On dissipation distances for reaction-diffusion equations — the Hellinger-Kantorovich distance

Matthias Liero, joint work with Alexander Mielke, Giuseppe Savaré
Outline

Introduction to Hellinger-Kantorovich distance

- Wasserstein distance on the space of probability measures
- Dissipation distance induced by reaction-diffusion problem
- Example: Optimal relocation of mass points
- Main results

Thorough discussion in Giuseppe’s talk
**The Wasserstein distance**

- **Kantorovich-Wasserstein distance** on space of prob. meas.

\[ \mathcal{W}_2(\mu_0, \mu_1)^2 = \min \left\{ \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \, d\eta(x, y) \right\} \]

Transport plan

\[ \eta(\cdot \times \Omega) = \mu_0, \quad \eta(\Omega \times \cdot) = \mu_1 \]

- **Equivalent dynamical formulation** (Benamou-Brenier-2000)

\[ \mathcal{W}_2(\mu_0, \mu_1)^2 = \min \left\{ \int_0^1 \int_{\mathbb{R}^d} |\mathbf{v}_t(x)|^2 \, d\mu_t(x) \, dt \right\} \]

continuity equation

\[ \partial_t \mu_t + \nabla \cdot (\mathbf{v}_t \mu_t) = 0 \]
The Wasserstein distance

Otto’s Riemannian interpretation

\[ g_\mu (\dot{\mu}, \dot{\mu}) = \int_{\mathbb{R}^d} \nabla \xi_\mu \cdot \nabla \xi_\mu \, d\mu \quad \dot{\mu} = -\nabla \cdot (\nabla \xi_\mu \mu) \]

Wasserstein gradient

\[ \mathcal{E}(u) = \int_{\Omega} E(u) \, dx \quad \text{grad}_W \mathcal{E}(u) = -\nabla \cdot (u \nabla \frac{\delta \mathcal{E}}{\delta u}) \]

Wasserstein gradient flows

\[ \partial_t u = -\text{grad}_W \mathcal{E}(u) \quad \partial_t u = \nabla \cdot (u E''(u) \nabla u) \]

see Ambrosio-Gigli-Savaré-2005
Generalizations of Wasserstein

Mielke-2011: Onsager operators

\[ -\nabla \cdot (\mu \nabla \xi) \quad \Rightarrow \quad K(u)\xi = -\nabla \cdot (M(u)\nabla \xi) + H(u)\xi \]

(e.g. reversible mass-action kinetics, semiconductor eqns.)

Can we characterize the associated reaction-diffusion distance?

General case too hard, consider

\[ K_{\alpha \beta}(u)\xi = -\nabla \cdot (\alpha u \nabla \xi) + \beta u \xi \]

Understanding geometry gives results for scalar RD equation

\[ \partial_t u = -\text{grad}_K \mathcal{E}(u) \]

\[ \partial_t u = \nabla \cdot (\alpha u E''(u) \nabla u) - \beta u E'(u) \]
**Competition between reaction/diffusion**

Reaction-diffusion distance

\[
D_{\alpha\beta}(\mu_0, \mu_1)^2 = \inf \left\{ \int_0^1 \int_\Omega \left( \alpha |\nabla \xi_t|^2 + \beta \xi_t^2 \right) d\mu_t dt \right\}
\]

generalized cont. eq.

\[
\partial_t \mu_t + \nabla \cdot (\alpha \nabla \xi_t \mu_t) = \beta \xi_t \mu_t
\]

Benamou-Brenier-trick

\[
v_t = \alpha \nabla \xi_t \quad w_t = \beta \xi_t
\]

inf-convolution of Wasserstein-Kantorovich and Hellinger distance

\[
D_{\alpha 0}(\mu_0, \mu_1) = \frac{1}{\sqrt{\alpha}} \mathcal{W}_2(\mu_0, \mu_1) \quad D_{0\beta}(\mu_0, \mu_1)^2 = \frac{4}{\beta} \int_\Omega \left( \sqrt{f_0} - \sqrt{f_1} \right)^2 d\lambda
\]

Competition between optimal transport and absorption/generation
Competition between reaction/diffusion

Reaction-diffusion distance

\[ D_{\alpha\beta}(\mu_0, \mu_1)^2 = \inf \left\{ \int_0^1 \int_\Omega \left( \frac{|\boldsymbol{v}_t|^2}{\alpha} + \frac{w_t^2}{\beta} \right) d\mu_t dt \right\} \]

generalized cont. eq. \n\[ \partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \]

Benamou-Brenier-trick \n\[ \boldsymbol{v}_t = \alpha \nabla \xi_t \quad w_t = \beta \xi_t \]

inf-convolution of Wasserstein-Kantorovich and Hellinger distance

\[ D_{\alpha\beta}(\mu_0, \mu_1) = \frac{1}{\sqrt{\alpha}} \mathcal{W}_2(\mu_0, \mu_1) \quad D_{0\beta}(\mu_0, \mu_1)^2 = \frac{4}{\beta} \int_\Omega (\sqrt{f_0} - \sqrt{f_1})^2 d\lambda \]

