A probabilistic view on unbalanced optimal transport

Hugo Lavenant[a]

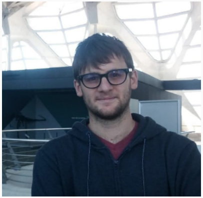
Decemeber 7, 2022

ISTA, Klosterneuburg, Austria

[a]Department of Decision Sciences, Bocconi University

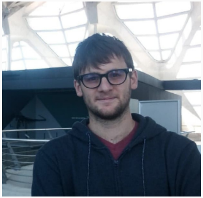
My coauthor

Joint work with **Aymeric Baradat** (Université Claude Bernard Lyon 1).



My coauthor

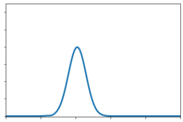
Joint work with **Aymeric Baradat** (Université Claude Bernard Lyon 1).

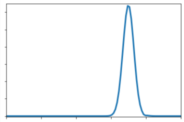


Disclaimer

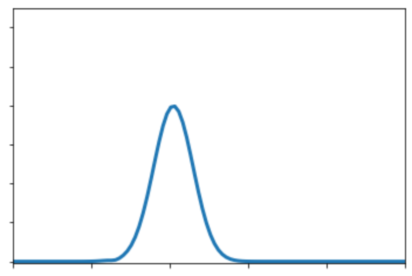
He is the one who knows about probability!

Optimal Transport

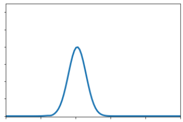




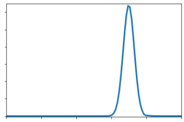
initial final



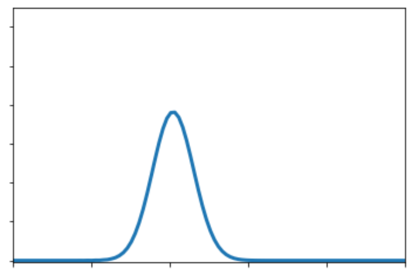
Regularized Optimal Transport



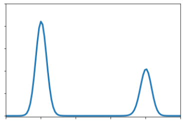
initial

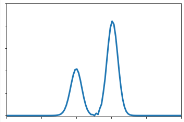


final

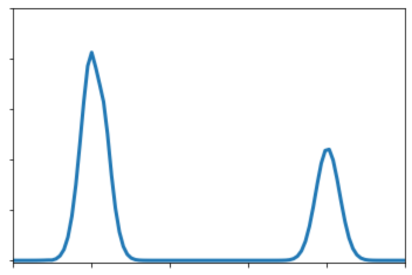


With bimodal inputs

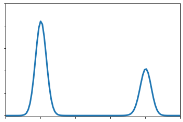


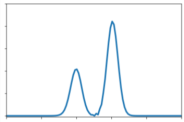


initial final

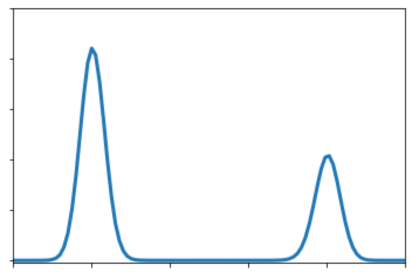


Solution: Unbalanced Optimal Transport

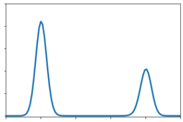


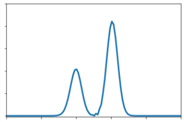


initial final

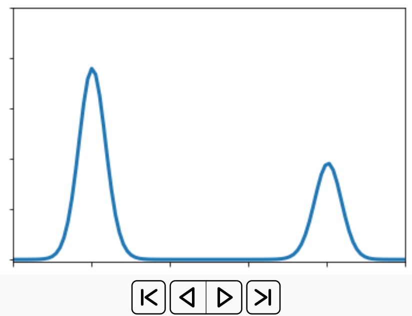


Today: Regularized Unbalanced Optimal Transport



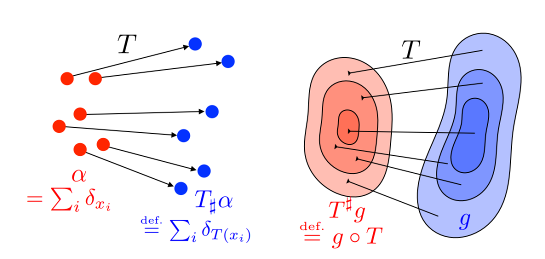


initial final



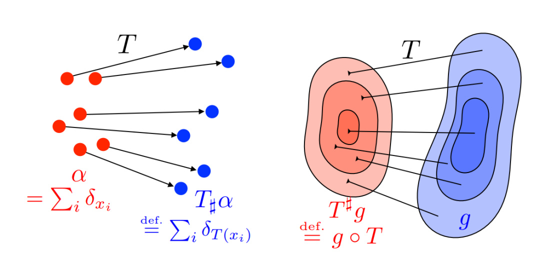
What is this talk about?

Regularized (a.k.a. entropic) Optimal Transport...

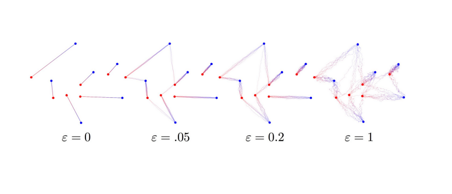


What is this talk about?

Regularized (a.k.a. entropic) Optimal Transport...

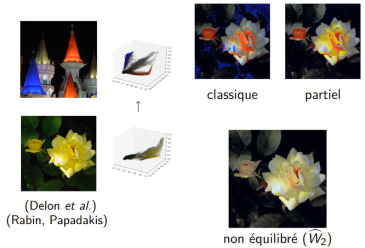


... as entropy minimization w.r.t. the law of Brownian Motion



What is this talk about?

