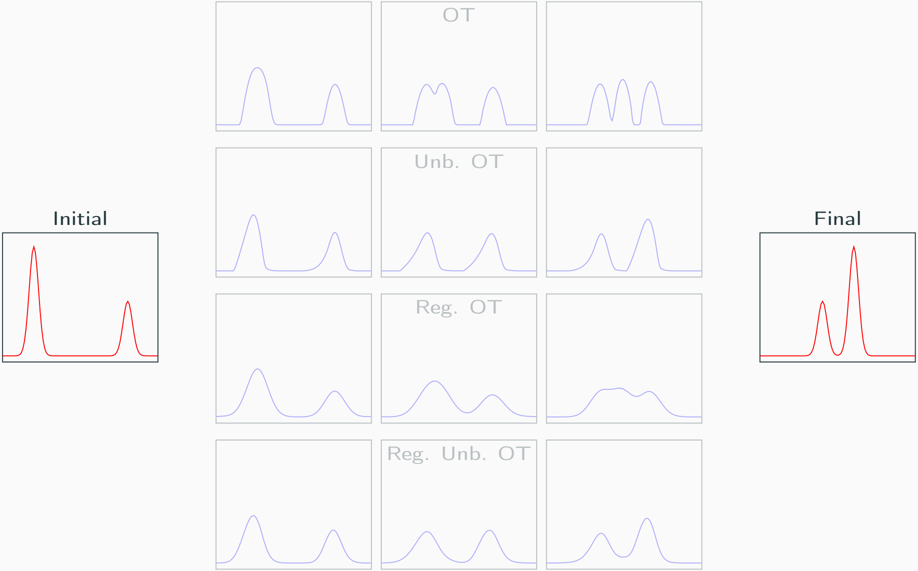
From branching Brownian motion to regularized unbalanced optimal transport

Aymeric Baradat

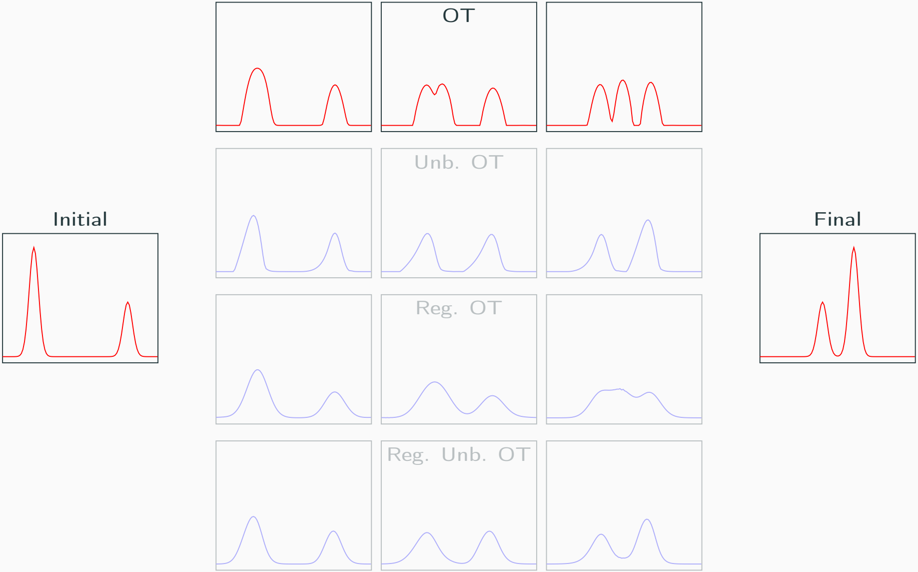
23 juin 2021

Work **in progress** in collaboration with **Hugo Lavenant**, from **Bocconi University** in Milan.

