}{t^2} S_\varepsilon(\mu, \mu_t) \xrightarrow[t\to 0]{} \langle \dot{\mu}, G_\mu[\dot{\mu}] \rangle$$

The Riemannian Geometry of S_ε

Notation.

- $k_c := \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,
- $k_\mu := \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Theorem. *Let $\mu_t = \mu + t\dot{\mu}$ with $\dot{\mu} \in \mathcal{M}_0(\mathcal{X})$ a signed balanced measure. Then*

$$\frac{1}{t^2} S_\varepsilon(\mu, \mu_t) \xrightarrow[t \to 0]{} \langle \dot{\mu}, G_\mu [\dot{\mu}] \rangle =: \mathbf{g}_\mu(\dot{\mu}, \dot{\mu}).$$

The Riemannian Geometry of S_ε

Notation.

- $k_c \coloneqq \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu,\mu)$,
- $k_\mu \coloneqq \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Theorem. *Let $\mu_t = \mu + t\dot{\mu}$ with $\dot{\mu} \in \mathcal{M}_0(\mathcal{X})$ a signed balanced measure. Then*

$$\frac{1}{t^2} S_\varepsilon(\mu, \mu_t) \xrightarrow[t\to 0]{} \langle \dot{\mu}, G_\mu[\dot{\mu}] \rangle \eqqcolon \mathbf{g}_\mu(\dot{\mu}, \dot{\mu}).$$

Theorem. *The completion of $\mathcal{M}_0(\mathcal{X})$ w.r.t. $\mathbf{g}_\mu$ is $\mathcal{H}^*_{\mu,0} \coloneqq \left\{\sigma \in \mathcal{H}^*_\mu \;\middle|\; \langle \sigma, 1\rangle = 0\right\}$.*

Notation.

- $k_c := \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,
- $k_\mu := \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Theorem. *Let $\mu_t = \mu + t\dot{\mu}$ with $\dot{\mu} \in \mathcal{M}_0(\mathcal{X})$ a signed balanced measure. Then*

$$\frac{1}{t^2} S_\varepsilon(\mu, \mu_t) \xrightarrow[t \to 0]{} \langle \dot{\mu}, G_\mu [\dot{\mu}] \rangle =: \mathbf{g}_\mu(\dot{\mu}, \dot{\mu}).$$

Theorem. *The completion of $\mathcal{M}_0(\mathcal{X})$ w.r.t. $\mathbf{g}_\mu$ is $\mathcal{H}^*_{\mu,0} := \left\{\sigma \in \mathcal{H}^*_\mu \,\middle|\, \langle \sigma, 1 \rangle = 0\right\}$.*

↳ Includes vertical and horizontal perturbations for c quadratic cost.

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Notation.

- $k_c := \exp\left(-\frac{c}{\varepsilon}\right)$ (assumed to be a universal kernel of RKHS $\mathcal{H}_c$),
- f_μ: Schrödinger potential for $\mathrm{OT}_\varepsilon(\mu, \mu)$,
- $k_\mu := \exp\left(\frac{1}{\varepsilon}(f_\mu \oplus f_\mu - c)\right)$, of RKHS $\mathcal{H}_\mu$.

Theorem. *Let $\mu_t = \mu + t\dot{\mu}$ with $\dot{\mu} \in \mathcal{M}_0(\mathcal{X})$ a signed balanced measure. Then*

$$\frac{1}{t^2} S_\varepsilon(\mu, \mu_t) \xrightarrow[t \to 0]{} \langle \dot{\mu}, G_\mu [\dot{\mu}] \rangle =: \mathbf{g}_\mu(\dot{\mu}, \dot{\mu}).$$

Theorem. *The completion of $\mathcal{M}_0(\mathcal{X})$ w.r.t. $\mathbf{g}_\mu$ is $\mathcal{H}^*_{\mu,0} := \left\{\sigma \in \mathcal{H}^*_\mu \,\middle|\, \langle \sigma, 1 \rangle = 0\right\}$.*

↳ Includes vertical and horizontal perturbations for c quadratic cost.

↳ $\mathcal{H}_\mu$ depends on $\mu \implies$ change of variables

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) := H_c \left[\exp\left(\frac{f_\mu}{\varepsilon} \right) \mu \right]$$

Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) \coloneqq H_c \left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

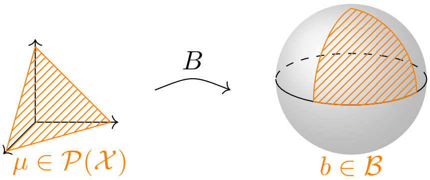
Change of variables

Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) \coloneqq H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} \coloneqq H_c\left[\mathcal{M}_+(\mathcal{X})\right] \cap \left\{b \in \mathcal{H}_c \,\middle|\, \|b\|_{\mathcal{H}_c} = 1\right\}$$



Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

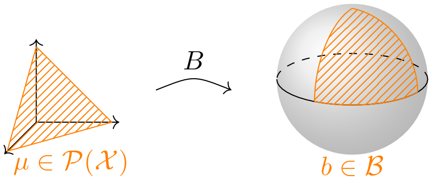
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) \coloneqq H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} \coloneqq \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \left\{b \in \mathcal{H}_c \;\middle|\; \|b\|_{\mathcal{H}_c} = 1\right\}$$

$\mu \geq 0$



Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

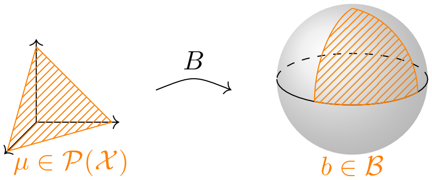
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) := H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} := \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \boxed{\left\{b \in \mathcal{H}_c \;\middle|\; \|b\|_{\mathcal{H}_c} = 1\right\}}$$

$\mu \geq 0$ $\qquad$ $\langle \mu, 1 \rangle = 1$



Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

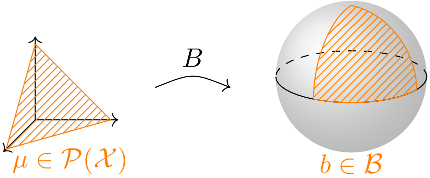
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) := H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} := \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \boxed{\left\{b \in \mathcal{H}_c \;\middle|\; \|b\|_{\mathcal{H}_c} = 1\right\}}$$

$\mu \geq 0$ $\qquad$ $\langle \mu, 1 \rangle = 1$



$\mathbf{g}_\mu \longleftrightarrow \tilde{\mathbf{g}}_\mu$

Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)
Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

