Gradient Flows and Optimal Transport in Discrete and Quantum Systems

Jan Maas (IST Austria)

Geometric Methods in Optimization and Sampling Boot Camp

Simons Institute

1 September 2021

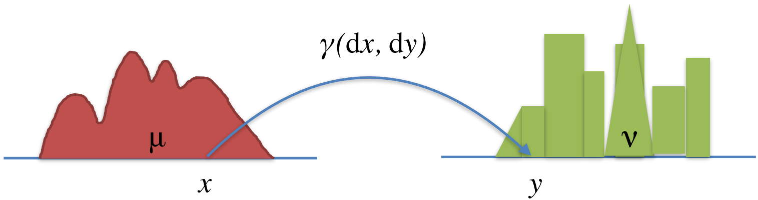
IST AUSTRIA

Institute of Science and Technology

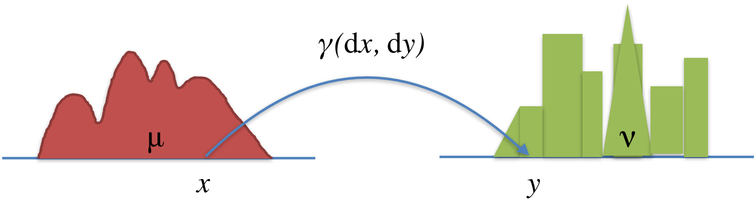
Starting point:

Diffusion equations and Ricci curvature
via optimal transport

Optimal transport

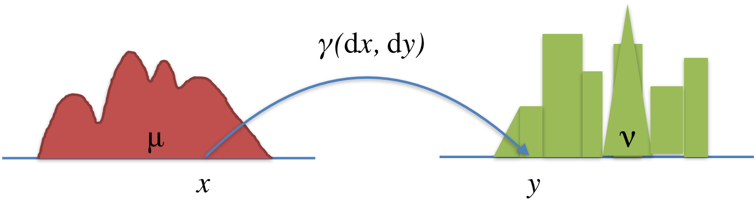


Optimal transport



- Let μ, ν be probability measures on a metric space $(\mathcal{X}, \mathrm{d})$.

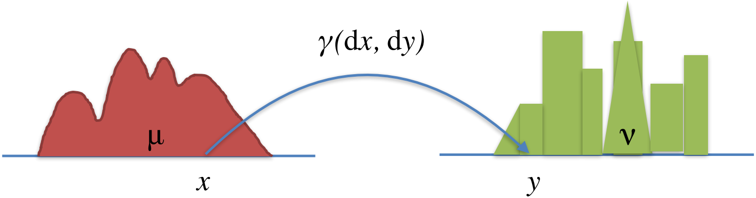
Optimal transport



- Let μ, ν be probability measures on a metric space $(\mathcal{X}, \mathrm{d})$.
- A transport plan (or coupling) between μ and ν is a probability measure $\gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ s.t.

$$\gamma(A \times \mathcal{X}) = \mu(A) \quad \text{and} \quad \gamma(\mathcal{X} \times A) = \nu(A) \quad \forall A \subseteq \mathcal{X} \ .$$

Optimal transport



- Let μ, ν be probability measures on a metric space $(\mathcal{X}, \mathrm{d})$.
- A transport plan (or coupling) between μ and ν is a probability measure $\gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ s.t.

$$\gamma(A \times \mathcal{X}) = \mu(A) \quad \text{and} \quad \gamma(\mathcal{X} \times A) = \nu(A) \quad \forall A \subseteq \mathcal{X} \ .$$

The Monge-Kantorovich problem (1781, 1942)

Minimize $\gamma \mapsto \int_{\mathcal{X} \times \mathcal{X}} c(x, y) \, \mathrm{d}\gamma(x, y)$ among all $\gamma \in \mathrm{Cpl}(\mu, \nu)$.

Diffusion equations via optimal transport

Jordan–Kinderlehrer–Otto '98: Beautiful connection between

- the 2-Kantorovich metric on the space of probability measures

$$W_2(\mu, \nu) = \inf_{\gamma \in \mathsf{Cpl}(\mu,\nu)} \sqrt{\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \, \mathrm{d}\gamma(x, y)}$$

Diffusion equations via optimal transport

Jordan–Kinderlehrer–Otto '98: Beautiful connection between

- the 2-Kantorovich metric on the space of probability measures

$$W_2(\mu,\nu) = \inf_{\gamma \in \mathsf{Cpl}(\mu,\nu)} \sqrt{\int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 \, \mathrm{d}\gamma(x,y)}$$

- the (negative of the) Boltzmann-Shannon entropy

$$\mathsf{Ent}(\mu) = \int_{\mathbb{R}^n} \rho(x) \log \rho(x) \, \mathrm{d}x, \qquad \text{if} \quad \mathrm{d}\mu(x) = \rho(x) \, \mathrm{d}x$$

Diffusion equations via optimal transport

Jordan–Kinderlehrer–Otto '98: Beautiful connection between

- the 2-Kantorovich metric on the space of probability measures

$$W_2(\mu,\nu) = \inf_{\gamma \in \mathsf{Cpl}(\mu,\nu)} \sqrt{\int_{\mathsf{R}^n \times \mathsf{R}^n} |x-y|^2 \, \mathrm{d}\gamma(x,y)}$$

- the (negative of the) Boltzmann-Shannon entropy

$$\mathsf{Ent}(\mu) = \int_{\mathsf{R}^n} \rho(x) \log \rho(x) \, \mathrm{d}x, \qquad \text{if} \quad \mathrm{d}\mu(x) = \rho(x) \, \mathrm{d}x$$

- the heat equation

$$\partial_t \mu = \Delta \mu$$

Diffusion equations via optimal transport

Jordan–Kinderlehrer–Otto '98: Beautiful connection between

- the 2-Kantorovich metric on the space of probability measures

$$W_2(\mu, \nu) = \inf_{\gamma \in \mathsf{Cpl}(\mu,\nu)} \sqrt{\int_{\mathsf{R}^n \times \mathsf{R}^n} |x - y|^2 \, \mathrm{d}\gamma(x, y)}$$

- the (negative of the) Boltzmann-Shannon entropy

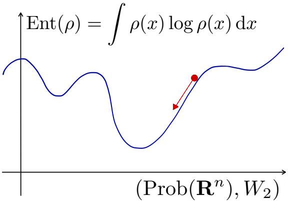
$$\mathsf{Ent}(\mu) = \int_{\mathsf{R}^n} \rho(x) \log \rho(x) \, \mathrm{d}x, \qquad \text{if} \quad \mathrm{d}\mu(x) = \rho(x) \, \mathrm{d}x$$

- the heat equation

$$\partial_t \mu = \Delta \mu$$

Theorem (J-K-O '98)

The heat flow is the gradient flow of the entropy w.r.t W_2.



Diffusion equations via optimal transport

Theorem (JORDAN–KINDERLEHRER–OTTO '98)

The heat flow is the gradient flow of the entropy w.r.t W_2

Diffusion equations via optimal transport

Theorem (Jordan–Kinderlehrer–Otto '98)

The heat flow is the gradient flow of the entropy w.r.t W_2

How to make sense of gradient flows in metric spaces?