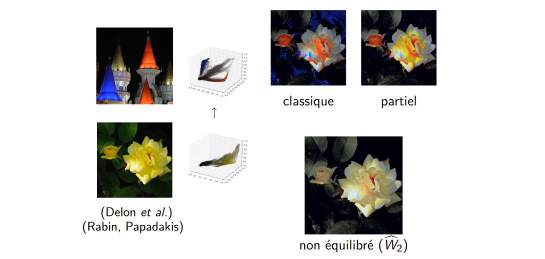
Regularized (a.k.a. entropic) Unbalanced Optimal Transport...



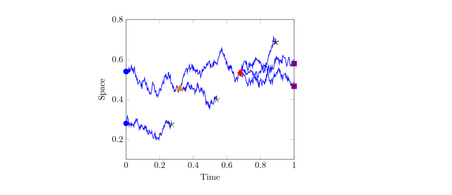
... as entropy minimization w.r.t. the law of **??**

What is this talk about?

Regularized (a.k.a. entropic) Unbalanced Optimal Transport...



... as entropy minimization w.r.t. the law of Branching Brownian Motion



Today

Goal of this presentation

Show an **equivalence** between two problems of calculus of variations:

- The dynamical formulation (a.k.a Benamou Brenier formulation) of **regularized unbalanced optimal transport**.
- Entropy minimization with respect to the law of **branching Brownian Motion** (“Branching Schrödinger problem”).

Outline

1. The Schrödinger problem

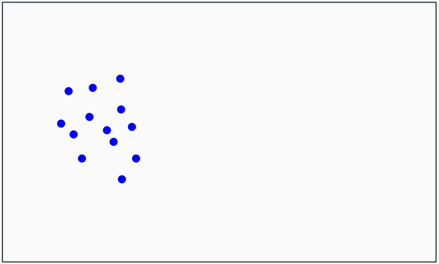
2. The branching Schrödinger problem

1. The Schrödinger problem

- Léonard (2013): A survey of the Schrödinger problem and some of its connections with optimal transport;
- Gentil, Léonard, and Ripani (2017): About the analogy between optimal transport and minimal entropy.

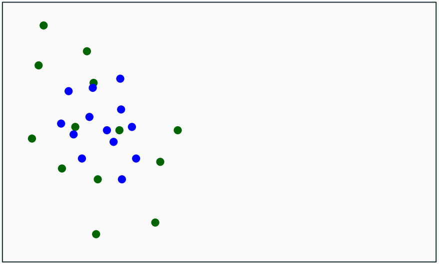
A problem coming from Large deviation

N particles $\sim \alpha$ at time $t = 0$. They follow **Brownian motion**.



A problem coming from Large deviation

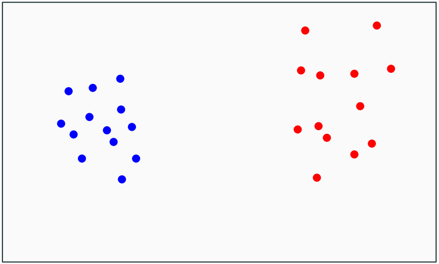
N particles $\sim \alpha$ at time $t = 0$. They follow **Brownian motion**.



Expected distribution at time $t = 1$, $\sim \mathcal{N}(0,1) \star \alpha$.

A problem coming from Large deviation

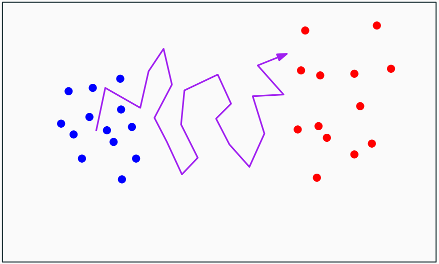
N particles $\sim \alpha$ at time $t = 0$. They follow **Brownian motion**.



Observed distribution at time $t = 1$, $\beta \neq \mathcal{N}(0,1) \star \alpha$.

A problem coming from Large deviation

N particles $\sim \alpha$ at time $t = 0$. They follow **Brownian motion**.



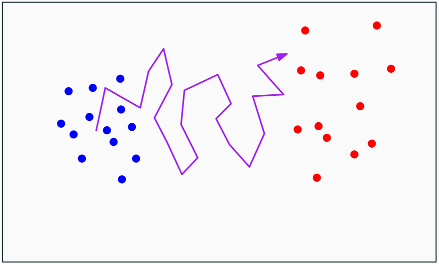
Observed distribution at time $t = 1$, $\beta \neq \mathcal{N}(0,1) \star \alpha$.

The problem

If $N \gg 1$, given this unlikely event, what is the most likely evolution?

A problem coming from Large deviation

N particles $\sim \alpha$ at time $t = 0$. They follow **Brownian motion**.



Observed distribution at time $t = 1$, $\beta \neq \mathcal{N}(0,1) \star \alpha$.

The problem

If $N \gg 1$, given this unlikely event, what is the most likely evolution?

Theory of Large Deviation: **entropy minimization** with respect to the law of Brownian motion.

Schrödinger problem and Regularized Optimal Transport

State space $\mathbb{T}^d$ the d-dimensional torus, $\alpha, \beta \in \mathcal{P}(\mathbb{T}^d)$ and $\nu > 0$.

Schrödinger problem and Regularized Optimal Transport

State space $\mathbb{T}^d$ the d-dimensional torus, $\alpha, \beta \in \mathcal{P}(\mathbb{T}^d)$ and $\nu > 0$.