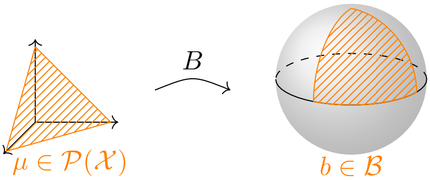
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) \coloneqq H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} \coloneqq \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \boxed{\left\{b \in \mathcal{H}_c \;\middle|\; \|b\|_{\mathcal{H}_c} = 1\right\}}$$

$\mu \geq 0$ $\qquad$ $\langle \mu, 1 \rangle = 1$



$\mathbf{g}_\mu \longleftrightarrow \tilde{\mathbf{g}}_\mu$

$G_\mu \longleftrightarrow \tilde{G}_\mu$

Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)
Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

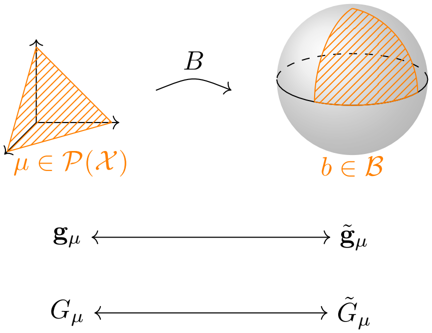
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) := H_c \left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} := \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \boxed{\left\{b \in \mathcal{H}_c \;\middle|\; \|b\|_{\mathcal{H}_c} = 1\right\}}$$

$\mu \geq 0$ $\qquad$ $\langle \mu, 1\rangle = 1$



Admissible paths: $(b_t)_t \in \mathscr{H}^1([0,1], \mathcal{H}_c)$ valued in $\mathcal{B}$.

Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)
Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

Change of variables

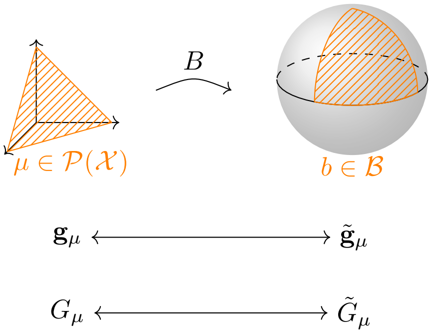
Definition. *With H_c the Riesz isometry $\mathcal{H}_c^* \to \mathcal{H}_c$,*

$$B(\mu) \coloneqq H_c\left[\exp\left(\frac{f_\mu}{\varepsilon}\right)\mu\right] = \exp\left(-\frac{f_\mu}{\varepsilon}\right).$$

Theorem. *The map B is a homeomorphism onto its image*

$$\mathcal{B} \coloneqq \boxed{H_c\left[\mathcal{M}_+(\mathcal{X})\right]} \cap \boxed{\left\{b \in \mathcal{H}_c \mid \|b\|_{\mathcal{H}_c} = 1\right\}}$$

$\mu \geq 0$ $\qquad$ $\langle \mu, 1 \rangle = 1$



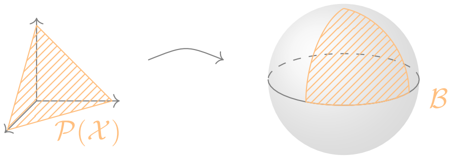
Admissible paths: $(b_t)_t \in \mathscr{H}^1([0,1], \mathcal{H}_c)$ valued in $\mathcal{B}$.

Geodesic distance d_S defined by minimizing $\displaystyle\int_0^1 \tilde{\mathbf{g}}_{\mu_t}\left(\dot{b}_t, \dot{b}_t\right) dt$ over admissible paths.

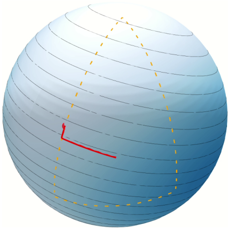
Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences. (2019)

Lavenant, Luckhardt, Mordant, Schmitzer, Tamanini. The Riemannian geometry of Sinkhorn divergences. (2024)

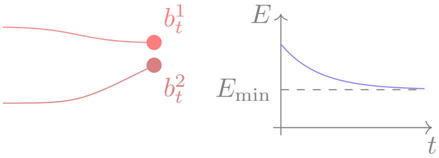
Plan



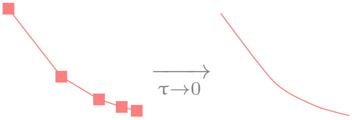
1. The Riemannian Geometry of S_ε



2. The equation and its structure



3. Well posedness and properties



4. Convergence of the SJKO scheme

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \, \langle \mu, V \rangle + \frac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \tfrac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \tfrac{1}{2\tau} \langle \mu - \mu_k^{\tau}, G_{\mu_k}(\mu - \mu_k^{\tau}) \rangle$$

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \frac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \frac{1}{2\tau} \langle \mu - \mu_k^{\tau}, G_{\mu_k}(\mu - \mu_k^{\tau}) \rangle$$

1st order conditions + formal limit:

$$\begin{cases} G_{\mu_t}[\dot{\mu}_t] + V + p_t = 0 \\ p_t \leq 0 \\ \langle \mu_t, p_t \rangle = 0 \\ \mu_t \in \mathcal{P}(\mathcal{X}) \end{cases}$$

Gradient flow of $\mu \mapsto \langle \mu, V \rangle$

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \tfrac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^{\tau} \in \underset{\mu \in \mathcal{P}(\mathcal{X})}{\arg\min} \langle \mu, V \rangle + \tfrac{1}{2\tau} \langle \mu - \mu_k^{\tau}, G_{\mu_k}(\mu - \mu_k^{\tau}) \rangle$$

1st order conditions + formal limit:

$$\begin{cases} G_{\mu_t}[\dot{\mu}_t] + V + p_t = 0 \\ p_t \leq 0 \\ \langle \mu_t, p_t \rangle = 0 \\ \mu_t \in \mathcal{P}(\mathcal{X}) \end{cases}$$

embedding in $\mathcal{B}$

Gradient flow of $\mu \mapsto \langle \mu, V \rangle$ $\longrightarrow$ Gradient flow of $b \mapsto \langle b, Vb \rangle_{\mathcal{H}_c}$

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V \rangle + \frac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V \rangle + \frac{1}{2\tau} \langle \mu - \mu_k^{\tau}, G_{\mu_k}(\mu - \mu_k^{\tau}) \rangle$$