Diffusion equations via optimal transport

Theorem (Jordan–Kinderlehrer–Otto '98)

The heat flow is the gradient flow of the entropy w.r.t W_2

How to make sense of gradient flows in metric spaces?

Gradient flows in $\mathbb{R}^n$

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ smooth and convex. For $u : \mathbb{R}_+ \to \mathbb{R}^n$ TFAE:

1. u solves the gradient flow equation $u'(t) = -\nabla\varphi(u(t))$.

Diffusion equations via optimal transport

Theorem (Jordan–Kinderlehrer–Otto '98)

The heat flow is the gradient flow of the entropy w.r.t W_2

How to make sense of gradient flows in metric spaces?

Gradient flows in $\mathbb{R}^n$

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ smooth and convex. For $u : \mathbb{R}_+ \to \mathbb{R}^n$ TFAE:

1. u solves the gradient flow equation $u'(t) = -\nabla\varphi(u(t))$.
2. u satisfies

$$\big(u(t) - y\big) \cdot u'(t) \leq \varphi(y) - \varphi(u(t)) \qquad \forall y \ .$$

Diffusion equations via optimal transport

Theorem (Jordan–Kinderlehrer–Otto '98)

The heat flow is the gradient flow of the entropy w.r.t W_2

How to make sense of gradient flows in metric spaces?

Gradient flows in $\mathbb{R}^n$

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ smooth and convex. For $u : \mathbb{R}_+ \to \mathbb{R}^n$ TFAE:

1. u solves the gradient flow equation $u'(t) = -\nabla\varphi(u(t))$.
2. u satisfies the evolution variational inequality

   $$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|u(t) - y|^2 = \big(u(t) - y\big) \cdot u'(t) \leq \varphi(y) - \varphi(u(t)) \qquad \forall y \; .$$

   (De Giorgi '93, Ambrosio–Gigli–Savaré '05)

Diffusion equations via optimal transport

Theorem (Jordan–Kinderlehrer–Otto '98)

The heat flow is the gradient flow of the entropy w.r.t W_2 , i.e.,

$$\partial_t \mu = \Delta \mu \quad \Longleftrightarrow \quad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} W_2(\mu_t, \nu)^2 \leq \mathrm{Ent}(\nu) - \mathrm{Ent}(\mu_t) \qquad \forall \nu \ .$$

How to make sense of gradient flows in metric spaces?

Gradient flows in R^n

Let $\varphi : \mathrm{R}^n \to \mathrm{R}$ smooth and convex. For $u : \mathrm{R}_+ \to \mathrm{R}^n$ TFAE:

1. u solves the gradient flow equation $u'(t) = -\nabla\varphi(u(t))$.
2. u satisfies the evolution variational inequality

   $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|u(t) - y|^2 = \big(u(t) - y\big) \cdot u'(t) \leq \varphi(y) - \varphi(u(t)) \qquad \forall y \ .$

   (De Giorgi '93, Ambrosio–Gigli–Savaré '05)

Diffusion equations via optimal transport

Many extensions have been proved:

Diffusion equations via optimal transport

Many extensions have been proved:

• $\mathbb{R}^n$	Jordan–Kinderlehrer–Otto
• Riemannian manifolds	Villani, Erbar
• Hilbert spaces	Ambrosio–Savaré–Zambotti
• Finsler spaces	Ohta–Sturm
• Wiener space	Fang–Shao–Sturm
• Heisenberg group	Juillet
• Alexandrov spaces	Gigli–Kuwada–Ohta
• Metric measures spaces	Ambrosio–Gigli–Savaré

Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, ...)

Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, ...)
- is physically appealing

Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, ...)
- is physically appealing
- comes with time-discrete approximation schemes

Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, . . .)
- is physically appealing
- comes with time-discrete approximation schemes
- applies to non-smooth problems

Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, . . .)
- is physically appealing
- comes with time-discrete approximation schemes
- applies to non-smooth problems
- yields functional inequalities and equilibration rates

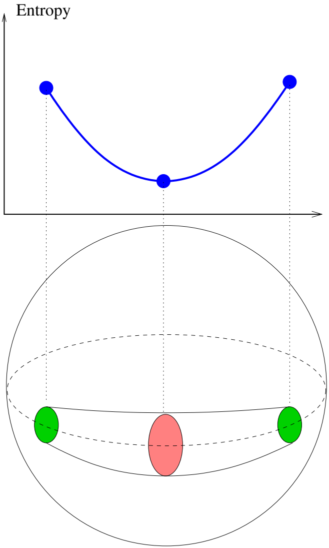
Diffusion equations via optimal transport

Advantages: The optimal transport approach to diffusion equations

- applies to a large class of dissipative equations (Fokker-Planck, porous medium, McKean–Vlasov equations, . . .)
- is physically appealing
- comes with time-discrete approximation schemes
- applies to non-smooth problems
- yields functional inequalities and equilibration rates
- is closely connected to geometry (Ricci curvature)

Optimal transport and curvature

The "lazy gas experiment" (see VILLANI '09)



Ricci curvature via optimal transport

Theorem (McCann '94)

The entropy is convex along geodesics in $(\mathcal{P}(\mathsf{R}^n), W_2)$.

Ricci curvature via optimal transport

Theorem (McCann '94)

The entropy is convex along geodesics in $(\mathcal{P}(\mathbb{R}^n), W_2)$.

Theorem (Otto–Villani '00, Cordero–McCann–Schmuckenschläger '01, von Renesse–Sturm '05)

For a Riemannian manifold $\mathcal{M}$, TFAE:

1. $\mathrm{Ric} \geq \kappa$ everywhere on $\mathcal{M}$;

Ricci curvature via optimal transport

Theorem (McCann '94)

The entropy is convex along geodesics in $(\mathcal{P}(\mathbb{R}^n), W_2)$.

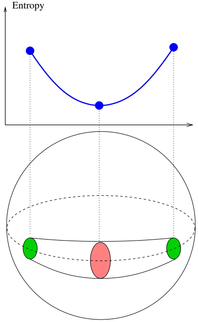
Theorem (Otto–Villani '00, Cordero–McCann–Schmuckenschläger '01, von Renesse–Sturm '05)

For a Riemannian manifold $\mathcal{M}$, TFAE:

1. $\mathrm{Ric} \geq \kappa$ everywhere on $\mathcal{M}$;
2. Displacement κ-convexity of the entropy:

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1)$$

for all W_2-geodesics $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(\mathcal{M})$.



Ricci curvature via optimal transport

Theorem (McCann '94)

The entropy is convex along geodesics in $(\mathcal{P}(\mathbb{R}^n), W_2)$.

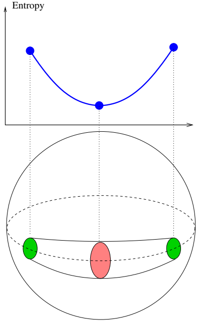
Theorem (Otto–Villani '00, Cordero–McCann–Schmuckenschläger '01, von Renesse–Sturm '05)

For a Riemannian manifold $\mathcal{M}$, TFAE:

1. $\mathrm{Ric} \geq \kappa$ everywhere on $\mathcal{M}$;
2. Displacement κ-convexity of the entropy:

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1)$$

for all W_2-geodesics $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(\mathcal{M})$.