Competition between optimal transport and absorption/generation
Two mass points

\[ \mu_0 = a_0 \delta_{x_0} \]

Optimal relocation

\[ \mu_1 = a_1 \delta_{x_1} \]

Naive approach

\[ b_0(s) \]

\[ b_1(s) \]

\[ c(s) \]
Two mass points

Consider only curves of the form

\[ \mu_t = b_0(t)\delta_{x_0} + c(t)\delta_{x(t)} + b_1(t)\delta_{x_1} \]

Initial and final conditions

\[ b_0(0) + c(0) = a_0, \quad b_1(0) = 0 \]
\[ b_1(1) + c(1) = a_1, \quad b_0(1) = 0 \]
\[ x(0) = x_0, \quad x(1) = x_1 \]

Solve gen. continuity equation

\[ \partial_t \mu_t + \nabla \cdot (\upsilon_t \mu_t) - w_t \mu_t = 0 \]

\[ w_t(x_i) = \frac{\dot{b}_i(t)}{b_i(t)}, \quad w_t(x(t)) = \frac{\dot{c}(s)}{c(s)}, \quad \upsilon_t(x(t)) = \dot{x}(t) \]
Two mass points

Plug into action functional and solve Euler-Lagrange equations

Absorption/Generation

\[ b_0(t) = (1-t)^2 b_0(0), \quad b_1(t) = t^2 b_1(1) \]

Transport

\[ R_{\alpha\beta} := \sqrt{\frac{\beta}{4\alpha}} |x_0 - x_1| \leq \pi \]

\[ c(t) = (1-t)^2 c(0) + t^2 c(1) + 2t(1-t) \sqrt{c(0)c(1)} \cos(R_{\alpha\beta}) \]

\[ x(t) = (1-\theta(t)) x_0 + \theta(t) x_1 \]

Speed of transport determined by

\[ \theta(s) = \frac{1}{R_{\alpha\beta}} \arctan \left( \frac{s \sqrt{c(1)} \sin(R_{\alpha\beta})}{(1-s) \sqrt{c(0)} + s \sqrt{c(1)} \cos(R_{\alpha\beta})} \right) \]
Two mass points

$$\mathcal{A}(b_0, b_1, c, x) = \frac{4}{\beta} \left\{ b_0(0) + c(0) + b_1(1) + c(1) - 2\sqrt{c(0)c(1) \cos(R_{\alpha\beta})} \right\}$$

Minimize with respect to initial and final distribution of mass

$$R_{\alpha\beta} < \frac{\pi}{2} \quad \text{Make } c(0) \text{ and } c(1) \text{ as big as possible}$$

$$c(0) = a_0, \quad c(1) = a_1, \quad b_0(0) = b_1(1) = 0$$

$$\min \mathcal{A} = \frac{4}{\beta} (a_0 + a_1 - 2\sqrt{a_0a_1 \cos(R_{\alpha\beta})})$$

$$R_{\alpha\beta} > \frac{\pi}{2} \quad \text{Make } c(0) \text{ and } c(1) \text{ as small as possible}$$

$$b_0(0) = a_0, \quad b_1(1) = a_1, \quad c(0) = c(1) = 0$$

$$\min \mathcal{A} = \frac{4}{\beta} (a_0 + a_1)$$

Hellinger-Kantorovich distance
Observations

- Sharp threshold
  \[ |x_0 - x_1| < \sqrt{\frac{\alpha}{\beta}} \pi \]  only (mixed) transport

  \[ |x_0 - x_1| > \sqrt{\frac{\alpha}{\beta}} \pi \]  only absorp./gen.

- Speed of transport depends on initial and final mass
- Mass is not conserved along transport
Observations

- **Sharp threshold**

  \[ |x_0 - x_1| < \sqrt{\frac{\alpha}{\beta}} \pi \]
  only (mixed) transport

  \[ |x_0 - x_1| > \sqrt{\frac{\alpha}{\beta}} \pi \]
  only absorp./gen.

- **Speed of transport depends on initial and final mass**

- **Mass is not conserved along transport**
Observations

- Sharp threshold
  \[ |x_0 - x_1| < \sqrt{\frac{\alpha}{\beta}} \pi \] only (mixed) transport
  \[ |x_0 - x_1| > \sqrt{\frac{\alpha}{\beta}} \pi \] only absorp./gen.