1st order conditions + formal limit:

$$\begin{cases} G_{\mu_t}[\dot{\mu}_t] + V + p_t = 0 \\ p_t \leq 0 \\ \langle \mu_t, p_t \rangle = 0 \\ \mu_t \in \mathcal{P}(\mathcal{X}) \end{cases} \xrightarrow[\text{embedding in } \mathcal{B}]{} \begin{cases} \tilde{G}_{\mu_t} \dot{b}_t + (V + V^*) b_t + p_t = 0 \\ p_t \leq 0 \\ \langle H_c^{-1} b_t, p_t \rangle = 0 \\ b_t \in \mathcal{B} \end{cases}$$

$$\text{Gradient flow of } \mu \mapsto \langle \mu, V \rangle \longrightarrow \text{Gradient flow of } b \mapsto \langle b, Vb \rangle_{\mathcal{H}_c}$$

Derivation of the equation

$$\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V \rangle + \tfrac{1}{2\tau} S_{\varepsilon}(\mu, \mu_k^{\tau})$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^{\tau} \in \arg\min_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V \rangle + \tfrac{1}{2\tau} \langle \mu - \mu_k^{\tau}, G_{\mu_k}(\mu - \mu_k^{\tau}) \rangle$$

1st order conditions + formal limit:

$$\begin{cases} G_{\mu_t}[\dot{\mu}_t] + V + p_t = 0 \\ p_t \leq 0 \\ \langle \mu_t, p_t \rangle = 0 \\ \mu_t \in \mathcal{P}(\mathcal{X}) \end{cases} \xrightarrow[\text{embedding in } \mathcal{B}]{} \begin{cases} \tilde{G}_{\mu_t} \dot{b}_t + (V + V^*) b_t + p_t = 0 \\ p_t \leq 0 \\ \left\langle H_c^{-1} b_t, p_t \right\rangle = 0 \\ b_t \in \mathcal{B} \end{cases}$$

Gradient flow of $\mu \mapsto \langle \mu, V \rangle$ $\longrightarrow$ Gradient flow of $b \mapsto \langle b, Vb \rangle_{\mathcal{H}_c}$

$$\triangle\!\!! \quad \underbrace{\tilde{G}_{\mu_t} \dot{b}_t}_{\in \mathcal{H}_c} + \underbrace{V b_t}_{\in \mathcal{C}(\mathcal{X})} + \underbrace{V^* b_t}_{\in \mathcal{H}_c} + \underbrace{p_t}_{\in \mathcal{C}(\mathcal{X})} = 0$$

Derivation of the equation

$$\mu_{k+1}^\tau \in \operatorname*{arg\,min}_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V\rangle + \frac{1}{2\tau} S_\varepsilon(\mu, \mu_k^\tau)$$

Intuitively: asymptotic equivalence with

$$\mu_{k+1}^\tau \in \operatorname*{arg\,min}_{\mu \in \mathcal{P}(\mathcal{X})} \langle \mu, V\rangle + \frac{1}{2\tau} \langle \mu - \mu_k^\tau, G_{\mu_k}(\mu - \mu_k^\tau)\rangle$$

1st order conditions + formal limit:

$$\begin{cases} G_{\mu_t}[\dot\mu_t] + V + p_t = 0 \\ p_t \leq 0 \\ \langle \mu_t, p_t\rangle = 0 \\ \mu_t \in \mathcal{P}(\mathcal{X}) \end{cases} \xrightarrow[\text{embedding in } \mathcal{B}]{} \begin{cases} \tilde G_{\mu_t} \dot b_t + (V + V^*) b_t + p_t = 0 \\ p_t \leq 0 \\ \left\langle H_c^{-1} b_t, p_t \right\rangle = 0 \\ b_t \in \mathcal{B} \end{cases}$$

Gradient flow of $\mu \mapsto \langle \mu, V\rangle$ $\longrightarrow$ Gradient flow of $b \mapsto \langle b, Vb\rangle_{\mathcal{H}_c} = \left\langle H_c^{-1} b, Vb\right\rangle$

⚠ $$\underbrace{\tilde G_{\mu_t} \dot b_t}_{\in \mathcal{H}_c} + \underbrace{V b_t}_{\in \mathcal{C}(\mathcal{X})} + \underbrace{V^* b_t}_{\in \mathcal{H}_c} + \underbrace{p_t}_{\in \mathcal{C}(\mathcal{X})} = 0$$

Structure of the equation

Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty), \mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \leq 0 \\ \langle H_c^{-1} b_t, p_t \rangle = 0 \\ b_t \in \mathcal{B} \end{cases}$$

Structure of the equation

Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty), \mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \left\langle H_c^{-1}b, p\right\rangle = 0\right\}$$

the set of ***pressure vectors*** *at b.*

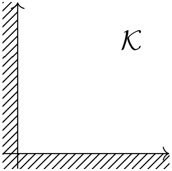
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \left\langle H_c^{-1}b, p\right\rangle = 0\right\}$$

the set of ***pressure vectors*** *at b.*



Structure of the equation

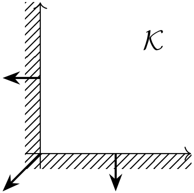
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \;\middle|\; p \leq 0 \text{ and } \left\langle H_c^{-1}b, p \right\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Structure of the equation

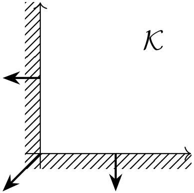
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \langle H_c^{-1}b, p\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

Structure of the equation

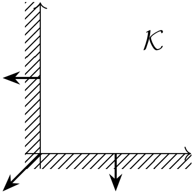
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty), \mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \left\langle H_c^{-1}b, p \right\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial\iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P}b$

Structure of the equation

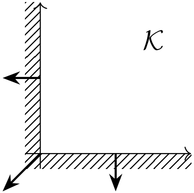
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \left\langle H_c^{-1}b, p \right\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial\iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P}b$

(b) $\tilde{G}_\mu^{-1}\mathfrak{P}b = \mathfrak{P}b$

Structure of the equation

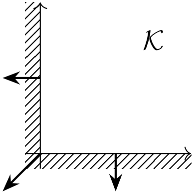
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \tilde{G}_{\mu_t}\dot{b}_t + (V+V^*)b_t + p_t = 0 \\ p_t \in \mathfrak{P}b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \;\middle|\; p \le 0 \text{ and } \left\langle H_c^{-1}b, p \right\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial\iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P}b$

(b) $\tilde{G}_\mu^{-1}\mathfrak{P}b = \mathfrak{P}b$

(c) $\tilde{G}_\mu^{-1}(V+V^*) = \frac{2}{\varepsilon}(V-V^*) =: \mathbf{W}$

Structure of the equation

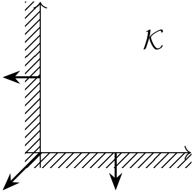
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \dot b_t + \mathbf{W} b_t + p_t = 0 \\ p_t \in \mathfrak{P} b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \le 0 \text{ and } \langle H_c^{-1} b, p\rangle = 0 \right\}$$

the set of ***pressure vectors*** *at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial \iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P}b$