$\rightsquigarrow$ Ricci curvature in metric measure spaces (Lott–Sturm–Villani)

Ricci curvature via optimal transport

Definition (Sturm '06, Lott–Villani '09)

A metric measure space $(\mathcal{X}, d, m)$ satisfies $\mathrm{Ric}(\mathcal{X}) \geq \kappa$ if

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1) .$$

along all W_2-geodesics $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(\mathcal{X})$.

Ricci curvature via optimal transport

Definition (Sturm '06, Lott–Villani '09)

A metric measure space $(\mathcal{X}, d, m)$ satisfies $\mathrm{Ric}(\mathcal{X}) \geq \kappa$ if

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1) \; .$$

along all W_2-geodesics $(\mu_t)_{t \in [0,1]}$ in $\mathcal{P}(\mathcal{X})$.

Crucial features

Ricci curvature via optimal transport

Definition (Sturm '06, Lott–Villani '09)

A metric measure space $(\mathcal{X}, d, m)$ satisfies $\mathrm{Ric}(\mathcal{X}) \geq \kappa$ if

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1) \; .$$

along all W_2-geodesics $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(\mathcal{X})$.

Crucial features

- Applicable to a wide class of metric measure spaces

Ricci curvature via optimal transport

Definition (Sturm '06, Lott–Villani '09)

A metric measure space $(\mathcal{X}, d, m)$ satisfies $\mathrm{Ric}(\mathcal{X}) \geq \kappa$ if

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1) \, .$$

along all W_2-geodesics $(\mu_t)_{t \in [0,1]}$ in $\mathcal{P}(\mathcal{X})$.

Crucial features

- Applicable to a wide class of metric measure spaces
- Many geometric, analytic and probabilistic consequences
  $\longrightarrow$ logarithmic Sobolev, Talagrand, Poincaré inequalities; concentration of measure.

Ricci curvature via optimal transport

Definition (Sturm '06, Lott–Villani '09)

A metric measure space $(\mathcal{X}, d, m)$ satisfies $\mathrm{Ric}(\mathcal{X}) \geq \kappa$ if

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\mathrm{Ent}(\mu_1) - \frac{\kappa}{2} t(1-t) W_2^2(\mu_0, \mu_1) \; .$$

along all W_2-geodesics $(\mu_t)_{t \in [0,1]}$ in $\mathcal{P}(\mathcal{X})$.

Crucial features

- Applicable to a wide class of metric measure spaces
- Many geometric, analytic and probabilistic consequences
  $\longrightarrow$ logarithmic Sobolev, Talagrand, Poincaré inequalities; concentration of measure.
- Stability under measured Gromov–Hausdorff convergence

$\rightsquigarrow$ rich theory, very active research area

What about discrete spaces?

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1 - \alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0, 1]$.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0, 1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\left(\mu_{\alpha(t)}\right)$ is a constant speed geodesic.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$W_2(\mu_{\alpha(t)}, \mu_{\alpha(s)}) = c|t-s|\ .$$

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2\big(\mu_{\alpha(t)}, \mu_{\alpha(s)}\big) = c|t-s|\ .$$

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2\big(\mu_{\alpha(t)}, \mu_{\alpha(s)}\big) = c|t-s|\ .$$

Thus: $t \mapsto \alpha(t)$ is 2-Hölder, hence constant.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0, 1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2(\mu_{\alpha(t)}, \mu_{\alpha(s)}) = c|t - s|\ .$$

  Thus: $t \mapsto \alpha(t)$ is 2-Hölder, hence constant.

- Conclusion: there are no non-trivial W_2-geodesics. In fact:

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2(\mu_{\alpha(t)}, \mu_{\alpha(s)}) = c|t - s|\ .$$

  Thus: $t \mapsto \alpha(t)$ is 2-Hölder, hence constant.
- Conclusion: there are no non-trivial W_2-geodesics. In fact:

$(\mathcal{P}(\mathcal{X}), W_2)$ is a geodesic space $\Leftrightarrow$ $(\mathcal{X}, d)$ is a geodesic space.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1 - \alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0, 1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $\big(\mu_{\alpha(t)}\big)$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2(\mu_{\alpha(t)}, \mu_{\alpha(s)}) = c|t - s|\ .$$

  Thus: $t \mapsto \alpha(t)$ is 2-Hölder, hence constant.

- Conclusion: there are no non-trivial W_2-geodesics. In fact:

$(\mathcal{P}(\mathcal{X}), W_2)$ is a geodesic space $\Leftrightarrow$ $(\mathcal{X}, d)$ is a geodesic space.

Moreover: no curves of finite length $\rightsquigarrow$ no gradient flows.

What about discrete spaces?

Example: 2-point space $\mathcal{X} = \{0, 1\}$.

- Set $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$ for $\alpha \in [0,1]$. Then:

$$W_2(\mu_\alpha, \mu_\beta) = \sqrt{|\alpha - \beta|}\ .$$

- Suppose that $(\mu_{\alpha(t)})$ is a constant speed geodesic. Then:

$$\sqrt{|\alpha(t) - \alpha(s)|} = W_2(\mu_{\alpha(t)}, \mu_{\alpha(s)}) = c|t-s|\ .$$

  Thus: $t \mapsto \alpha(t)$ is 2-Hölder, hence constant.

- Conclusion: there are no non-trivial W_2-geodesics. In fact:

$(\mathcal{P}(\mathcal{X}), W_2)$ is a geodesic space $\Leftrightarrow$ $(\mathcal{X}, d)$ is a geodesic space.

Moreover: no curves of finite length $\rightsquigarrow$ no gradient flows.

Discrete setting

Let $\mathcal{L}$ be generator of a reversible Markov chain on a finite set $\mathcal{X}$:

$$\mathcal{L}\psi(x) = \sum_y Q(x, y)(\psi(y) - \psi(x))$$

Discrete setting

Let $\mathcal{L}$ be generator of a reversible Markov chain on a finite set $\mathcal{X}$:

$$\mathcal{L}\psi(x) = \sum_y Q(x,y)(\psi(y) - \psi(x))$$

Let π be its invariant measure, and consider the relative entropy

$$\mathrm{Ent}_\pi(\mu) := \sum_{x\in\mathcal{X}} \mu(x) \log \frac{\mu(x)}{\pi(x)} \ , \qquad \mu \in \mathcal{P}(\mathcal{X}) \ .$$

Discrete setting

Let $\mathcal{L}$ be generator of a reversible Markov chain on a finite set $\mathcal{X}$:

$$\mathcal{L}\psi(x) = \sum_y Q(x,y)(\psi(y) - \psi(x))$$

Let π be its invariant measure, and consider the relative entropy

$$\mathrm{Ent}_\pi(\mu) := \sum_{x \in \mathcal{X}} \mu(x) \log \frac{\mu(x)}{\pi(x)} \ , \qquad \mu \in \mathcal{P}(\mathcal{X}) \ .$$

Question: Is the Kolmogorov forward equation $\partial_t \mu = \mathcal{L}^\dagger \mu$ the gradient flow of Ent_π w.r.t. a suitable metric on $\mathcal{P}(\mathcal{X})$?