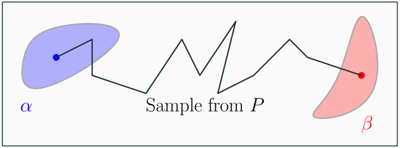
Space $\Omega = C([0,1], \mathbb{T}^d)$. $R^\nu \in \mathcal{P}(\Omega)$ Wiener measure with diffusivity ν and $X_0 \sim \mathcal{L} = \mathrm{d}x$ under R^ν.

The Schrödinger problem

Given $\alpha, \beta \in \mathcal{P}(\mathbb{T}^d)$, find $P \in \mathcal{P}(\Omega)$ which minimizes

$$H(P|R^\nu) := \int_\Omega \log\left(\frac{\mathrm{d}P}{\mathrm{d}R^\nu}(X)\right) \mathrm{d}P(X).$$

such that $X_0 \sim \alpha$ and $X_1 \sim \beta$ under P.



Schrödinger problem and Regularized Optimal Transport

State space $\mathbb{T}^d$ the d-dimensional torus, $\alpha, \beta \in \mathcal{P}(\mathbb{T}^d)$ and $\nu > 0$.

Space $\Omega = C([0,1], \mathbb{T}^d)$. $R^\nu \in \mathcal{P}(\Omega)$ Wiener measure with diffusivity ν and $X_0 \sim \mathcal{L} = \mathrm{d}x$ under R^ν.

The Schrödinger problem

Given $\alpha, \beta \in \mathcal{P}(\mathbb{T}^d)$, find $P \in \mathcal{P}(\Omega)$ which minimizes

$$H(P|R^\nu) := \int_\Omega \log\left(\frac{\mathrm{d}P}{\mathrm{d}R^\nu}(X)\right) \mathrm{d}P(X).$$

such that $X_0 \sim \alpha$ and $X_1 \sim \beta$ under P.

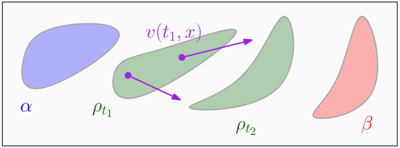
Regularized Optimal Transport

Look for ρ and v time-dependent density and velocity field which minimize

$$\mathcal{A}(\rho, v) = \int_0^1 \int_{\mathbb{T}^d} \frac{|v(t,x)|^2}{2} \rho(t,x)\, \mathrm{d}t\mathrm{d}x$$

such that $\rho_0 = \alpha$, $\rho_1 = \beta$ and

$$\partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2} \Delta \rho$$



Equivalence between the problems

Both problems are well-posed if $H(\alpha|\mathcal{L}), H(\beta|\mathcal{L}) < +\infty$.

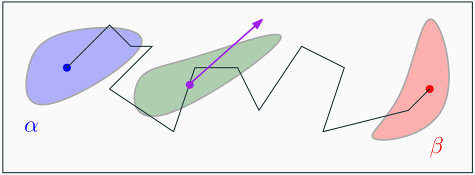
From Schrödinger to ROT

Given $P \in \mathcal{P}(\Omega)$ with $H(P|R^\nu) < +\infty$, define $\rho_t := \mathrm{Law}_P(X_t)$,

$$v(t, X_t) := \lim_{h \to 0, h > 0} \mathbb{E}_P \left[\frac{X_{t+h} - X_t}{h} \middle| X_t \right].$$

Then (ρ, v) admissible and

$$\nu H(\alpha|\mathcal{L}) + \mathcal{A}(\rho, v) \leqslant \nu H(P|R^\nu).$$



Equivalence between the problems

Both problems are well-posed if $H(\alpha|\mathcal{L}), H(\beta|\mathcal{L}) < +\infty$.

From Schrödinger to ROT

Given $P \in \mathcal{P}(\Omega)$ with $H(P|R^\nu) < +\infty$, define $\rho_t := \mathrm{Law}_P(X_t)$,

$$v(t, X_t) := \lim_{h \to 0, h > 0} \mathbb{E}_P \left[\frac{X_{t+h} - X_t}{h} \middle| X_t \right].$$

Then (ρ, v) admissible and

$$\nu H(\alpha|\mathcal{L}) + \mathcal{A}(\rho, v) \leqslant \nu H(P|R^\nu).$$

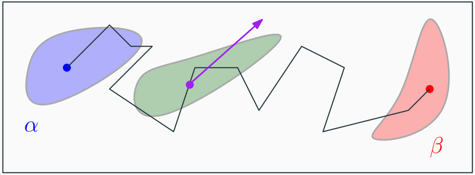
From ROT to Schrödinger

If (ρ, v) admissible with v smooth, P the law of the SDE

$$\mathrm{d}X_t = v(t, X_t)\,\mathrm{d}t + \sqrt{\nu}\,\mathrm{d}B_t.$$

Then P admissible and

$$\nu H(\alpha|\mathcal{L}) + \mathcal{A}(\rho, v) = \nu H(P|R^\nu).$$



Consequence: equality of the values

Theorem

For any α, β with $H(\alpha|\mathcal{L}), H(\beta|\mathcal{L}) < +\infty$, there holds

$$\nu H(\alpha|\mathcal{L}) + \min_{\rho, v} \left\{ \mathcal{A}(\rho, v) \ : \ \partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2} \Delta \rho, \ \rho_0 = \alpha, \rho_1 = \beta \right\}$$
$$= \min_{P} \left\{ \nu H(P|R^\nu) \ : \ X_0 \sim \alpha \text{ and } X_1 \sim \beta \text{ under } P \right\}.$$

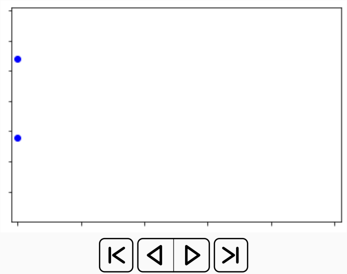
Moreover, if (ρ, v) and P optimal then P is the law of the SDE with drift v.