(b) $\tilde G_\mu^{-1} \mathfrak{P} b = \mathfrak{P} b$

(c) $\tilde G_\mu^{-1}(V + V^*) = \frac{2}{\varepsilon}(V - V^*) =: \mathbf{W}$

Structure of the equation

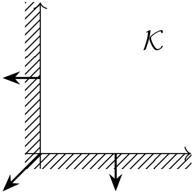
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \dot{b}_t + \mathbf{W} b_t + p_t = 0 \\ p_t \in \mathfrak{P} b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P} b := \left\{ p \in \mathcal{C}(\mathcal{X}) \,\middle|\, p \leq 0 \text{ and } \langle H_c^{-1} b, p \rangle = 0 \right\}$$

*the set of **pressure vectors** at b.*



Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial \iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P} b$

(b) $\tilde{G}_\mu^{-1} \mathfrak{P} b = \mathfrak{P} b$

(c) $\tilde{G}_\mu^{-1}(V + V^*) = \boxed{\frac{2}{\varepsilon}(V - V^*) =: \mathbf{W}}$

Antisymmetric: generates rotational motion!

Structure of the equation

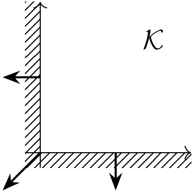
Sinkhorn Potential Flow:

$$(b_t)_t \in \mathscr{H}^1([0,+\infty),\mathcal{H}_c), \begin{cases} \dot b_t + \mathbf{W} b_t + p_t = 0 \\ p_t \in \mathfrak{P} b_t \\ b_t \in \mathcal{B} \end{cases}$$

Definition. *For $b \in \mathcal{K} := H_c[\mathcal{M}_+(\mathcal{X})]$, we call*

$$\mathfrak{P}b := \left\{ p \in \mathcal{C}(\mathcal{X}) \;\middle|\; p \le 0 \text{ and } \langle H_c^{-1} b, p \rangle = 0 \right\}$$

*the set of **pressure vectors** at b.*

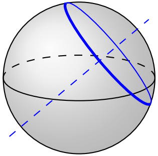


Proposition. *Denote $\iota_{\mathcal{K}}$ the convex indicator of $\mathcal{K}$. Then for $b \in \mathcal{K}$,*

(a) $\partial \iota_{\mathcal{K}}(b) = \mathcal{H}_c \cap \mathfrak{P}b$

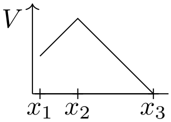
(b) $\tilde G_\mu^{-1} \mathfrak{P} b = \mathfrak{P} b$

(c) $\tilde G_\mu^{-1}(V + V^*) = \boxed{\frac{2}{\varepsilon}(V - V^*) =: \mathbf{W}}$



Antisymmetric: generates rotational motion!

Numerics on the 3-point space

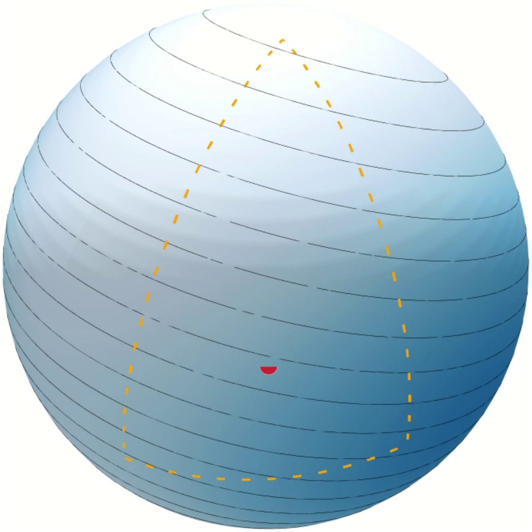


Boundary of $\mathcal{B}$

Theoretical rotation lines

SJKO Flow (embedded)

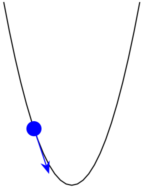
Potential energy $\langle H_c^{-1} b, V b \rangle$



Flow of a Dirac mass

Proposition. *If* $(x_t)_t \subset \mathcal{X}$ *is a smooth trajectory and* $b_t = B(\delta_{x_t})$ *then*

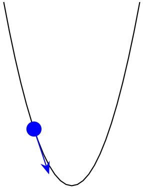
$$\dot{b}_t + \mathbf{W} b_t + \mathfrak{P} b_t \ni 0 \iff \dot{x}_t \in -\partial V(x_t).$$



Flow of a Dirac mass

Proposition. *If $(x_t)_t \subset \mathcal{X}$ is a smooth trajectory and $b_t = B(\delta_{x_t})$ then*

$$\dot{b}_t + \mathbf{W} b_t + \mathfrak{P} b_t \ni 0 \iff \dot{x}_t \in -\partial V(x_t).$$

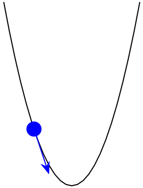


PROOF: Direct computations.

Flow of a Dirac mass

Proposition. *If* $(x_t)_t \subset \mathcal{X}$ *is a smooth trajectory and* $b_t = B(\delta_{x_t})$ *then*

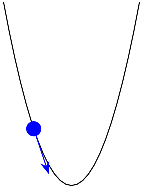
$$\dot{b}_t + \mathbf{W} b_t + \mathfrak{P} b_t \ni 0 \iff \dot{x}_t \in -\partial V(x_t).$$



Proof: Direct computations.

Corollary. *For* V *convex and any* $x_0 \in \mathcal{X}$, *the Sinkhorn potential flow starting at* $b_0 = B(\delta_{x_0})$ *is given by* $B(\delta_{x_t})$ *with* $(x_t)_t$ *the subgradient flow of* V.

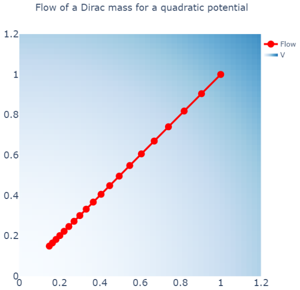
Flow of a Dirac mass

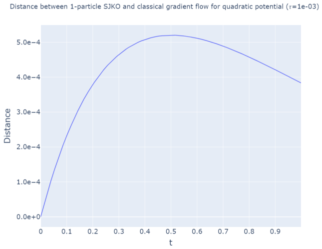


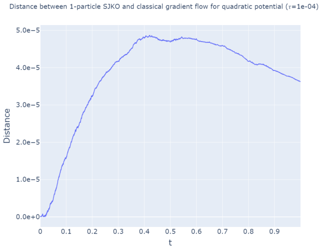
> **Proposition.** *If $(x_t)_t \subset \mathcal{X}$ is a smooth trajectory and $b_t = B(\delta_{x_t})$ then*
> $$\dot{b}_t + \mathbf{W} b_t + \mathfrak{P} b_t \ni 0 \iff \dot{x}_t \in -\partial V(x_t).$$

PROOF: Direct computations.