Discrete setting

Let $\mathcal{L}$ be generator of a reversible Markov chain on a finite set $\mathcal{X}$:

$$\mathcal{L}\psi(x) = \sum_y Q(x,y)(\psi(y) - \psi(x))$$

Let π be its invariant measure, and consider the relative entropy

$$\mathrm{Ent}_\pi(\mu) := \sum_{x \in \mathcal{X}} \mu(x) \log \frac{\mu(x)}{\pi(x)} \ , \qquad \mu \in \mathcal{P}(\mathcal{X}) \ .$$

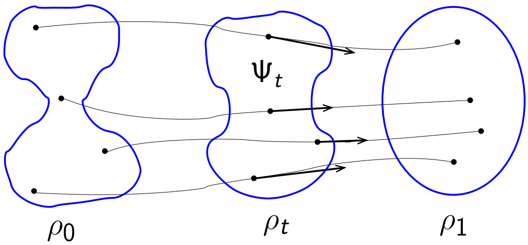
Question: Is the Kolmogorov forward equation $\partial_t \mu = \mathcal{L}^\dagger \mu$ the gradient flow of Ent_π w.r.t. a suitable metric on $\mathcal{P}(\mathcal{X})$?

Teaser: On the two point space: **Yes!**

$$\mathcal{W}(\mu_\alpha, \mu_\beta) = \int_\alpha^\beta \sqrt{\frac{\mathrm{arctanh}(2r-1)}{2r-1}} \, \mathrm{d}r, \qquad 0 \le \alpha \le \beta \le 1.$$

Back to $\mathbb{R}^n$: dynamical characterisation of W_2

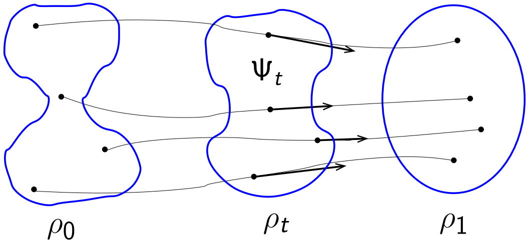
Back to $\mathbb{R}^n$: dynamical characterisation of W_2



Back to R^n: dynamical characterisation of W_2

Benamou-Brenier formula in R^n

$$W_2(\rho_0,\rho_1)^2 = \inf_{(\rho_t,\Psi_t)_t} \left\{ \int_0^1 \int_{\mathsf{R}^n} |\Psi_t(x)|^2 \, \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \ : \ \partial_t \rho + \nabla \cdot (\rho \Psi) = 0 \ , \quad \rho|_{t=0} = \rho_0 \ , \quad \rho|_{t=1} = \rho_1 \right\} .$$



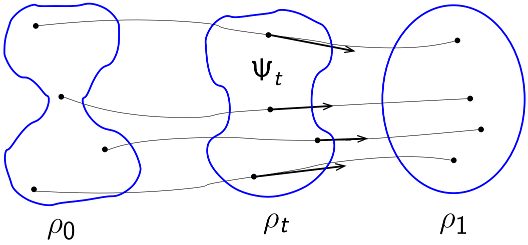
Back to $\mathbb{R}^n$: dynamical characterisation of W_2

Benamou-Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{(\rho_t,\psi_t)_t} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \, \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \ : \right.$$

$$\partial_t \rho + \nabla \cdot (\rho \nabla \psi) = 0 \ ,$$

$$\left. \rho|_{t=0} = \rho_0 \ , \quad \rho|_{t=1} = \rho_1 \right\} .$$



Definition of the metric $\mathcal{W}$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \, \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \operatorname{div}(\rho \nabla \psi) = 0 \text{ and } \rho_{t=0} = \rho_0, \ \rho_{t=1} = \rho_1 \ .$$

Definition of the metric $\mathcal{W}$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \rho_t(x) \,\mathrm{d}x \,\mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \operatorname{div}(\rho \nabla \psi) = 0 \text{ and } \rho_{t=0} = \rho_0, \ \rho_{t=1} = \rho_1 \ .$$

Definition in the discrete case (write $\omega(x, y) = Q(x, y)\pi(x)$)

$$\mathcal{W}(\mu_0, \mu_1)^2$$

$$:= \inf_{\mu,\psi} \left\{ \int_0^1 \sum_{x,y} \big(\psi_t(x) - \psi_t(y)\big)^2 \qquad \omega(x, y) \,\mathrm{d}t \right\}$$

s.t.

Definition of the metric $\mathcal{W}$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \, \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \mathrm{div}(\rho \nabla \psi) = 0 \text{ and } \rho_{t=0} = \rho_0, \; \rho_{t=1} = \rho_1 \; .$$

Definition in the discrete case (write $\omega(x, y) = Q(x, y)\pi(x)$)

$$\mathcal{W}(\mu_0, \mu_1)^2 := \inf_{\mu,\psi} \left\{ \int_0^1 \sum_{x,y} \big(\psi_t(x) - \psi_t(y)\big)^2 \qquad \omega(x, y) \, \mathrm{d}t \right\}$$

s.t.

Problem: μ is defined on vertices, $\nabla\psi$ is defined on edges.

Definition of the metric $\mathcal{W}$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \, \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \mathrm{div}(\rho \nabla \psi) = 0 \text{ and } \rho_{t=0} = \rho_0, \; \rho_{t=1} = \rho_1 \; .$$

Definition in the discrete case (write $\omega(x, y) = Q(x, y)\pi(x)$)

$$\mathcal{W}(\mu_0, \mu_1)^2$$
$$:= \inf_{\mu,\psi} \left\{ \int_0^1 \sum_{x,y} \big(\psi_t(x) - \psi_t(y)\big)^2 \Lambda\left(\frac{\mu_t(x)}{\pi(x)}, \frac{\mu_t(y)}{\pi(y)} \right) \omega(x, y) \, \mathrm{d}t \right\}$$

s.t.

Use the *logarithmic mean* of the densities to define the mobility!

$$\Lambda(a, b) := \int_0^1 a^{1-p} b^p \, \mathrm{d}p \; .$$

Definition of the metric $\mathcal{W}$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \rho_t(x) \, \mathrm{d}x \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \mathrm{div}(\rho \nabla \psi) = 0 \text{ and } \rho_{t=0} = \rho_0, \; \rho_{t=1} = \rho_1 \; .$$

Definition in the discrete case (write $\omega(x,y) = Q(x,y)\pi(x)$)

$$\mathcal{W}(\mu_0, \mu_1)^2 := \inf_{\mu,\psi} \left\{ \int_0^1 \sum_{x,y} \big(\psi_t(x) - \psi_t(y)\big)^2 \Lambda\left(\frac{\mu_t(x)}{\pi(x)}, \frac{\mu_t(y)}{\pi(y)} \right) \omega(x,y) \, \mathrm{d}t \right\}$$

$$\text{s.t. } \partial_t \mu(x) + \sum_y \Lambda\left(\frac{\mu(x)}{\pi(x)}, \frac{\mu(y)}{\pi(y)} \right) \big(\psi(x) - \psi(y)\big) \omega(x,y) = 0 \quad \forall x$$

Use the *logarithmic mean* of the densities to define the mobility!

$$\Lambda(a,b) := \int_0^1 a^{1-p} b^p \, \mathrm{d}p \; .$$

Discrete heat flow as gradient flow

- $\mathcal{W}$ is induced by a Riemannian metric on $\mathcal{P}(\mathcal{X})$.