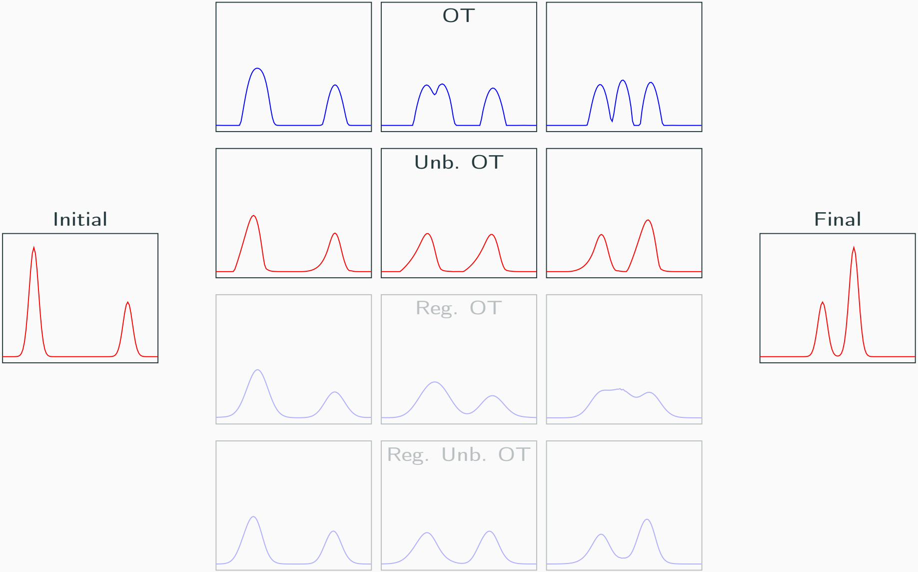
Different models of optimal transport



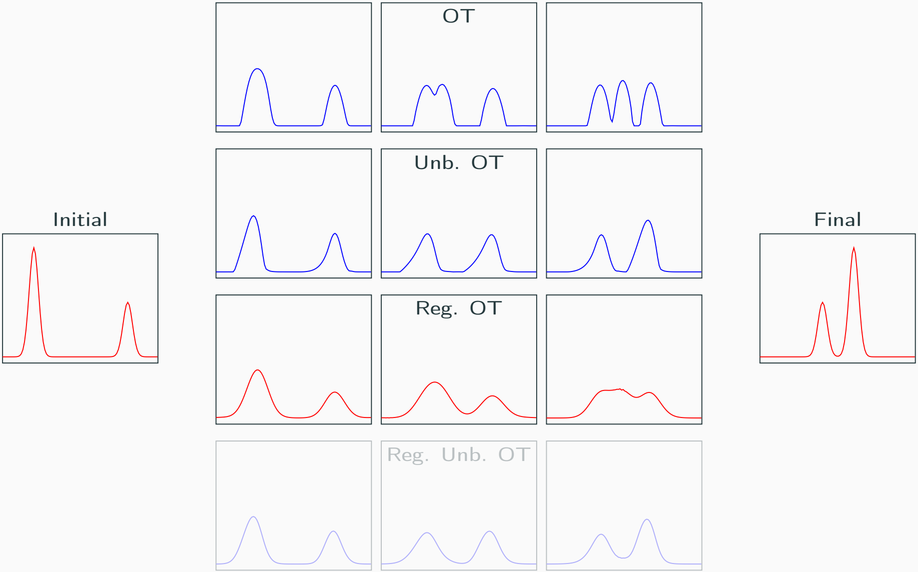
Different models of optimal transport



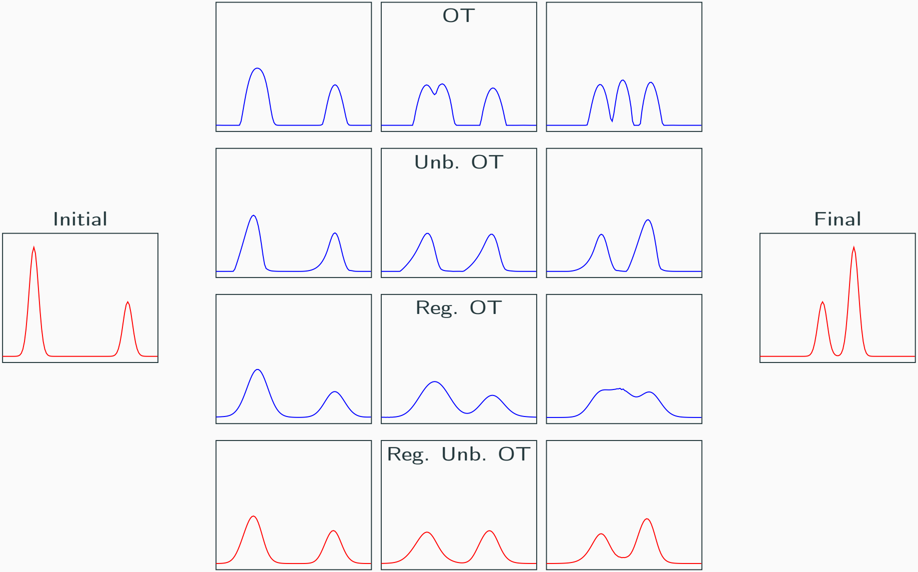
Different models of optimal transport



Different models of optimal transport



Different models of optimal transport

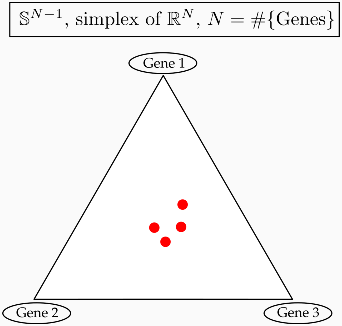


The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

The mechanism of cell specialization

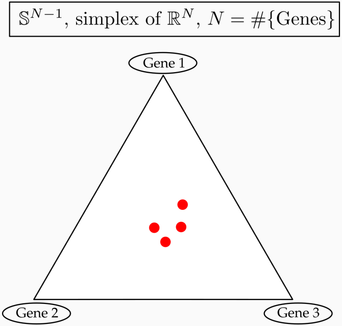
Can we describe quantitatively the process of specialization of the cells of an individual during its development?



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

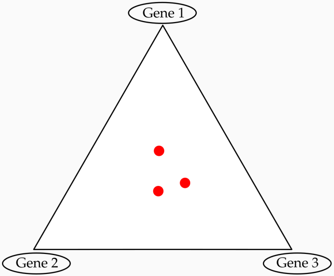


- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The i-th coordinate of a point measures the **expression rate** of gene i in the associated cell.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

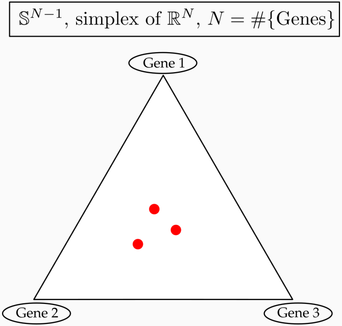
$\mathbb{S}^{N-1}$, simplex of $\mathbb{R}^N$, $N = \#\{\text{Genes}\}$



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

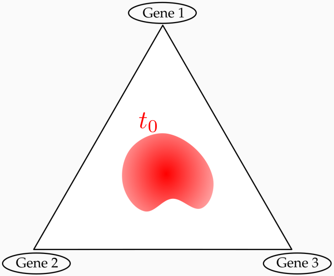


- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The i-th coordinate of a point measures the **expression rate** of gene i in the associated cell.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

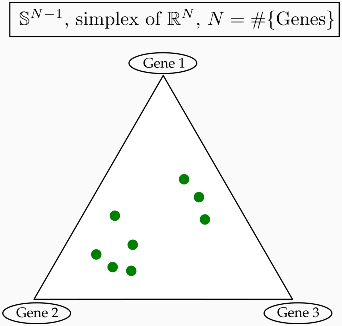
$\mathbb{S}^{N-1}$, simplex of $\mathbb{R}^N$, $N = \#\{\text{Genes}\}$