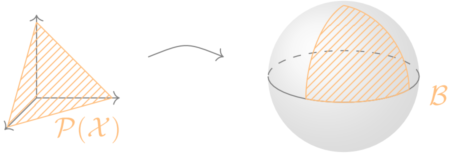
> **Corollary.** *For V convex and any $x_0 \in \mathcal{X}$, the Sinkhorn potential flow starting at $b_0 = B(\delta_{x_0})$ is given by $B(\delta_{x_t})$ with $(x_t)_t$ the subgradient flow of V.*



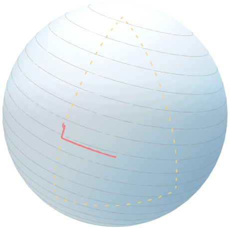




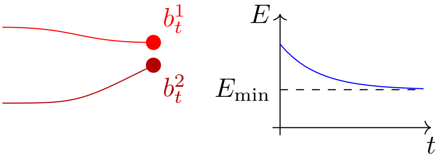
Plan



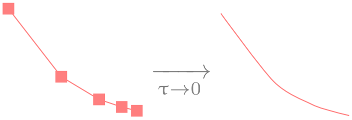
1. The Riemannian Geometry of S_ε



2. The equation and its structure



3. Well posedness and properties



4. Convergence of the SJKO scheme

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

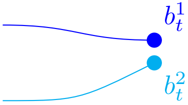
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



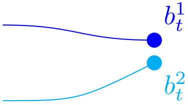
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

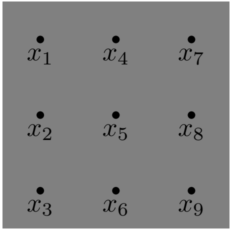
Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

For $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$



$\mathcal{X}$

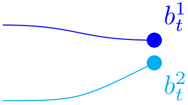
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

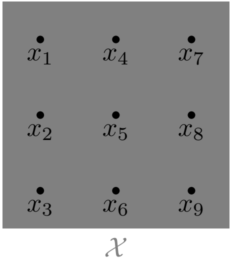
(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

For $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$

$\hookrightarrow$ $\mathbf{W}$ antisymmetric, continuous $\mathcal{H}_c \to \mathcal{H}_c$ and $\mathfrak{P} = \partial \iota_{\mathcal{K}}$,



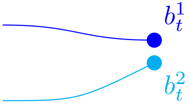
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*

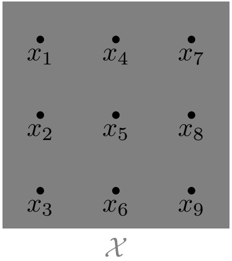


Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

For $\mathcal{X} = \{x_1, \dots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$

$\hookrightarrow$ $\mathbf{W}$ antisymmetric, continuous $\mathcal{H}_c \to \mathcal{H}_c$ and $\mathfrak{P} = \partial \iota_{\mathcal{K}}$,

$\hookrightarrow$ $\mathbf{W} + \mathfrak{P}$ maximal monotone (gives contractivity),



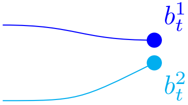
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



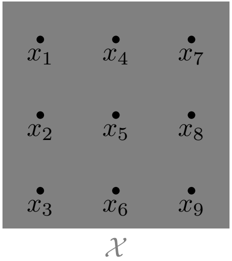
Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

For $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$

$\hookrightarrow$ $\mathbf{W}$ antisymmetric, continuous $\mathcal{H}_c \to \mathcal{H}_c$ and $\mathfrak{P} = \partial \iota_{\mathcal{K}}$,

$\hookrightarrow$ $\mathbf{W} + \mathfrak{P}$ maximal monotone (gives contractivity),

$\hookrightarrow$ We can use the Hille-Yosida theorem.



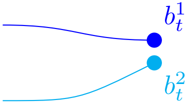
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

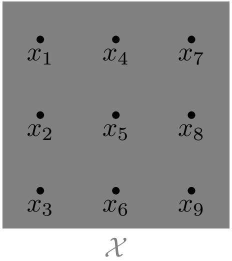
For $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$

↳ $\mathbf{W}$ antisymmetric, continuous $\mathcal{H}_c \to \mathcal{H}_c$ and $\mathfrak{P} = \partial \iota_{\mathcal{K}}$,

↳ $\mathbf{W} + \mathfrak{P}$ maximal monotone (gives contractivity),

↳ We can use the Hille-Yosida theorem.

Then, use classical compactness results and verify that the limit $n \to \infty$ verifies the equation.



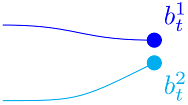
Existence, uniqueness, contractivity

Theorem. *There exists a unique Sinkhorn potential flow $(b_t)_t$ starting at $b^0 \in \mathcal{B}$ which additionally verifies*

(a) *The norm $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ decreases.*

Moreover,

(b) *The flow is contractive i.e. two flows b^1, b^2 have decreasing $\left\|b_t^1 - b_t^2\right\|_{\mathcal{H}_c}$.*



Sketch of proof: Discretize the space to utilize the structure of $\mathcal{H}_c$:

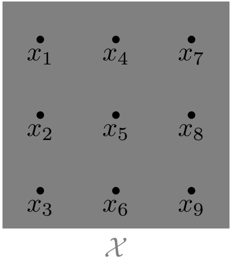
For $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{C}(\mathcal{X}) = \mathcal{H}_c = \mathbb{R}^n$

$\hookrightarrow$ $\mathbf{W}$ antisymmetric, continuous $\mathcal{H}_c \to \mathcal{H}_c$ and $\mathfrak{P} = \partial\iota_{\mathcal{K}}$,

$\hookrightarrow$ $\mathbf{W} + \mathfrak{P}$ maximal monotone (gives contractivity),

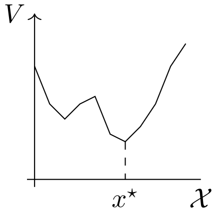
$\hookrightarrow$ We can use the Hille-Yosida theorem.