Discrete heat flow as gradient flow

- $\mathcal{W}$ is induced by a Riemannian metric on $\mathcal{P}(\mathcal{X})$.

Discrete JKO-Theorem (M. , Mielke, Chow-Huang-Li-Zhou '11)

The heat flow is the gradient flow of the entropy w.r.t. $\mathcal{W}$.

Discrete heat flow as gradient flow

- $\mathcal{W}$ is induced by a Riemannian metric on $\mathcal{P}(\mathcal{X})$.

Discrete JKO-Theorem (M. , Mielke, Chow-Huang-Li-Zhou '11)

The heat flow is the gradient flow of the entropy w.r.t. $\mathcal{W}$.

Why the logarithmic mean?

Discrete heat flow as gradient flow

- $\mathcal{W}$ is induced by a Riemannian metric on $\mathcal{P}(\mathcal{X})$.

Discrete JKO-Theorem (M. , Mielke, Chow-Huang-Li-Zhou '11)

The heat flow is the gradient flow of the entropy w.r.t. $\mathcal{W}$.

Why the logarithmic mean?

- Represent heat equation as continuity equation:

$$\partial_t \rho = \Delta \rho \quad \iff \quad \begin{cases} \partial_t \rho + \mathrm{div}(\rho \Psi) = 0 \\ \Psi = -\nabla \log \rho \end{cases}$$

Discrete heat flow as gradient flow

- $\mathcal{W}$ is induced by a Riemannian metric on $\mathcal{P}(\mathcal{X})$.

Discrete JKO-Theorem (M. , Mielke, Chow-Huang-Li-Zhou '11)

The heat flow is the gradient flow of the entropy w.r.t. $\mathcal{W}$.

Why the logarithmic mean?

- Represent heat equation as continuity equation:

$$\partial_t \rho = \Delta \rho \quad \Longleftrightarrow \quad \begin{cases} \partial_t \rho + \mathrm{div}(\rho \Psi) = 0 \\ \Psi = -\nabla \log \rho \end{cases}$$

- Log-mean compensates for the lack of discrete chain rule:

$$\Lambda(\rho(x), \rho(y)) = \int_0^1 \rho(x)^{1-p} \rho(y)^p \, \mathrm{d}p = \frac{\rho(x) - \rho(y)}{\log \rho(x) - \log \rho(y)}$$

Ricci curvature of Markov chains

Definition (à la Lott–Sturm–Villani) (Erbar, M.)

A Markov chain $(\mathcal{X}, Q, \pi)$ is said to have Ricci curvature bounded from below by $\kappa \in \mathbb{R}$ if the relative entropy Ent_π is κ-convex along $\mathcal{W}$-geodesics.

Ricci curvature of Markov chains

Definition (à la Lott–Sturm–Villani) (Erbar, M.)

A Markov chain $(\mathcal{X}, Q, \pi)$ is said to have Ricci curvature bounded from below by $\kappa \in \mathbb{R}$ if the relative entropy Ent_π is κ-convex along $\mathcal{W}$-geodesics.

Simple examples with positive curvature

- discrete hypercube $\{-1, 1\}^n$: $\frac{2}{n}$
- Bernoulli-Laplace model (with k particles on n sites): $\frac{n+2}{k(n-k)}$
- random transposition model on S_n: $\frac{4}{n(n-1)}$
- zero-range processes on complete graph

Ricci curvature of Markov chains

Definition (à la Lott–Sturm–Villani) (Erbar, M.)

A Markov chain $(\mathcal{X}, Q, \pi)$ is said to have Ricci curvature bounded from below by $\kappa \in \mathbb{R}$ if the relative entropy Ent_π is κ-convex along $\mathcal{W}$-geodesics.

Simple examples with positive curvature

- discrete hypercube $\{-1, 1\}^n$: $\frac{2}{n}$
- Bernoulli-Laplace model (with k particles on n sites): $\frac{n+2}{k(n-k)}$
- random transposition model on S_n: $\frac{4}{n(n-1)}$
- zero-range processes on complete graph

Remark Many other notions of discrete Ricci curvature exist, e.g.:

- Ollivier's course Ricci curvature
- Bakry-Émery curvature (in various versions).

Consequences: Sharp functional inequalities

Bakry–Émery Theorem (Erbar, M.)

Let $(\mathcal{X}, Q, \pi)$ be a reversible Markov chain. Let $\kappa > 0$.

If $\mathrm{Ric}(K) \geq \kappa$, then the logarithmic Sobolev inequality holds:

$$\mathrm{Ent}_\pi(\rho\pi) \leq \frac{1}{2\kappa}\mathcal{E}(\rho, \log \rho) \,,$$

where $\mathcal{E}(\varphi, \psi) = -\langle \mathcal{L}\varphi, \psi \rangle_{L^2(\pi)}$ is the Dirichlet form.

This implies exponential decay of the relative entropy:

$$\mathrm{Ent}_\pi(e^{tL^\dagger}\mu) \leq e^{-2\kappa t}\,\mathrm{Ent}_\pi(\mu) \qquad \forall \mu \in \mathcal{P}(\mathcal{X}) \,.$$

Related gradient flow structures

Closely related gradient flow structures have been discovered for

Related gradient flow structures

Closely related gradient flow structures have been discovered for

- Jump processes on general state spaces (Erbar)
  e.g., geodesic convexity of the entropy for Lévy processes

Related gradient flow structures

Closely related gradient flow structures have been discovered for

- Jump processes on general state spaces (Erbar)
  e.g., geodesic convexity of the entropy for Lévy processes
- Nonlocal-interaction equations on graphs
  (Esposito-Patacchini-Schlichting-Slepčev)

Related gradient flow structures

Closely related gradient flow structures have been discovered for

- Jump processes on general state spaces (Erbar)
  e.g., geodesic convexity of the entropy for Lévy processes
- Nonlocal-interaction equations on graphs
  (Esposito-Patacchini-Schlichting-Slepčev)
- Discrete porous medium equations (Erbar-M.)
  allows for structure-preserving discretisations of PDEs

Related gradient flow structures

Closely related gradient flow structures have been discovered for

- Jump processes on general state spaces (Erbar)
  e.g., geodesic convexity of the entropy for Lévy processes
- Nonlocal-interaction equations on graphs
  (Esposito-Patacchini-Schlichting-Slepčev)
- Discrete porous medium equations (Erbar-M.)
  allows for structure-preserving discretisations of PDEs
- Chemical reaction networks (Mielke)
  gradient flow structure for chemical master equation and Liouville equation

Related gradient flow structures

Closely related gradient flow structures have been discovered for

- Jump processes on general state spaces (Erbar)
  e.g., geodesic convexity of the entropy for Lévy processes
- Nonlocal-interaction equations on graphs
  (Esposito-Patacchini-Schlichting-Slepčev)
- Discrete porous medium equations (Erbar-M.)
  allows for structure-preserving discretisations of PDEs
- Chemical reaction networks (Mielke)
  gradient flow structure for chemical master equation and Liouville equation
- Dissipative quantum mechanics (Carlen-M., Mielke-Mittnenzweig, Chen-Gangbo-Georgiou-Tannenbaum)
  non-commutative analogue of $\mathcal{W}$ for density matrices

Is there a JKO theorem for dissipative quantum systems?