2. The branching Schrödinger problem

- Liero, Mielke, and Savaré (2018): Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures;
- Chizat (2017): Unbalanced optimal transport: Models, numerical methods, applications;
- Kondratyev, Monsaingeon, and Vorotnikov (2016): A new optimal transport distance on the space of finite Radon measures;
- Baradat and Lavenant (2021): Arxiv 2111.01666.

The Branching Brownian motion

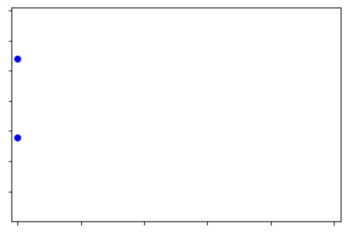
Parameters: diffusivity $\nu > 0$, branching rate $\lambda > 0$, law $(p_k)_{k=0,1,\dots} \in \mathcal{P}(\mathbb{N})$.



Particles diffuse (ν), at temporal rate λ they “branch” and have a k offsprings, drawn from $(p_k)_{k=0,1,\dots} \in \mathcal{P}(\mathbb{N})$.

The Branching Brownian motion

Parameters: diffusivity $\nu > 0$, branching rate $\lambda > 0$, law $(p_k)_{k=0,1,\ldots} \in \mathcal{P}(\mathbb{N})$.

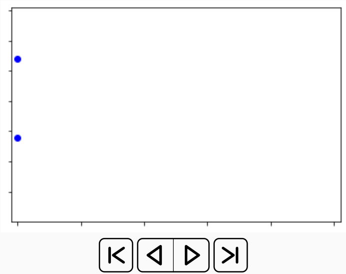


Particles diffuse (ν), at temporal rate λ they "branch" and have a k offsprings, drawn from $(p_k)_{k=0,1,\ldots} \in \mathcal{P}(\mathbb{N})$.

At time t, random **measure** $M_t = \sum_{X \in \{\text{particles alive at time } t\}} \delta_X$.

The Branching Brownian motion

Parameters: diffusivity $\nu > 0$, branching rate $\lambda > 0$, law $(p_k)_{k=0,1,\ldots} \in \mathcal{P}(\mathbb{N})$.



Particles diffuse (ν), at temporal rate λ they “branch” and have a k offsprings, drawn from $(p_k)_{k=0,1,\ldots} \in \mathcal{P}(\mathbb{N})$.

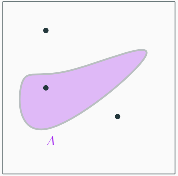
At time t, random **measure** $M_t = \sum_{X \in \{\text{particles alive at time } t\}} \delta_X$.

Description

The Branching Brownian Motion is a probability distribution on $\Omega := \text{càdlàg}([0,1], \mathcal{M}_+(\mathbb{T}^d))$.

Assumptions: $0 < \nu, \lambda < \infty$ and $\sum k p_k < +\infty$.

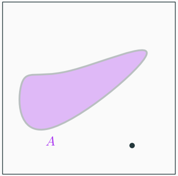
The Branching Schrödinger problem



$M_t(A) = 1$

$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

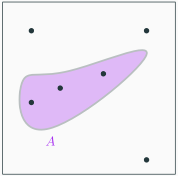
The Branching Schrödinger problem



$M_t(A) = 0$

$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

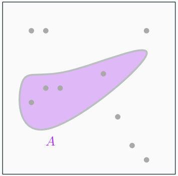
The Branching Schrödinger problem



$M_t(A) = 3$

$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

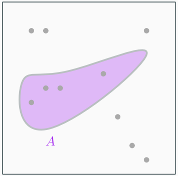
The Branching Schrödinger problem



$\mathbb{E}[M_t](A) = 4/3$

$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

The Branching Schrödinger problem



$\mathbb{E}[M_t](A) = 4/3$

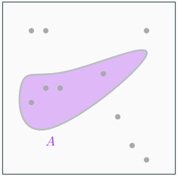
$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

R law of the Branching Brownian Motion with parameters ν, λ and (p_k).

Branching Schrödinger problem

Given $\alpha, \beta \in \mathcal{M}_+(\mathbb{T}^d)$, find $P \in \mathcal{P}(\Omega)$ which minimizes $H(P|R)$ under the constraints $\mathbb{E}_P[M_0] = \alpha$ and $\mathbb{E}_P[M_1] = \beta$.

The Branching Schrödinger problem



$\mathbb{E}[M_t](A) = 4/3$

$\mathbb{E}_P[M_t]$ is the deterministic measure $\mathbb{E}_P[M_t](A) = \mathbb{E}_P[M_t(A)]$.

R law of the Branching Brownian Motion with parameters ν, λ and (p_k).

Branching Schrödinger problem

Given $\alpha, \beta \in \mathcal{M}_+(\mathbb{T}^d)$, find $P \in \mathcal{P}(\Omega)$ which minimizes $H(P|R)$ under the constraints $\mathbb{E}_P[M_0] = \alpha$ and $\mathbb{E}_P[M_1] = \beta$.

Important remark. Ill-posed problem as the constraints are not closed:

$$\{P \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta\}$$

is not closed for a topology making $H(\cdot|R)$ continuous.