Then, use classical compactness results and verify that the limit $n \to \infty$ verifies the equation. **The results hold for any $V \in \mathcal{C}(\mathcal{X})$!**



Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$.*

Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$.*

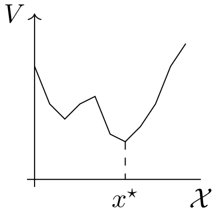


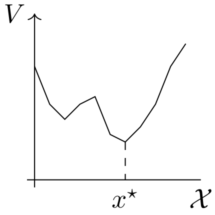
Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$. Then,*

$$b_t \xrightarrow[t\to\infty]{} b_{\min} \coloneqq B(\delta_{x^\star})$$

and thus

$$\mu_t \xrightharpoonup[t\to\infty]{*} \delta_{x^\star}.$$





Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$. Then,*

$$b_t \xrightarrow[t\to\infty]{} b_{\min} := B(\delta_{x^\star})$$

and thus

$$\mu_t \xrightharpoonup[t\to\infty]{*} \delta_{x^\star}.$$

SKETCH OF PROOF:

- $\mathcal{B}$ compact $\implies$ convergent subsequences exist

Asymptotics

Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$. Then,*

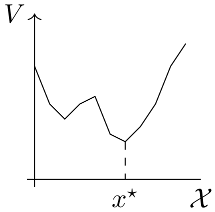
$$b_t \xrightarrow[t\to\infty]{} b_{\min} := B(\delta_{x^\star})$$

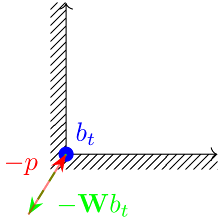
and thus

$$\mu_t \xrightarrow[t\to\infty]{*} \delta_{x^\star}.$$

Sketch of proof:

- $\mathcal{B}$ compact $\implies$ convergent subsequences exist
- $\dot{b}_t \to 0$ and $\mathbf{W} + \mathfrak{P}$ closed $\implies$ accumulation points are critical



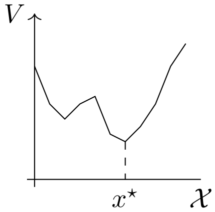


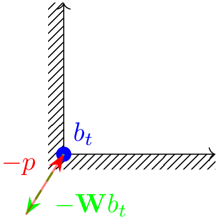
Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$. Then,*

$$b_t \xrightarrow[t\to\infty]{} b_{\min} := B(\delta_{x^\star})$$

and thus

$$\mu_t \xrightarrow[t\to\infty]{*} \delta_{x^\star}.$$





SKETCH OF PROOF:

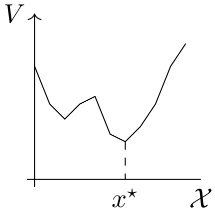
- $\mathcal{B}$ compact $\implies$ convergent subsequences exist
- $\dot{b}_t \to 0$ and $\mathbf{W} + \mathfrak{P}$ closed $\implies$ accumulation points are critical
- $x^\star$ is the unique minimizer $\implies$ $b_{\min}$ is the only critical point.

Theorem. *Assume V has a unique minimizer $x^\star$ on $\mathcal{X}$. Then,*

$$b_t \xrightarrow[t\to\infty]{} b_{\min} := B(\delta_{x^\star})$$

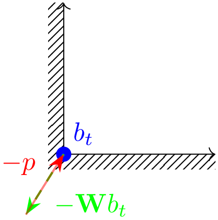
and thus

$$\mu_t \xrightharpoonup[t\to\infty]{*} \delta_{x^\star}.$$

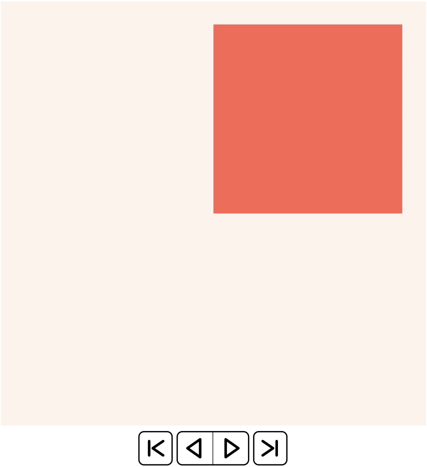


Sketch of proof:

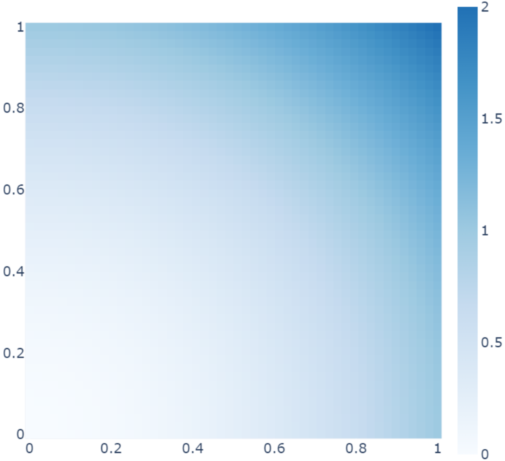
- $\mathcal{B}$ compact $\implies$ convergent subsequences exist
- $\dot{b}_t \to 0$ and $\mathbf{W} + \mathfrak{P}$ closed $\implies$ accumulation points are critical
- $x^\star$ is the unique minimizer $\implies$ $b_{\min}$ is the only critical point.



Because vertical perturbations are admissible for Sinkhorn!

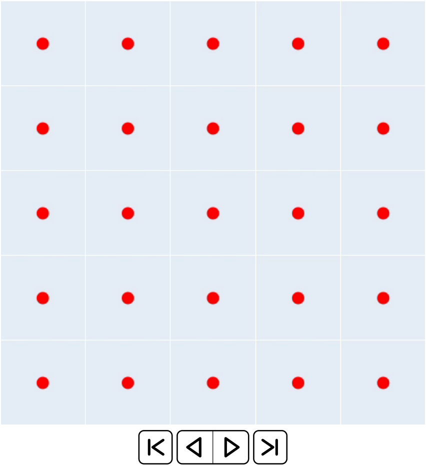


Sinkhorn flow ($\varepsilon = 0.2$)