Dissipative quantum mechanics

Dynamics of open quantum systems

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space
- Let $\mathfrak{P}(\mathfrak{H}) = \{\rho \in B(\mathfrak{H}) \ : \ \rho = \rho^* \geq 0 \, , \mathrm{Tr}[\rho] = 1\}$ be the set of density matrices

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space
- Let $\mathfrak{P}(\mathfrak{H}) = \{\rho \in B(\mathfrak{H}) \ : \ \rho = \rho^* \geq 0 \ , \mathrm{Tr}[\rho] = 1\}$ be the set of density matrices
- Let $\mathcal{P}_t = e^{t\mathcal{L}}$ be a trace preserving and completely positive semigroup acting on $\mathfrak{P}(\mathfrak{H})$.

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space
- Let $\mathfrak{P}(\mathfrak{H}) = \{\rho \in B(\mathfrak{H}) \ : \ \rho = \rho^* \geq 0 \ , \mathrm{Tr}[\rho] = 1\}$ be the set of density matrices
- Let $\mathcal{P}_t = e^{t\mathcal{L}}$ be a trace preserving and completely positive semigroup acting on $\mathfrak{P}(\mathfrak{H})$.
- Then, $\mathcal{L}$ can be written in Lindblad form

$$\mathcal{L}\rho = -i[H, \rho] + \sum_j \ [V_j, \rho V_j^*] + [V_j \rho, V_j^*] \ ,$$

where the Hamiltonian H is self-adjoint, and $V_j \in B(\mathfrak{H})$.

[Gorini/Kossakowski/Sudarshan, Lindblad '76]

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space
- Let $\mathfrak{P}(\mathfrak{H}) = \{\rho \in B(\mathfrak{H}) \; : \; \rho = \rho^* \geq 0 \; , \mathrm{Tr}[\rho] = 1\}$ be the set of density matrices
- Let $\mathcal{P}_t = e^{t\mathcal{L}}$ be a trace preserving and completely positive semigroup acting on $\mathfrak{P}(\mathfrak{H})$.
- Then, $\mathcal{L}$ can be written in Lindblad form

$$\mathcal{L}\rho = -i[H, \rho] + \sum_j [V_j, \rho V_j^*] + [V_j \rho, V_j^*] \; ,$$

where the Hamiltonian H is self-adjoint, and $V_j \in B(\mathfrak{H})$.

[Gorini/Kossakowski/Sudarshan, Lindblad '76]

- Let $\sigma \in \mathfrak{P}(\mathfrak{H})$ be a stationary state, i.e., $\mathcal{L}\sigma = 0$. Consider the quantum relative entropy $\mathrm{Ent}(\rho|\sigma) = \mathrm{Tr}[\rho(\log \rho - \log \sigma)]$.

Dissipative quantum mechanics

Dynamics of open quantum systems

- Let $\mathfrak{H}$ be a (finite-dimensional) Hilbert space
- Let $\mathfrak{P}(\mathfrak{H}) = \{\rho \in B(\mathfrak{H}) \ : \ \rho = \rho^* \geq 0 \ , \mathrm{Tr}[\rho] = 1\}$ be the set of density matrices
- Let $\mathcal{P}_t = e^{t\mathcal{L}}$ be a trace preserving and completely positive semigroup acting on $\mathfrak{P}(\mathfrak{H})$.
- Then, $\mathcal{L}$ can be written in Lindblad form

  $$\mathcal{L}\rho = -i[H, \rho] + \sum_j \ [V_j, \rho V_j^*] + [V_j \rho, V_j^*] \ ,$$

  where the Hamiltonian H is self-adjoint, and $V_j \in B(\mathfrak{H})$.

  [Gorini/Kossakowski/Sudarshan, Lindblad '76]
- Let $\sigma \in \mathfrak{P}(\mathfrak{H})$ be a stationary state, i.e., $\mathcal{L}\sigma = 0$. Consider the quantum relative entropy $\mathrm{Ent}(\rho|\sigma) = \mathrm{Tr}[\rho(\log \rho - \log \sigma)]$.
- H-Theorem [Spohn '78]: $t \mapsto \mathrm{Ent}(\mathcal{P}_t \rho|\sigma)$ is decreasing.

Gradient flow structures

Gradient flow structures

- Can we formulate the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ as gradient flow of the relative entropy?

Gradient flow structures

- Can we formulate the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ as gradient flow of the relative entropy?
- Assume first: $\mathcal{L}$ is self-adjoint w.r.t. the scalar product

$$\langle A, B \rangle = \mathrm{Tr}[A^* B] \quad \text{on } B(\mathfrak{H}) \ .$$

Gradient flow structures

- Can we formulate the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ as gradient flow of the relative entropy?
- Assume first: $\mathcal{L}$ is self-adjoint w.r.t. the scalar product

$$\langle A, B \rangle = \mathrm{Tr}[A^* B] \quad \text{on } B(\mathfrak{H}) \ .$$

- Then: $\mathcal{L}$ has the divergence form representation

$$\mathcal{L}A = -\sum_j \partial_j^\dagger \partial_j A \quad \text{where} \ \ \partial_j A = [V_j, A] \ .$$

Gradient flow structures

- Can we formulate the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ as gradient flow of the relative entropy?
- Assume first: $\mathcal{L}$ is self-adjoint w.r.t. the scalar product

$$\langle A, B \rangle = \mathrm{Tr}[A^* B] \quad \text{on } B(\mathfrak{H}) \ .$$

- Then: $\mathcal{L}$ has the divergence form representation

$$\mathcal{L}A = -\sum_j \partial_j^\dagger \partial_j A \quad \text{where } \ \partial_j A = [V_j, A] \ .$$

- Ansatz: define a distance $\mathcal{W}$ on $\mathfrak{P}(\mathfrak{H})$ by

$$\mathcal{W}(\rho_0, \rho_1)^2 = \inf_{\rho, A} \left\{ \int_0^1 \sum_j \mathrm{Tr}[(\partial_j A)^* \rho \bullet \partial_j A] \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \textstyle\sum_j \partial_j^\dagger (\rho \bullet \partial_j A) = 0, \quad \rho : \rho_0 \rightsquigarrow \rho_1 \ .$$

Gradient flow structures

- Can we formulate the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ as gradient flow of the relative entropy?
- Assume first: $\mathcal{L}$ is self-adjoint w.r.t. the scalar product

$$\langle A, B \rangle = \mathrm{Tr}[A^* B] \quad \text{on } B(\mathfrak{H}) \ .$$

- Then: $\mathcal{L}$ has the divergence form representation

$$\mathcal{L}A = -\sum_j \partial_j^\dagger \partial_j A \quad \text{where } \ \partial_j A = [V_j, A] \ .$$

- Ansatz: define a distance $\mathcal{W}$ on $\mathfrak{P}(\mathfrak{H})$ by

$$\mathcal{W}(\rho_0, \rho_1)^2 = \inf_{\rho, A} \left\{ \int_0^1 \sum_j \mathrm{Tr}[(\partial_j A)^* \rho \bullet \partial_j A] \, \mathrm{d}t \right\}$$

$$\text{s.t.} \quad \partial_t \rho + \textstyle\sum_j \partial_j^\dagger (\rho \bullet \partial_j A) = 0, \quad \rho : \rho_0 \rightsquigarrow \rho_1 \ .$$