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

The mechanism of cell specialization

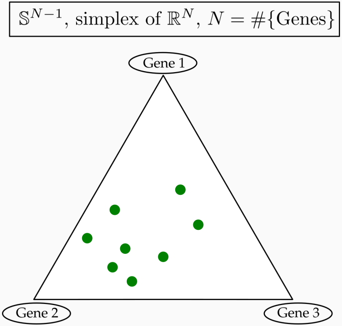
Can we describe quantitatively the process of specialization of the cells of an individual during its development?



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

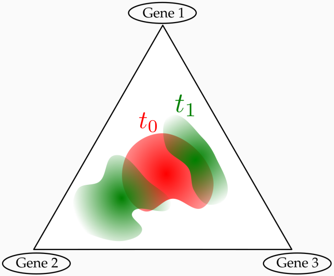


- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The i-th coordinate of a point measures the **expression rate** of gene i in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

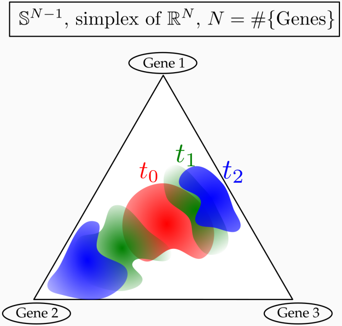
$\mathbb{S}^{N-1}$, simplex of $\mathbb{R}^N$, $N = \#\{\text{Genes}\}$



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

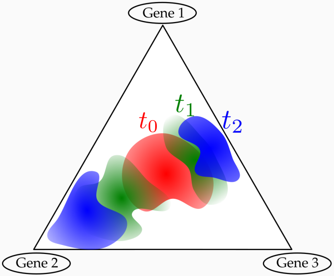


- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?

$\mathbb{S}^{N-1}$, simplex of $\mathbb{R}^N$, $N = \#\{\text{Genes}\}$

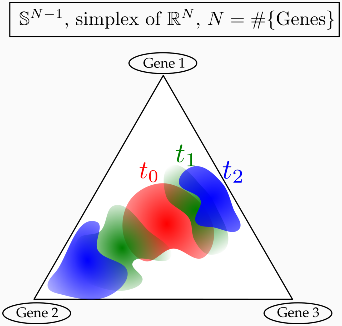


- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

Question: How to interpolate between these different intensities?

The mechanism of cell specialization

Can we describe quantitatively the process of specialization of the cells of an individual during its development?



- An **individual** is represented by a point cloud in $\mathbb{S}^{N-1}$, each point being a **cell**.
- The *i*-th coordinate of a point measures the **expression rate** of gene *i* in the associated cell.
- To a **population** corresponds a **measure** ρ on $\mathbb{S}^{N-1}$, called **intensity** of the cloud: $\rho(A) =$ average number of cells in A in our population.

Question: How to interpolate between these different intensities? Schiebinger *& al.* 2019 (Cell), Lavenant *& al.* 2021 (preprint).

Regularized unbalanced optimal transport (RUOT)

Let $\rho_0, \rho_1 \in \mathcal{M}_+(\mathbb{T}^d)$. Our optimization problem reads

Among those fields $\rho = \rho(t,x)$, $v = v(t,x)$, $r = r(t,x)$ satisfying:

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = \dfrac{\nu}{2}\Delta\rho + \rho r, \\ \rho|_{t=0} = \rho_0, \quad \rho|_{t=1} = \rho_1, \end{cases}$$

find the ones minimizing:

$$\int_0^1 \int \left\{ \frac{1}{2}|v(t,x)|^2 + \Psi(r(t,x)) \right\} \rho(t, \mathrm{d}x)\, \mathrm{d}x\, \mathrm{d}t.$$