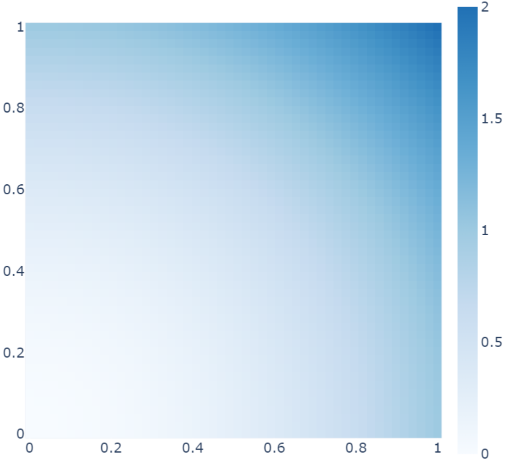


V

Some numerics

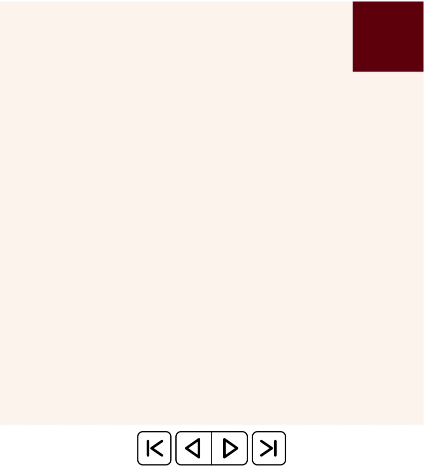


Sinkhorn flow ($\varepsilon = 0.2$)

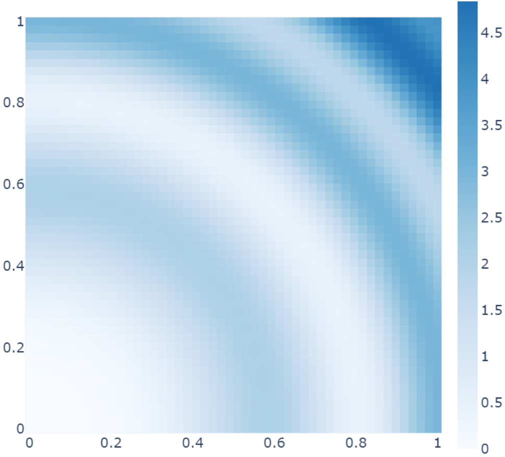


V

Some numerics

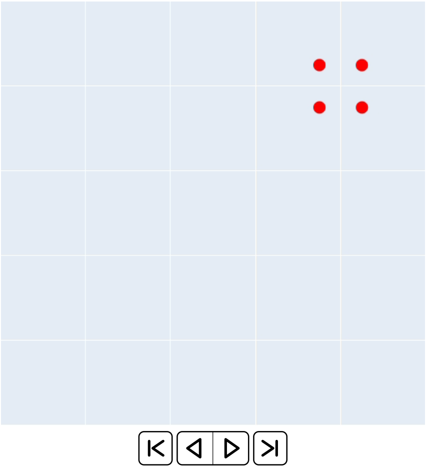


Sinkhorn flow ($\varepsilon = 5$)

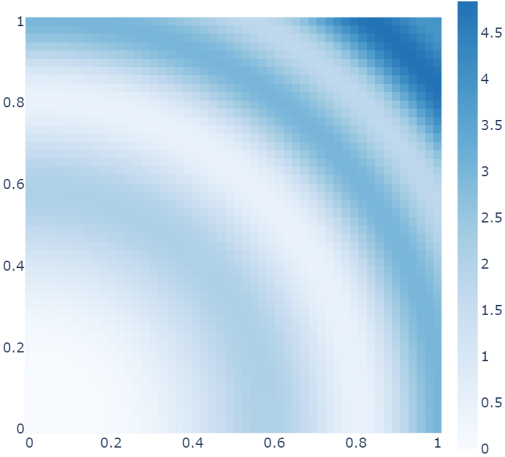


V

Some numerics

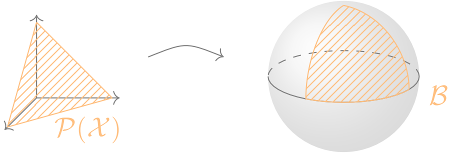


Sinkhorn flow ($\varepsilon = 5$)

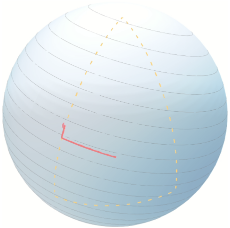


V

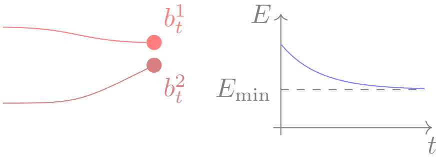
Plan



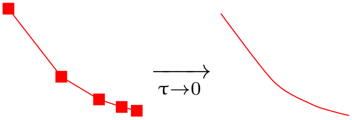
1. The Riemannian Geometry of S_ε



2. The equation and its structure



3. Well posedness and properties



4. Convergence of the SJKO scheme

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Sketch of proof:

- Classical estimates and compactness yields a limit

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau, \mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

SKETCH OF PROOF:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau,\mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$

$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right)$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b^\tau_k)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}^\tau_t\right)_t$ the piecewise constant interpolation, $(b^\tau_t)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

SKETCH OF PROOF:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu^\tau_{k+1},\mu^\tau_k} - f_{\mu^\tau_{k+1}}\right) + V + p^\tau_{k+1} = 0$

$$\frac{1}{2\tau}\left(f^\tau_{t,t-\tau} - f^\tau_{t,t}\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \frac{\partial f^\tau_{t,s}}{\partial s} ds$$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau, \mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$

$$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \frac{\partial f_{t,s}^\tau}{\partial s} ds$$

$$\xrightarrow[\tau \to 0]{} -\frac{1}{2}\frac{\partial f_{t,t}}{\partial t} = G_{\mu_t}\left[\dot{\mu}_t\right]$$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau\to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau,\mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$

$$p_{k+1}^\tau \xrightarrow[\tau\to 0]{} p_t$$

$$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \frac{\partial f_{t,s}^\tau}{\partial s} ds$$

$$\xrightarrow[\tau\to 0]{} -\frac{1}{2}\frac{\partial f_{t,t}}{\partial t} = G_{\mu_t}\left[\dot{\mu}_t\right]$$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau, \mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$, with $p_{k+1}^\tau \xrightarrow[\tau \to 0]{} p_t$

$$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \boxed{\frac{\partial f_{t,s}^\tau}{\partial s}} ds$$

Only valid in finite space !
In the general case, tangent spaces are incompatible.

$$\xrightarrow[\tau \to 0]{} -\frac{1}{2}\frac{\partial f_{t,t}}{\partial t} = G_{\mu_t}\left[\dot{\mu}_t\right]$$

Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0, T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Hope:

$$\begin{array}{ccc} \mathrm{SJKO}_n & \xrightarrow[\tau \to 0]{} & \mathrm{Flow}_n \\ & & \downarrow n \to \infty \\ & & \mathrm{Flow} \end{array}$$

Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau, \mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$, with $p_{k+1}^\tau \xrightarrow[\tau \to 0]{} p_t$

$$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \boxed{\frac{\partial f_{t,s}^\tau}{\partial s}} ds$$

Only valid in finite space !