- Speed of transport depends on initial and final mass
- Mass is not conserved along transport
The cone distance

\[ D(x_0, a_0; x_1, a_1) := \frac{4}{\beta} \left( a_0 + a_1 - 2\sqrt{a_0a_1} \cos(R_{\alpha\beta}) \right) \]

gives distance on cone space over \( \Omega \)

Consider fiber \([0, \infty]\) in each point
The cone distance

Identify all points in $\Omega \times \{0\}$
The cone distance

\[ \mathcal{E}_\Omega = \left( \Omega \times [0, \infty[ \right) / (\Omega \times \{0\}) \]

\[ d_{\text{Cone}}([x_0, a_0], [x_1, a_1])^2 = \begin{cases} a_0 + a_1 - 2\sqrt{a_0a_1} \cos(R_{\alpha\beta}) & R_{\alpha\beta} < \pi \\ (\sqrt{a_0} + \sqrt{a_1})^2 & R_{\alpha\beta} \geq \pi \end{cases} \]

Law of cosines:

\[ R_{\alpha\beta} := \sqrt{\frac{\beta}{4\alpha}} |x_0 - x_1| \]

Solutions of Euler-Lagrange eqns are geodesics in cone space
Two mass points

**Transport:** $|x_0 - x_1| < \sqrt{\frac{\alpha}{\beta}} \pi$

cost of transport given by

$$\frac{4}{\beta} d_{\text{Cone}}([x_0, a_0], [x_1, a_1])^2 = \frac{4}{\beta} \left( a_0 + a_1 - 2\sqrt{a_0 a_1} \cos \left( \sqrt{\beta/(4\alpha)} |x_0 - x_1| \right) \right)$$

**Absorption/generation:** $|x_0 - x_1| > \sqrt{\frac{\alpha}{\beta}} \pi$

cost of absorption/generation given by

$$\frac{4}{\beta} d_{\text{Cone}}([x_0, a_0], 0 e)^2 + \frac{4}{\beta} d_{\text{Cone}}(0 e, [x_1, a_1])^2 = \frac{4}{\beta} (a_0 + a_1)$$
Extension to general measures

Optimal relocation in $\Omega$ induced by optimal transport in $\mathcal{C}_\Omega$

$$|x_0 - x_1| < \sqrt{\frac{\alpha}{\beta}} \pi$$

$$\nu_0 = \delta_{0\Omega} + \delta_{[x_0,a_0]} \quad \nu_1 = \delta_{0\Omega} + \delta_{[x_1,a_1]}$$

$$|x_0 - x_1| > \sqrt{\frac{\alpha}{\beta}} \pi$$

Connection to measures on $\Omega$ provided by projection

$$\mathcal{P} : \mathcal{M}_{\geq 0}(\mathcal{C}_\Omega) \to \mathcal{M}_{\geq 0}(\Omega)$$

$$\int_{\Omega} \varphi(x) \, d\mathcal{P} \nu_i = \int_{\mathcal{C}_\Omega} \hat{a}(z) \varphi(\hat{x}(z)) \, d\nu_i(z)$$

Define $\text{HK}(\mu_0, \mu_1) = \frac{2}{\sqrt{\beta}} \min \{ \mathcal{W}_{\mathcal{C}_\Omega}(\nu_0, \nu_1) \mid \mu_i = \mathcal{P} \nu_i \}$
Main result

Theorem [L.-Mielke-Savaré-2014]

\[
\HK(\mu_0, \mu_1)^2 = \min_{\nu \in \mathcal{H}(\mu_0, \mu_1)} \int \int d_{\text{Cone}}(z_0, z_1)^2 \, d\nu
\]

\[
\nu = \mathcal{H}(\mu_0, \mu_1) \quad \iff \quad \mathcal{P}(\nu_i) = \mu_i
\]

- defines distance on \( \mathcal{M}_{\geq 0}(\Omega) \)
- complete, separable metric space
- metrizes the weak topology
- minimizer exists
- geodesic curves induced by geodesics in \( \mathcal{C}_\Omega \)
- equivalent to dynamical formulation
Another equivalent characterization

**Theorem [L.-Mielke-Savaré-2014]**

**Cost function**

\[
\Gamma_{\alpha\beta}(R) = \begin{cases} 
-2\log \left( \cos \left( \sqrt{\frac{\beta}{4\alpha}} R \right) \right) & R < \pi \sqrt{\frac{\alpha}{\beta}} \\
\infty & R \geq \pi \sqrt{\frac{\alpha}{\beta}}
\end{cases}
\]

**Boltzmann function**

\[
F(z) = z \log z - z + 1
\]

**Equivalent formulation**

\[
\text{HK}(\mu_0, \mu_1)^2 = \frac{4}{\beta} \min \left\{ \int_{\Omega} F\left(\frac{d\eta_0}{d\mu_0}\right) d\mu_0 + \int_{\Omega} F\left(\frac{d\eta_1}{d\mu_1}\right) d\mu_1 + \int_{\Omega \times \Omega} \Gamma_{\alpha\beta}(\|x_0-x_1\|) d\eta \right\}
\]

**Marginals**

\[
\eta_i = P^i_\# \eta \ll \mu_i
\]
Geodesic curves