- How to define the product $\bullet$?
  Need: non-commutative version of the classical chain rule

$$\nabla \rho = \rho \nabla \log \rho \ ?$$

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

- Recall that $\partial_j A = [V_j, A]$

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

- Recall that $\partial_j A = [V_j, A]$
- Observe: $\partial_j(AB) = (\partial_j A)B + A\,\partial_j B$

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

- Recall that $\partial_j A = [V_j, A]$
- Observe: $\partial_j(AB) = (\partial_j A)B + A\,\partial_j B$
- Consequently:

$$\partial_j(A^n) = \sum_{k=0}^{n-1} A^k\,(\partial_j A)A^{n-k-1}$$

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

- Recall that $\partial_j A = [V_j, A]$
- Observe: $\partial_j (AB) = (\partial_j A)B + A\,\partial_j B$
- Consequently:

$$\partial_j (A^n) = \sum_{k=0}^{n-1} A^k\,(\partial_j A) A^{n-k-1}$$

- Set $A = \rho^{1/n}$. Then:

$$\partial_j \rho = \sum_{k=0}^{n-1} \rho^{k/n}\,(\partial_j \rho^{1/n}) \rho^{1-(k+1)/n}$$

A non-commutative chain rule

Is there a non-commutative chain rule "$\partial_j \rho = \rho \bullet \partial_j \log \rho$"?

- Recall that $\partial_j A = [V_j, A]$
- Observe: $\partial_j(AB) = (\partial_j A)B + A\,\partial_j B$
- Consequently:

$$\partial_j(A^n) = \sum_{k=0}^{n-1} A^k\,(\partial_j A)A^{n-k-1}$$

- Set $A = \rho^{1/n}$. Then:

$$\partial_j \rho = \sum_{k=0}^{n-1} \rho^{k/n}\,(\partial_j \rho^{1/n})\rho^{1-(k+1)/n}$$

$n \to \infty$:

$$\partial_j \rho = \int_0^1 \rho^s\,(\partial_j \log \rho)\rho^{1-s}\,\mathrm{d}s$$

Quantum JKO

Let $\mathcal{L}$ be a Lindblad operator given by
$\mathcal{L}\rho = \sum_j [V_j, \rho V_j^*] + [V_j\rho, V_j^*]$ with $(V_j)_j \subset B(\mathfrak{H})$.

Quantum JKO

Let $\mathcal{L}$ be a Lindblad operator given by
$\mathcal{L}\rho = \sum_j [V_j, \rho V_j^*] + [V_j \rho, V_j^*]$ with $(V_j)_j \subset B(\mathfrak{H})$.

Definition (Quantum transport metric)

For density matrices $\rho_0, \rho_1 \in \mathfrak{P}(\mathfrak{H})$ we set

$$\mathcal{W}^2(\rho_0, \rho_1) = \inf_{\rho, A} \Bigg\{ \int_0^1 \sum_j \mathrm{Tr}[(\partial_j A)^* \rho \bullet \partial_j A] \, \mathrm{d}t \ : \\ \partial_t \rho + \sum_j \partial_j^\dagger (\rho \bullet \partial_j A) = 0, \quad \rho : \rho_0 \rightsquigarrow \rho_1 \Bigg\}$$

where $\quad \partial_j A = [V_j, A] \quad$ and $\quad \rho \bullet A := \int_0^1 \rho^s \, A \rho^{1-s} \, \mathrm{d}s \ .$

Quantum JKO

Let $\mathcal{L}$ be a Lindblad operator given by
$\mathcal{L}\rho = \sum_j [V_j, \rho V_j^*] + [V_j\rho, V_j^*]$ with $(V_j)_j \subset B(\mathfrak{H})$.

Definition (Quantum transport metric)

For density matrices $\rho_0, \rho_1 \in \mathfrak{P}(\mathfrak{H})$ we set

$$\mathcal{W}^2(\rho_0, \rho_1) = \inf_{\rho, A} \Big\{ \int_0^1 \sum_j \mathrm{Tr}[(\partial_j A)^* \rho \bullet \partial_j A] \, \mathrm{d}t \ : \ \partial_t \rho + \sum_j \partial_j^\dagger (\rho \bullet \partial_j A) = 0, \ \ \rho : \rho_0 \rightsquigarrow \rho_1 \Big\}$$

where $\quad \partial_j A = [V_j, A] \quad$ and $\quad \rho \bullet A := \int_0^1 \rho^s A \rho^{1-s} \, \mathrm{d}s$.

Quantum JKO-Theorem I (Carlen-M. '17)

If $\mathcal{L}$ is self-adjoint w.r.t. trace duality, then the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ is the gradient flow of the von Neumann entropy $\mathrm{Ent}(\rho) = \mathrm{Tr}[\rho \log \rho]$ w.r.t $\mathcal{W}$.

Quantum JKO

Let $\mathcal{L}$ be a Lindblad operator given by
$\mathcal{L}\rho = \sum_j [V_j, \rho V_j^*] + [V_j\rho, V_j^*]$ with $(V_j)_j \subset B(\mathfrak{H})$.

Definition (Quantum transport metric)

For density matrices $\rho_0, \rho_1 \in \mathfrak{P}(\mathfrak{H})$ we set

$$\mathcal{W}^2(\rho_0, \rho_1) = \inf_{\rho, A} \Big\{ \int_0^1 \sum_j \mathrm{Tr}[(\partial_j A)^* \rho \bullet \partial_j A] \,\mathrm{d}t \ : \ \partial_t \rho + \sum_j \partial_j^\dagger (\rho \bullet \partial_j A) = 0, \quad \rho : \rho_0 \rightsquigarrow \rho_1 \Big\}$$

where $\quad \partial_j A = [V_j, A] \quad$ and $\quad \rho \bullet A := \int_0^1 \rho^s A \rho^{1-s} \,\mathrm{d}s$.

Quantum JKO-Theorem I (Carlen-M. '17)

If $\mathcal{L}$ is self-adjoint w.r.t. trace duality, then the Lindblad equation $\partial_t \rho = \mathcal{L}\rho$ is the gradient flow of the von Neumann entropy $\mathrm{Ent}(\rho) = \mathrm{Tr}[\rho \log \rho]$ w.r.t $\mathcal{W}$.

Quantum JKO: the general case

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B \rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B \rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.
- Then: $\mathcal{L} = \sum_j e^{\omega_j/2} \mathcal{L}_j$, $\qquad \mathcal{L}_j \rho = [V_j, \rho V_j^*] + [V_j \rho, V_j^*]$, where $\{V_j\} = \{V_j^*\}$ and $[V_j, \log \sigma] = -\omega_j V_j$ for some $\omega_j \in \mathrm{R}$.