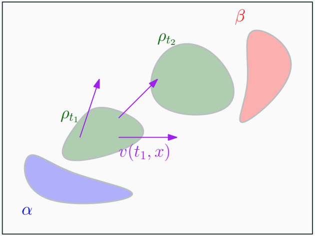
The regularized unbalanced optimal transport problem

Regularized Optimal Transport

Look for ρ, v time-dependent density, velocity field which minimize

$$\mathcal{A}(\rho, v) = \iint \frac{|v(t,x)|^2}{2} \rho(t,x)\, \mathrm{d}t \mathrm{d}x$$

under the constraint $\rho_0 = \alpha$, $\rho_1 = \beta$ and $\partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2} \Delta \rho$.



The regularized unbalanced optimal transport problem

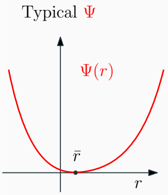
$\Psi : \mathbb{R} \to [0, +\infty]$ convex function. The field $r = r(t,x)$ is the **growth rate**.

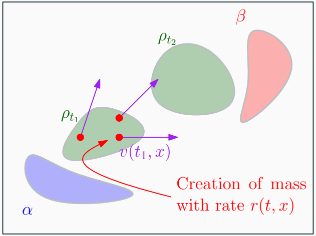
Regularized Unbalanced Optimal Transport

Look for ρ, v, r time-dependent density, velocity and scalar field which minimize

$$\mathcal{A}(\rho, v, r) = \iint \frac{|v(t,x)|^2}{2} \rho(t,x) \, \mathrm{d}t\mathrm{d}x \; + \iint \Psi(r(t,x))\rho(t,x) \, \mathrm{d}t\mathrm{d}x$$

under the constraint $\rho_0 = \alpha$, $\rho_1 = \beta$ and $\partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2} \Delta \rho + r\rho$.





The regularized unbalanced optimal transport problem

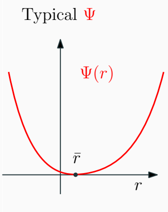
$\Psi : \mathbb{R} \to [0, +\infty]$ convex function. The field $r = r(t, x)$ is the **growth rate**.

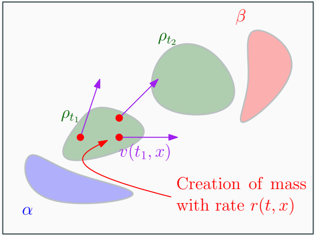
Regularized Unbalanced Optimal Transport

Look for ρ, v, r time-dependent density, velocity and scalar field which minimize

$$\mathcal{A}(\rho, v, r) = \iint \frac{|v(t,x)|^2}{2} \rho(t,x) \, \mathrm{d}t\mathrm{d}x \; + \iint \Psi(r(t,x))\rho(t,x) \, \mathrm{d}t\mathrm{d}x$$

under the constraint $\rho_0 = \alpha$, $\rho_1 = \beta$ and $\partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2} \Delta \rho + r\rho$.





If Ψ grows polynomially at $+\infty$ and $H(\beta|\mathcal{L}) < +\infty$, then well posed.

Equivalence of the values

Choose Ψ depending on λ, ν and (p_k) (see after). Write

$$\mathrm{Ruot}(\alpha, \beta) := \min_{\rho, v, r} \left\{ \mathcal{A}(\rho, v, r) \ : \ \partial_t \rho + \nabla \cdot (\rho v) = \frac{\nu}{2} \Delta \rho + r\rho, \ \rho_0 = \alpha, \rho_1 = \beta \right\}$$

$$\mathrm{BrSch}(\alpha, \beta) := \inf_{P} \{ \nu H(P|R) \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta \} .$$

Equivalence of the values

Choose Ψ depending on λ, ν and (p_k) (see after). Write

$$\mathrm{Ruot}(\alpha, \beta) := \min_{\rho, v, r} \left\{ \mathcal{A}(\rho, v, r) \ : \ \partial_t \rho + \nabla \cdot (\rho v) = \frac{\nu}{2} \Delta \rho + r\rho,\ \rho_0 = \alpha, \rho_1 = \beta \right\}$$

$$\mathrm{BrSch}(\alpha, \beta) := \inf_{P} \{ \nu H(P|R) \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta \} .$$

Define $L : \varphi \to \log \mathbb{E}_R \left[\exp \left(\langle \varphi, M_0 \rangle \right) \right]$ log-Laplace transform of R_0. We expect:

$$\nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta) = \mathrm{BrSch}(\alpha, \beta)$$

Equivalence of the values

Choose Ψ depending on λ, ν and (p_k) (see after). Write

$$\mathrm{Ruot}(\alpha, \beta) := \min_{\rho, v, r} \left\{ \mathcal{A}(\rho, v, r) \ : \ \partial_t \rho + \nabla \cdot (\rho v) = \frac{\nu}{2} \Delta \rho + r\rho, \ \rho_0 = \alpha, \rho_1 = \beta \right\}$$

$$\mathrm{BrSch}(\alpha, \beta) := \inf_{P} \{ \nu H(P|R) \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta \} .$$

Define $L : \varphi \to \log \mathbb{E}_R \left[\exp \left(\langle \varphi, M_0 \rangle \right) \right]$ log-Laplace transform of R_0. We expect:

$$\nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta) \neq \mathrm{BrSch}(\alpha, \beta)$$

Cannot hold for **all** α, β. (e.g. $\alpha = 0$)

Equivalence of the values

Choose Ψ depending on λ, ν and (p_k) (see after). Write

$$\mathrm{Ruot}(\alpha, \beta) := \min_{\rho, v, r} \left\{ \mathcal{A}(\rho, v, r) \ : \ \partial_t \rho + \nabla \cdot (\rho v) = \frac{\nu}{2} \Delta \rho + r\rho, \ \rho_0 = \alpha, \rho_1 = \beta \right\}$$

$$\mathrm{BrSch}(\alpha, \beta) := \inf_{P} \{ \nu H(P|R) \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta \} .$$

Define $L : \varphi \to \log \mathbb{E}_R \left[\exp \left(\langle \varphi, M_0 \rangle \right) \right]$ log-Laplace transform of R_0. We expect:

$$\nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta) \neq \mathrm{BrSch}(\alpha, \beta)$$

Cannot hold for **all** α, β. (e.g. $\alpha = 0$)

Theorem (equivalence of the values)

The function $(\alpha, \beta) \mapsto \nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta)$ is the lower semi continuous envelope of $(\alpha, \beta) \mapsto \mathrm{BrSch}(\alpha, \beta)$ for the topology of weak convergence.