Regularized unbalanced optimal transport (RUOT)

Let $\rho_0, \rho_1 \in \mathcal{M}_+(\mathbb{T}^d)$. Our optimization problem reads

Among those fields $\rho = \rho(t,x)$, $v = v(t,x)$, $r = r(t,x)$ satisfying:

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = \dfrac{\nu}{2} \Delta \rho + \rho r, \\ \rho|_{t=0} = \rho_0, \quad \rho|_{t=1} = \rho_1, \end{cases}$$

find the ones minimizing:

$$\int_0^1 \int \left\{ \frac{1}{2} |v(t,x)|^2 + \Psi(r(t,x)) \right\} \rho(t, \mathrm{d}x) \, \mathrm{d}x \, \mathrm{d}t.$$

Question: How to chose Ψ?

Branching Brownian motion (BBM) as a reference model

We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

Branching Brownian motion (BBM) as a reference model

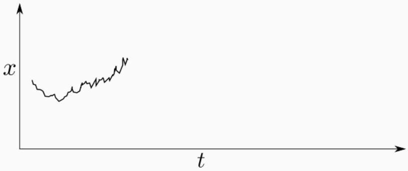
We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.

Branching Brownian motion (BBM) as a reference model

We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

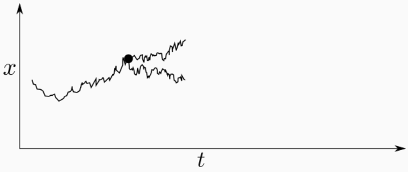
We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.



Branching Brownian motion (BBM) as a reference model

We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

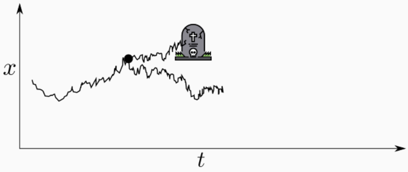
We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.



Branching Brownian motion (BBM) as a reference model

We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

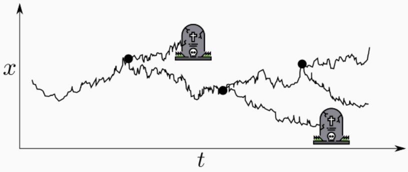
We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.



Branching Brownian motion (BBM) as a reference model

We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

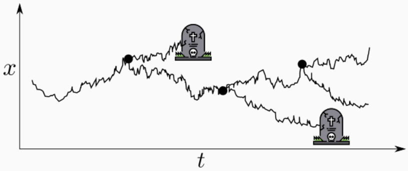
We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.



Branching Brownian motion (BBM) as a reference model

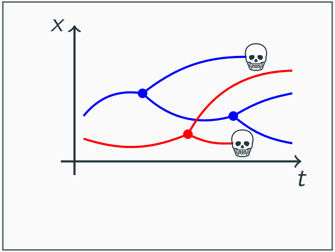
We want a reference model including both random trajectories for the particles, and the possibility for them to die or divide.

We choose the model of **branching Brownian motion**: each particle of diffusivity $\boldsymbol{\nu} > 0$ has a random lifetime of law $\mathcal{E}xp(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} > 0$. When it dies, it gives birth to $k \in \mathbb{N}$ particles with probability $\boldsymbol{p_k} \in [0, 1]$.

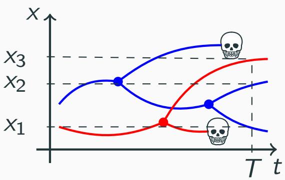


Remark: λ and $\boldsymbol{p}$ could depend on t and x: $\lambda(t, x)\,\mathrm{d}t$ represents the probability for a particle in (t, x) to die between t and $t + \mathrm{d}t$. Then, $p_k(t, x)$ is the probability that it is replaced by k particles.