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B\rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.
- Then: $\mathcal{L} = \sum_j e^{\omega_j/2}\mathcal{L}_j$, $\qquad \mathcal{L}_j\rho = [V_j, \rho V_j^*] + [V_j\rho, V_j^*]$, where $\{V_j\} = \{V_j^*\}$ and $[V_j, \log\sigma] = -\omega_j V_j$ for some $\omega_j \in \mathrm{R}$.
- Need: a non-commutative chain rule of the form
$$\sigma\nabla(\rho/\sigma) = \rho\nabla(\log\rho - \log\sigma)$$

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B \rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.
- Then: $\mathcal{L} = \sum_j e^{\omega_j/2} \mathcal{L}_j$, $\quad \mathcal{L}_j \rho = [V_j, \rho V_j^*] + [V_j \rho, V_j^*]$, where $\{V_j\} = \{V_j^*\}$ and $[V_j, \log \sigma] = -\omega_j V_j$ for some $\omega_j \in \mathrm{R}$.
- Need: a non-commutative chain rule of the form
$$\sigma \nabla(\rho/\sigma) = \rho \nabla(\log \rho - \log \sigma)$$
- Key identity:
$$\sigma^{1/2} \partial_j \big(\sigma^{-1/2} \rho \sigma^{-1/2}\big) \sigma^{1/2} = \rho \bullet_j \big(\partial_j(\log \rho - \log \sigma)\big) \ ,$$

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B \rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.
- Then: $\mathcal{L} = \sum_j e^{\omega_j/2} \mathcal{L}_j$, $\qquad \mathcal{L}_j \rho = [V_j, \rho V_j^*] + [V_j \rho, V_j^*]$, where $\{V_j\} = \{V_j^*\}$ and $[V_j, \log \sigma] = -\omega_j V_j$ for some $\omega_j \in \mathrm{R}$.
- Need: a non-commutative chain rule of the form
$$\sigma \nabla(\rho/\sigma) = \rho \nabla(\log \rho - \log \sigma)$$
- Key identity:
$$\sigma^{1/2} \partial_j (\sigma^{-1/2} \rho \sigma^{-1/2}) \sigma^{1/2} = \rho \bullet_j \big(\partial_j (\log \rho - \log \sigma)\big) ,$$
where $\qquad \rho \bullet_j A = \int_0^1 \big(e^{-\omega_j/2}\rho\big)^{1-s} A \big(e^{\omega_j/2}\rho\big)^s \, \mathrm{d}s$

Quantum JKO: the general case

- Assume: $\mathcal{L}$ satisfies detailed balance with respect to $\sigma \in \mathfrak{P}(\mathfrak{H})$, i.e., $\mathcal{L}$ is selfadjoint w.r.t. $\langle A, B\rangle_\sigma = \mathrm{Tr}[\sigma A^* B]$.

- Then: $\mathcal{L} = \sum_j e^{\omega_j/2}\mathcal{L}_j$, $\qquad \mathcal{L}_j\rho = [V_j, \rho V_j^*] + [V_j\rho, V_j^*]$,
where $\{V_j\} = \{V_j^*\}$ and $[V_j, \log\sigma] = -\omega_j V_j$ for some $\omega_j \in \mathrm{R}$.

- Need: a non-commutative chain rule of the form
$$\sigma\nabla(\rho/\sigma) = \rho\nabla(\log\rho - \log\sigma)$$

- Key identity:
$$\sigma^{1/2}\partial_j\big(\sigma^{-1/2}\rho\sigma^{-1/2}\big)\sigma^{1/2} = \rho \bullet_j \big(\partial_j(\log\rho - \log\sigma)\big) ,$$
where $$\rho \bullet_j A = \int_0^1 \big(e^{-\omega_j/2}\rho\big)^{1-s} A\big(e^{\omega_j/2}\rho\big)^s \,\mathrm{d}s$$

Quantum JKO-Theorem II (Carlen-M., Mielke-Mittnenzweig 2017)

If $\mathcal{L}$ satisfies detailed balance w.r.t. a state σ, then the Lindblad equation $\partial_t\rho = \mathcal{L}\rho$ is the gradient flow of the quantum relative entropy $\mathrm{Ent}(\rho|\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]$ w.r.t $\mathcal{W}$.

Example: the quantum Ornstein-Uhlenbeck semigroup

Example: the quantum Ornstein-Uhlenbeck semigroup

- Let a be an operator satisfying the canonical commutation relation $[a, a^*] = I$

Example: the quantum Ornstein-Uhlenbeck semigroup

- Let a be an operator satisfying the canonical commutation relation $[a, a^*] = I$
- Concrete realisation: $\mathfrak{H} = L^2(\mathrm{R}, \gamma)$, γ Gaussian measure,
$$a = \partial_x, \quad a^* = x - \partial_x$$

Example: the quantum Ornstein-Uhlenbeck semigroup

- Let a be an operator satisfying the canonical commutation relation $[a, a^*] = I$
- Concrete realisation: $\mathfrak{H} = L^2(\mathrm{R}, \gamma)$, γ Gaussian measure,
$$a = \partial_x, \quad a^* = x - \partial_x$$
- For $\beta > 0$, consider the quantum OU-operator

$$\mathcal{L}_\beta \rho = \frac{1}{2} e^{\beta/2} \underbrace{\Big([a, \rho a^*] + [a\rho, a^*]\Big)}_{\textit{attenuator}} + \frac{1}{2} e^{-\beta/2} \underbrace{\Big([a^*, \rho a] + [a^* \rho, a]\Big)}_{\textit{amplifier}}$$

Example: the quantum Ornstein-Uhlenbeck semigroup

- Let a be an operator satisfying the canonical commutation relation $[a, a^*] = I$
- Concrete realisation: $\mathfrak{H} = L^2(\mathrm{R}, \gamma)$, γ Gaussian measure,
$$a = \partial_x, \qquad a^* = x - \partial_x$$
- For $\beta > 0$, consider the quantum OU-operator

$$\mathcal{L}_\beta \rho = \frac{1}{2} e^{\beta/2} \underbrace{\Big([a, \rho a^*] + [a\rho, a^*]\Big)}_{attenuator} + \frac{1}{2} e^{-\beta/2} \underbrace{\Big([a^*, \rho a] + [a^* \rho, a]\Big)}_{amplifier}$$

- $\exists!$ (Gaussian) stationary state: $\sigma_\beta = Z^{-1} e^{-\beta H}$, $H = a^* a$

Theorem: [Carlen/M. '17]

$$\mathrm{Ent}(e^{t\mathcal{L}_\beta} \rho | \sigma_\beta) \leq e^{-2\lambda_\beta t}\, \mathrm{Ent}(\rho | \sigma_\beta) \quad \text{where} \quad \lambda_\beta = \sinh(\beta/2) \,.$$

Geodesic convexity of the quantum entropy

Key ingredients of the proof:

Geodesic convexity of the quantum entropy

Key ingredients of the proof:

- Intertwining relations:

$$\partial_j \circ \mathcal{P}_t = e^{-\lambda_\beta t} \mathcal{P}_t \circ \partial_j$$

where $\lambda_\beta = \sinh(\beta/2)$

Geodesic convexity of the quantum entropy

Key ingredients of the proof:

- Intertwining relations:

$$\partial_j \circ \mathcal{P}_t = e^{-\lambda_\beta t} \mathcal{P}_t \circ \partial_j$$

  where $\lambda_\beta = \sinh(\beta/2)$

- Matrix convexity inequalities:

$$(R, A) \mapsto \mathrm{Tr}\left[\int_0^\infty (tI + e^{-\omega/2} R)^{-1} A^* (tI + e^{\omega/2} R)^{-1} A \, \mathrm{d}t \right]$$

  is jointly convex on $\mathcal{M}_n^+ \times \mathcal{M}_n$ for all $\omega \in \mathrm{R}$.

Thank you!