Equivalence of the values

Choose Ψ depending on λ, ν and (p_k) (see after). Write

$$\mathrm{Ruot}(\alpha, \beta) := \min_{\rho, v, r} \left\{ \mathcal{A}(\rho, v, r) \ : \ \partial_t \rho + \nabla \cdot (\rho v) = \frac{\nu}{2} \Delta \rho + r\rho, \ \rho_0 = \alpha, \rho_1 = \beta \right\}$$

$$\mathrm{BrSch}(\alpha, \beta) := \inf_{P} \{ \nu H(P|R) \ : \ \mathbb{E}_P[M_0] = \alpha \text{ and } \mathbb{E}_P[M_1] = \beta \} .$$

Define $L : \varphi \to \log \mathbb{E}_R \left[\exp\left(\langle \varphi, M_0 \rangle \right) \right]$ log-Laplace transform of R_0. We expect:

$$\nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta) \neq \mathrm{BrSch}(\alpha, \beta)$$

Cannot hold for **all** α, β. (e.g. $\alpha = 0$)

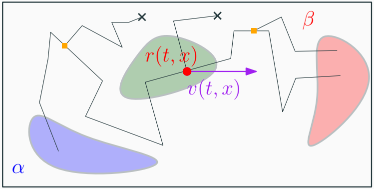
Theorem (equivalence of the values)

The function $(\alpha, \beta) \mapsto \nu L^*(\alpha) + \mathrm{Ruot}(\alpha, \beta)$ is the lower semi continuous envelope of $(\alpha, \beta) \mapsto \mathrm{BrSch}(\alpha, \beta)$ for the topology of weak convergence.

Idea of the proof: **duality**.

Equivalence of the competitors

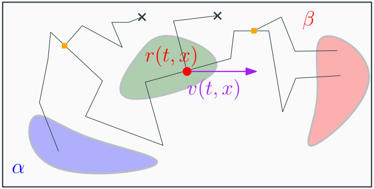
Additional assumption: one finite exponential moment for M_0 and (p_k).



Intuition: as before v drift, $r = \sum_{k=0}^{+\infty}(k-1)\tilde{\lambda}\tilde{p}_k$ for modified branching rate $\tilde{\lambda}$, modified law of offsprings $(\tilde{p}_k)_{k\in\mathbb{N}}$.

Equivalence of the competitors

Additional assumption: one finite exponential moment for M_0 and (p_k).



Intuition: as before v drift, $r = \sum_{k=0}^{+\infty}(k-1)\tilde{\lambda}\tilde{p}_k$ for modified branching rate $\tilde{\lambda}$, modified law of offsprings $(\tilde{p}_k)_{k\in\mathbb{N}}$.

From Branching Schrödinger to RUOT

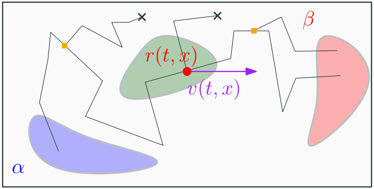
Given P with $H(P|R) < +\infty$ we build (ρ, v, r) competitor for RUOT with

$$\nu L^*(\alpha) + \mathcal{A}(\rho, v, r) \leqslant \nu H(P|R).$$

If $H(P|R) < +\infty$ then P is the law of BBM with random (predictable) space time dependent drift $\tilde{v}$, $\tilde{\lambda}$ and $(\tilde{p}_k)_{k\in\mathbb{N}}$.

Equivalence of the competitors

Additional assumption: one finite exponential moment for M_0 and (p_k).



Intuition: as before v drift, $r = \sum_{k=0}^{+\infty}(k-1)\tilde{\lambda}\tilde{p}_k$ for modified branching rate $\tilde{\lambda}$, modified law of offsprings $(\tilde{p}_k)_{k\in\mathbb{N}}$.

From Branching Schrödinger to RUOT

Given P with $H(P|R) < +\infty$ we build (ρ, v, r) competitor for RUOT with

$$\nu L^*(\alpha) + \mathcal{A}(\rho, v, r) \leqslant \nu H(P|R).$$

If $H(P|R) < +\infty$ then P is the law of BBM with random (predictable) space time dependent drift $\tilde{v}$, $\tilde{\lambda}$ and $(\tilde{p}_k)_{k\in\mathbb{N}}$.

From RUOT to Branching Schrödinger

Up to smoothing everything (including α, β) from (ρ, v, r) admissible we build a BBM with drift v and $\tilde{\lambda}$, $(\tilde{p}_k)_{k\in\mathbb{N}}$ depending on r such that

$$\nu L^*(\alpha) + \mathcal{A}(\rho, v, r) \geqslant \nu H(P|R).$$

Choosing the right growth penalization

Definition (growth penalization)

Given ν, λ and (p_k) choose

$$\Psi(r) = \nu \inf_{\tilde{\lambda}, (\tilde{p}_k)} \left\{ H(\tilde{\lambda}(\tilde{p}_k) | \lambda(p_k)) \text{ such that } \sum_{k=0}^{+\infty} (k-1) \tilde{\lambda} \tilde{p}_k = r \right\}.$$

Equivalently with $\Phi_p(X) = \sum p_k X^k$ then $\Psi^*(s) = \nu\lambda \left(e^{-s/\nu} \Phi_p(e^{s/\nu}) - 1 \right)$.