Entropic minimization w.r.t. BBM

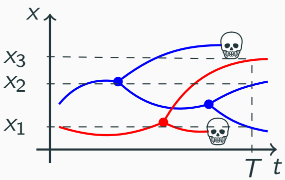


Entropic minimization w.r.t. BBM



Here, $\omega_T = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$.

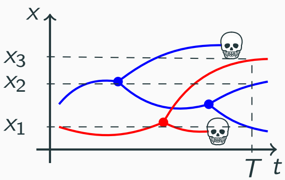
Entropic minimization w.r.t. BBM



Here, $\omega_T = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$.

- Ω := *càdlàg*$([0,1];\mathcal{M}_\delta(\mathbb{T}^d))$, $\mathcal{M}_\delta(\mathbb{T}^d)$ being the set of sums of unitary Diracs on $\mathbb{T}^d$.

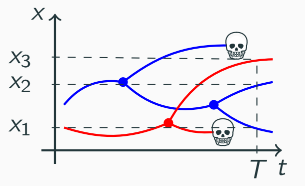
Entropic minimization w.r.t. BBM



Here, $\omega_T = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$.

- $\Omega := c\grave{a}dl\grave{a}g([0,1]; \mathcal{M}_\delta(\mathbb{T}^d))$, $\mathcal{M}_\delta(\mathbb{T}^d)$ being the set of sums of unitary Diracs on $\mathbb{T}^d$.
- If $\omega \in \Omega$ and $t \in [0,1]$, the value ω_t of the curve ω at time t is therefore a sum of Diracs.

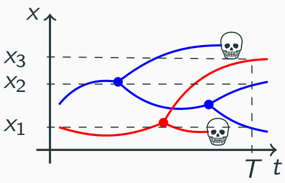
Entropic minimization w.r.t. BBM



Here, $\omega_T = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$.

- $\Omega := \textit{càdlàg}([0,1]; \mathcal{M}_\delta(\mathbb{T}^d))$, $\mathcal{M}_\delta(\mathbb{T}^d)$ being the set of sums of unitary Diracs on $\mathbb{T}^d$.
- If $\omega \in \Omega$ and $t \in [0,1]$, the value ω_t of the curve ω at time t is therefore a sum of Diracs.
- $R \in \mathcal{P}(\Omega)$: law of BBM of parameters ν, λ and $\boldsymbol{p} = (p_k)_k$, of initial law $R_0 \in \mathcal{P}(\mathcal{M}_\delta(\mathbb{T}^d))$.

Entropic minimization w.r.t. BBM



Here, $\omega_T = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$.

- $\Omega := \textit{càdlàg}([0,1]; \mathcal{M}_\delta(\mathbb{T}^d))$, $\mathcal{M}_\delta(\mathbb{T}^d)$ being the set of sums of unitary Dirac on $\mathbb{T}^d$.
- If $\omega \in \Omega$ and $t \in [0,1]$, the value ω_t of the curve ω at time t is therefore a sum of Diracs.
- $R \in \mathcal{P}(\Omega)$: law of BBM of parameters ν, λ and $\boldsymbol{p} = (p_k)_k$, of initial law $R_0 \in \mathcal{P}(\mathcal{M}_\delta(\mathbb{T}^d))$.

Fixing the marginals $\rho_0, \rho_1 \in \mathcal{M}_+(\mathbb{T}^d)$, our minimization problem reads:

Among those $P \in \mathcal{P}(\Omega)$ satisfying for all $A \in \mathcal{B}(\mathbb{T}^d)$:

$$\mathbb{E}_P[\omega_0(A)] = \rho_0(A) \qquad \text{and} \qquad \mathbb{E}_P[\omega_1(A)] = \rho_1(A).$$

find the ones minimizing

$$H(P|R) := \mathbb{E}_R\left[\frac{\mathrm{d}P}{\mathrm{d}R}\log\frac{\mathrm{d}P}{\mathrm{d}R}\right] = \mathbb{E}_P\left[\log\frac{\mathrm{d}P}{\mathrm{d}R}\right].$$

Equivalence of the models

RUOT

Let $\nu > 0$ and Ψ a function.