In the general case, tangent spaces are incompatible.

$$\xrightarrow[\tau \to 0]{} -\frac{1}{2}\frac{\partial f_{t,t}}{\partial t} = G_{\mu_t}\left[\dot{\mu}_t\right]$$

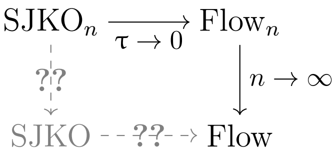
Limit of the SJKO scheme as $\tau \to 0$ (finite $\mathcal{X}$)

Theorem. *On a finite space, with $(b_k^\tau)_k$ given by the SJKO scheme after embedding, $\left(\bar{b}_t^\tau\right)_t$ the piecewise constant interpolation, $(b_t^\tau)_t$ the piecewise geodesic interpolation, then*

$$b^\tau, \bar{b}^\tau \xrightarrow[\tau \to 0]{} b \text{ uniformly on } [0,T]$$

with b the Sinkhorn potential flow of V starting at b_0.

Hope:



Sketch of proof:

- Classical estimates and compactness yields a limit
- Optimality condition of SJKO: $\frac{1}{2\tau}\left(f_{\mu_{k+1}^\tau, \mu_k^\tau} - f_{\mu_{k+1}^\tau}\right) + V + p_{k+1}^\tau = 0$, with $p_{k+1}^\tau \xrightarrow[\tau \to 0]{} p_t$

$$\frac{1}{2\tau}\left(f_{t,t-\tau}^\tau - f_{t,t}^\tau\right) = -\frac{1}{2\tau}\int_{t-\tau}^{t} \boxed{\frac{\partial f_{t,s}^\tau}{\partial s}} ds$$

Only valid in finite space !
In the general case, tangent spaces are incompatible.

$$\xrightarrow[\tau \to 0]{} -\frac{1}{2}\frac{\partial f_{t,t}}{\partial t} = G_{\mu_t}\left[\dot{\mu}_t\right]$$

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive
- The solution converges to the energy minimizer when it is unique, sometimes through ”teleportation”

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive
- The solution converges to the energy minimizer when it is unique, sometimes through "teleportation"

Future research directions:

- Limit $\tau \to 0$ for any compact space

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive
- The solution converges to the energy minimizer when it is unique, sometimes through "teleportation"

Future research directions:

- Limit $\tau \to 0$ for any compact space
- Quantitative asymptotics

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive
- The solution converges to the energy minimizer when it is unique, sometimes through "teleportation"

Future research directions:

- Limit $\tau \to 0$ for any compact space
- Quantitative asymptotics
- Fokker-Planck case $\left(E(\mu) = \langle \mu, V \rangle + \int \log(\mu) d\mu\right)$

Conclusion and perspective

What we have seen:

- SJKO limit (valid for finite $\mathcal{X}$) after embedding is a constrained rotational motion
- The equation is well posed for any $V \in \mathcal{C}(\mathcal{X})$ and contractive
- The solution converges to the energy minimizer when it is unique, sometimes through "teleportation"

Future research directions:

- Limit $\tau \to 0$ for any compact space
- Quantitative asymptotics
- Fokker-Planck case $\left(E(\mu) = \langle \mu, V \rangle + \int \log(\mu) d\mu\right)$

Thank you for listening!

Appendix

Challenges in the general case

$$\frac{\partial f_{t,s}}{\partial s} = -\varepsilon \left(\mathrm{Id} - K_{t,s} K_{s,t}\right)^{-1} H_{t,s}\left[\dot{\mu}_s\right]$$

Challenges in the general case

$$\frac{\partial f_{t,s}}{\partial s} = -\varepsilon \left(\mathrm{Id} - K_{t,s}K_{s,t}\right)^{-1} H_{t,s} \boxed{\dot{\mu}_s} \in \mathcal{H}^*_{\mu_s,0}$$

Challenges in the general case

$$\frac{\partial f_{t,s}}{\partial s} = -\varepsilon\left(\mathrm{Id} - K_{t,s}K_{s,t}\right)^{-1} \boxed{H_{t,s}} \boxed{\dot{\mu}_s} \in \mathcal{H}^*_{\mu_s,0}$$

$$\mathcal{H}^*_{\mu_t,\mu_s} \to \mathcal{H}_{\mu_t,\mu_s}$$

$$\text{where } \mathcal{H}_{\mu,\nu} := \exp\left(\frac{f_{\mu,\nu}}{\varepsilon}\right)\mathcal{H}_c$$

Challenges in the general case

$$\frac{\partial f_{t,s}}{\partial s} = -\varepsilon\left(\mathrm{Id} - K_{t,s}K_{s,t}\right)^{-1} \boxed{H_{t,s}}\, \boxed{\dot{\mu}_s} \in \mathcal{H}^*_{\mu_s,0}$$

But generally speaking $\mathcal{H}^*_{\mu_s,0} \not\subset \mathcal{H}^*_{\mu_t,\mu_s}$!

$$\mathcal{H}^*_{\mu_t,\mu_s} \to \mathcal{H}_{\mu_t,\mu_s}$$

where $\mathcal{H}_{\mu,\nu} \coloneqq \exp\left(\frac{f_{\mu,\nu}}{\varepsilon}\right)\mathcal{H}_c$

Finite space case

Theorem. *On a finite space, the differential inclusion*

$$\begin{cases} \dot{b}_t + \mathbf{W} b_t + p_t = 0 \\ p_t \in \partial \iota_{\mathcal{K}}(b_t) \\ b_t \in \mathcal{B} \end{cases}$$

has a solution that additionally verifies

(a) $p_t = \underset{p \in \partial \iota_{\mathcal{K}}(b_t)}{\arg\min} \|\mathbf{W} b_t + p\|_{\mathcal{H}_c}$

(b) $\frac{d}{dt} E(b_t) = -\tilde{\mathbf{g}}_{\mu_t}\left(\dot{b}_t, \dot{b}_t\right) \implies \begin{cases} (E(b_t))_t \text{ decreases} \\ \int_0^T \left\|\dot{b}_t\right\|_{\mathcal{H}_c}^2 dt \text{ bounded} \end{cases}$

(c) $\left\|\dot{b}_t\right\|_{\mathcal{H}_c}$ *decreases.*

The flow is contractive.