Choosing the right growth penalization

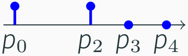
Definition (growth penalization)

Given ν, λ and (p_k) choose

$$\Psi(r) = \nu \inf_{\tilde{\lambda}, (\tilde{p}_k)} \left\{ H(\tilde{\lambda}(\tilde{p}_k) | \lambda(p_k)) \text{ such that } \sum_{k=0}^{+\infty} (k-1)\tilde{\lambda}\tilde{p}_k = r \right\}.$$

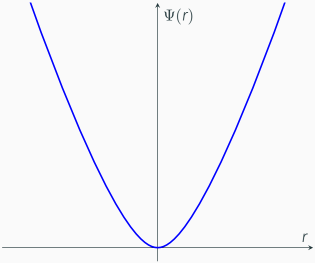
Equivalently with $\Phi_p(X) = \sum p_k X^k$ then $\Psi^*(s) = \nu\lambda \left(e^{-s/\nu} \Phi_p(e^{s/\nu}) - 1 \right)$.

If $p_0 = p_2 = 1/2$ then



$$\Psi^*(s) = \lambda\nu \left(\cosh\left[\frac{s}{\nu}\right] - 1 \right),$$

Ψ convex, minimal for $r = 0$.



Choosing the right growth penalization

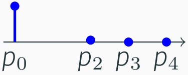
Definition (growth penalization)

Given ν, λ and (p_k) choose

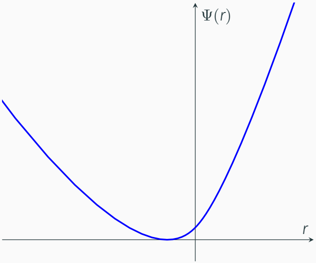
$$\Psi(r) = \nu \inf_{\tilde{\lambda}, (\tilde{p}_k)} \left\{ H(\tilde{\lambda}(\tilde{p}_k) | \lambda(p_k)) \text{ such that } \sum_{k=0}^{+\infty} (k-1)\tilde{\lambda}\tilde{p}_k = r \right\}.$$

Equivalently with $\Phi_p(X) = \sum p_k X^k$ then $\Psi^*(s) = \nu\lambda \left(e^{-s/\nu} \Phi_p(e^{s/\nu}) - 1 \right)$.

If $p_0 = 0.95$, $p_2 = 0.05$



then Ψ minimal for $\bar{r} < 0$.



Choosing the right growth penalization

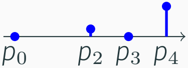
Definition (growth penalization)

Given ν, λ and (p_k) choose

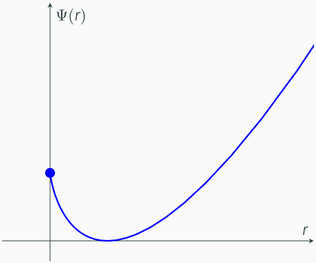
$$\Psi(r) = \nu \inf_{\tilde{\lambda}, (\tilde{p}_k)} \left\{ H(\tilde{\lambda}(\tilde{p}_k) | \lambda(p_k)) \text{ such that } \sum_{k=0}^{+\infty} (k-1)\tilde{\lambda}\tilde{p}_k = r \right\}.$$

Equivalently with $\Phi_p(X) = \sum p_k X^k$ then $\Psi^*(s) = \nu\lambda \left(e^{-s/\nu} \Phi_p(e^{s/\nu}) - 1 \right)$.

If $p_2 = 0.2$, $p_4 = 0.8$ (no killing allowed),



then $\Psi(r) = +\infty$ for $r < 0$.



Choosing the right growth penalization

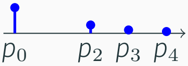
Definition (growth penalization)

Given ν, λ and (p_k) choose

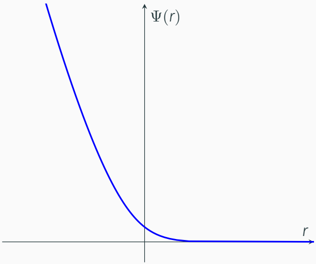
$$\Psi(r) = \nu \inf_{\tilde{\lambda}, (\tilde{p}_k)} \left\{ H(\tilde{\lambda}(\tilde{p}_k) | \lambda(p_k)) \text{ such that } \sum_{k=0}^{+\infty} (k-1) \tilde{\lambda} \tilde{p}_k = r \right\}.$$

Equivalently with $\Phi_p(X) = \sum p_k X^k$ then $\Psi^*(s) = \nu\lambda \left(e^{-s/\nu} \Phi_p(e^{s/\nu}) - 1 \right)$.

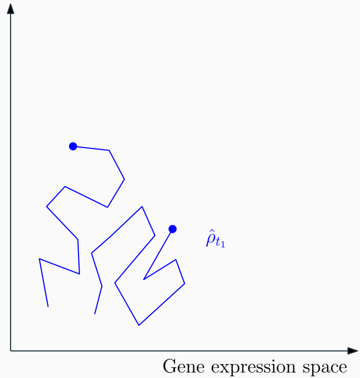
If $p_k = 1/(k-1)^{2.2}$, and $p_0 = 1 - \sum_{k \geq 2} p_k$ (no exponential moment)



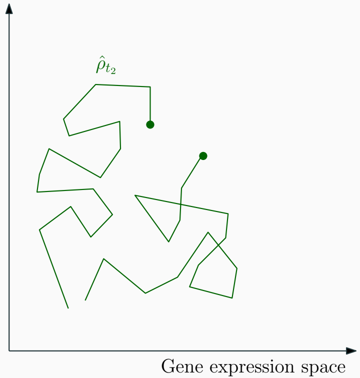
then $\Psi(r) = 0$ for $r \geq \bar{r}$.