Find (ρ, v, r) s.t.

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = \dfrac{\nu}{2}\Delta\rho + \rho r, \\ \rho|_{t=0} = \rho_0, \quad \rho|_{t=1} = \rho_1, \end{cases}$$

minimizing

$$\int_0^1 \int \left\{ \frac{|v|^2}{2} + \Psi(r) \right\} \rho \, \mathrm{d}x \, \mathrm{d}t.$$

Branching Schrödinger

R BBM of initial law R_0, of parameter ν, λ and $\boldsymbol{p}$.

Find $P \in \mathcal{P}(\Omega)$ s.t.

$$E_P[\omega_i(A)] = \rho_i(A), \quad i = 0, 1,$$

minimizing $H(P|R)$.

Equivalence of the models

RUOT

Let $\nu > 0$ and Ψ a function.
Find (ρ, v, r) s.t.

$$\begin{cases} \partial_t \rho + \mathrm{div}(\rho v) = \frac{\nu}{2}\Delta\rho + \rho r, \\ \rho|_{t=0} = \rho_0, \quad \rho|_{t=1} = \rho_1, \end{cases}$$

minimizing

$$\int_0^1 \int \left\{ \frac{|v|^2}{2} + \Psi(r) \right\} \rho \, \mathrm{d}x \, \mathrm{d}t.$$

Branching Schrödinger

R BBM of initial law R_0, of parameter ν, λ and $\boldsymbol{p}$.
Find $P \in \mathcal{P}(\Omega)$ s.t.

$$E_P[\omega_i(A)] = \rho_i(A), \quad i = 0, 1,$$

minimizing $H(P|R)$.

Main result (BL): If Ψ is defined *via* its Legendre transform by:

$$\Psi^*(s) = \nu\lambda \sum p_k \left\{ e^{(k-1)\frac{s}{\nu}} - 1 \right\},$$

The minimum in RUOT is the l.s.c. relaxation of the infimum in Branching Schrödinger, up to an initial term.

A few ideas

Changing the law: Let us call $\boldsymbol{q} = \lambda \boldsymbol{p}$.

A few ideas

Changing the law: Let us call $\boldsymbol{q} = \lambda \boldsymbol{p}$. Let P be the law of the BBM of drift $v = v(t,x)$, whose branching events are describe by $\tilde{\boldsymbol{q}} = \tilde{\boldsymbol{q}}(t,x)$, and of initial law P_0, then calling $\rho(t) := \mathbb{E}_P[\omega_t]$ and $r(t,x) := \sum_k (k-1) q_k$, we have:

A few ideas

Changing the law: Let us call $\boldsymbol{q} = \lambda \boldsymbol{p}$. Let P be the law of the BBM of drift $v = v(t,x)$, whose branching events are describe by $\tilde{\boldsymbol{q}} = \tilde{\boldsymbol{q}}(t,x)$, and of initial law P_0, then calling $\rho(t) := \mathbb{E}_P[\omega_t]$ and $r(t,x) := \sum_k (k-1) q_k$, we have:

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

$$\nu H(P|R) = \nu H(P_0|R_0) + \int_0^1 \int \left\{ \frac{1}{2} |v(t,x)|^2 + h(\tilde{\boldsymbol{q}}(t,x)|\boldsymbol{q}) \right\} \rho(t, \mathrm{d}x) \, \mathrm{d}x \, \mathrm{d}t,$$

$$\text{where} \qquad h(\tilde{\boldsymbol{q}}|\boldsymbol{q}) := \sum_k \tilde{q}_k \log \frac{\tilde{q}_k}{q_k} + q_k - \tilde{q}_k.$$