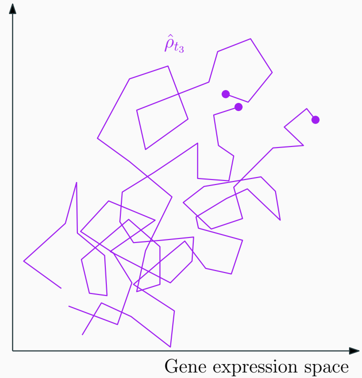
One motivation: biology



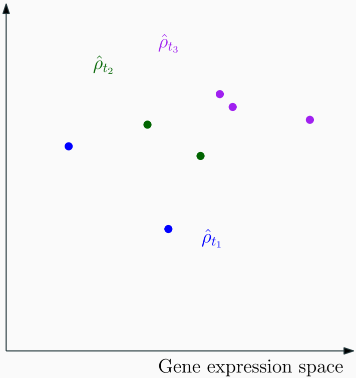
One motivation: biology



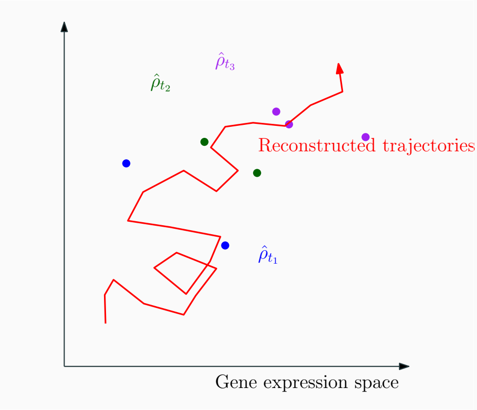
One motivation: biology



One motivation: biology



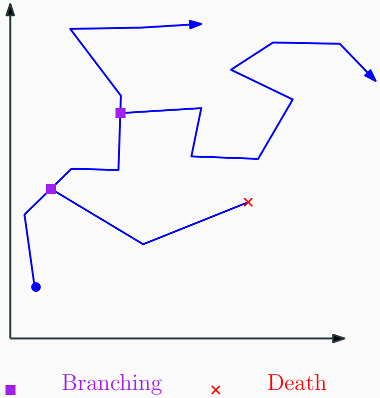
One motivation: biology



Idea: use the optimal transport to reconstruct the temporal couplings.

- Schiebinger et al, *Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming* (2019).
- Lavenant, Zhang, Kim and Schiebinger, *Towards a mathematical theory of trajectory inference* (2021).

One motivation: biology

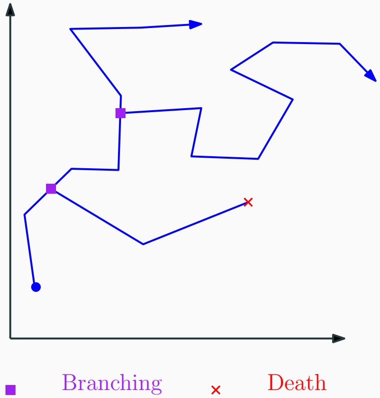


In reality cells divided and die.

Idea: use the optimal transport to reconstruct the temporal couplings.

- Schiebinger et al, *Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming* (2019).
- Lavenant, Zhang, Kim and Schiebinger, *Towards a mathematical theory of trajectory inference* (2021).

One motivation: biology



In reality cells divided and die.

Idea: use the optimal transport to reconstruct the temporal couplings.

- Schiebinger et al, *Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming* (2019).
- Lavenant, Zhang, Kim and Schiebinger, *Towards a mathematical theory of trajectory inference* (2021).

Use **unbalanced** optimal transport to account for cell division.

Conclusion

What I have not presented:

- Proofs of the equivalence (convex analysis, stochastic analysis).
- Small noise limit $\nu, \lambda \to 0$: partial optimal transport ($\Psi(r) = |r|$).
- Numerical simulations with the dynamical formulation of RUOT.
- Formal computations for other measure valued processes.

Conclusion

What I have not presented:

- Proofs of the equivalence (convex analysis, stochastic analysis).
- Small noise limit $\nu, \lambda \to 0$: partial optimal transport ($\Psi(r) = |r|$).
- Numerical simulations with the dynamical formulation of RUOT.
- Formal computations for other measure valued processes.

Thank you for your attention

Other measure valued processes?

Given a process R, need for the computation of $\mathbb{E}_R\left[\exp(\langle\theta, M_1\rangle)\middle|\, M_0\right]$.

Other measure valued processes?

Given a process R, need for the computation of $\mathbb{E}_R\left[\exp(\langle\theta, M_1\rangle)\middle| M_0\right]$.

Example (Dawson-Watanabe)

If R Dawson-Watanabe superprocess then the associated PDE is

$$\partial_t\phi + \frac{1}{2}\Delta\phi + \frac{1}{2}\phi^2 = 0$$

as

$$\mathbb{E}_R\left[\exp(\langle\phi(1,\cdot), M_1\rangle)\middle| M_0\right] = \exp(\langle\phi(0,\cdot), M_0\rangle).$$

We expect the value of the Schrödinger problem to coincide with

$$L^*(\alpha) + \min_{\rho,r}\left\{\iint r^2\rho \ : \ \partial_t\rho = \frac{\nu}{2}\Delta\rho + r\rho\right\}$$

(that is Ψ quadratic and $v = 0$).