A few ideas

Changing the law: Let us call $\boldsymbol{q} = \lambda \boldsymbol{p}$. Let P be the law of the BBM of drift $v = v(t,x)$, whose branching events are describe by $\tilde{\boldsymbol{q}} = \tilde{\boldsymbol{q}}(t,x)$, and of initial law P_0, then calling $\rho(t) := \mathbb{E}_P[\omega_t]$ and $r(t,x) := \sum_k (k-1) q_k$, we have:

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

$$\nu H(P|R) = \nu H(P_0|R_0) + \int_0^1 \int \left\{ \frac{1}{2} |v(t,x)|^2 + h(\tilde{\boldsymbol{q}}(t,x)|\boldsymbol{q}) \right\} \rho(t, \mathrm{d}x)\, \mathrm{d}x\, \mathrm{d}t,$$

$$\text{where} \qquad h(\tilde{\boldsymbol{q}}|\boldsymbol{q}) := \sum_k \tilde{q}_k \log \frac{\tilde{q}_k}{q_k} + q_k - \tilde{q}_k.$$

Observation: Provided Ψ is defined as before,

$$\Psi(r) = \inf \left\{ h(\tilde{\boldsymbol{q}}|\boldsymbol{q}) \ : \ \sum (k-1) \tilde{q}_k = r \right\}.$$

A few ideas

Changing the law: Let us call $\boldsymbol{q} = \lambda \boldsymbol{p}$. Let P be the law of the BBM of drift $v = v(t,x)$, whose branching events are describe by $\tilde{\boldsymbol{q}} = \tilde{\boldsymbol{q}}(t,x)$, and of initial law P_0, then calling $\rho(t) := \mathbb{E}_P[\omega_t]$ and $r(t,x) := \sum_k (k-1) q_k$, we have:

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

$$\nu H(P|R) = \nu H(P_0|R_0) + \int_0^1 \int \left\{ \frac{1}{2} |v(t,x)|^2 + h(\tilde{\boldsymbol{q}}(t,x)|\boldsymbol{q}) \right\} \rho(t, \mathrm{d}x) \, \mathrm{d}x \, \mathrm{d}t,$$

$$\text{where} \qquad h(\tilde{\boldsymbol{q}}|\boldsymbol{q}) := \sum_k \tilde{q}_k \log \frac{\tilde{q}_k}{q_k} + q_k - \tilde{q}_k.$$

Observation: Provided Ψ is defined as before,

$$\Psi(r) = \inf \left\{ h(\tilde{\boldsymbol{q}}|\boldsymbol{q}) \ : \ \sum (k-1) \tilde{q}_k = r \right\}.$$

A counter-example to equality: In RUOT, we can connect 0 to $\rho_1 \neq 0$. In branching Schrödinger it is not possible.

Conclusion

- By studying the rare events of the branching Brownian motion, we derived a specific cost in a model of regularized and unbalanced optimal transport model, that seems natural in some applications.

Conclusion

- By studying the rare events of the branching Brownian motion, we derived a specific cost in a model of regularized and unbalanced optimal transport model, that seems natural in some applications.
- In the regime $\nu \to 0$, $\lambda \sim \exp(-1/\nu)$, we recover **optimal partial transport** at the **PDE level**.

Conclusion

- By studying the rare events of the branching Brownian motion, we derived a specific cost in a model of regularized and unbalanced optimal transport model, that seems natural in some applications.
- In the regime $\nu \to 0$, $\lambda \sim \exp(-1/\nu)$, we recover **optimal partial transport** at the **PDE level**.
- Now, we want to include more biology in the model. Ex: replace the Brownian trajectories in the space of gene expressions by more relevant ones.

Conclusion

- By studying the rare events of the branching Brownian motion, we derived a specific cost in a model of regularized and unbalanced optimal transport model, that seems natural in some applications.
- In the regime $\nu \to 0$, $\lambda \sim \exp(-1/\nu)$, we recover **optimal partial transport** at the **PDE level**.
- Now, we want to include more biology in the model. Ex: replace the Brownian trajectories in the space of gene expressions by more relevant ones.