

Quelques modèles en Transport Optimal de mesures à valeurs vectorielles

Xavier Bacon

Soutenance de thèse, sous la direction de Bruno Nazaret

2 décembre 2022



1 Introduction au Transport Optimal

2 Transport Optimal multi-espèces

B., Journal of Optimization Theory and Applications, 2020.

3 Un problème spatial d'échanges économiques à la Pareto

B., G. Carlier, B. Nazaret, accepted for publication in Applied Mathematics and Optimization Journal, 2022.

4 Transport Optimal de matrices SDP et distance de Bures-Wasserstein

B., F-X. Vialard

1 Introduction au Transport Optimal

2 Transport Optimal multi-espèces

B., Journal of Optimization Theory and Applications, 2020.

3 Un problème spatial d'échanges économiques à la Pareto

B., G. Carlier, B. Nazaret, accepted for publication in Applied Mathematics and Optimization Journal, 2022.

4 Transport Optimal de matrices SDP et distance de Bures-Wasserstein

B., F-X. Vialard

Introduction au Transport Optimal (1/2)

- $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ désignent respectivement les distributions **initiale** et **finale**.
- $c = c(x, y)$ est le coût de transport.

Le problème de Transport Optimal (TO) est défini par

$$\mathcal{T}_c(\mu, \nu) = \inf \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y) \text{ t.q. } \gamma \in \Pi(\mu, \nu) \right\}$$

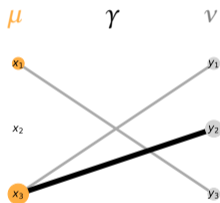
où $\Pi(\mu, \nu)$ désigne l'ensemble des mesures $\gamma \in \mathcal{P}(X \times Y)$ admettant μ et ν comme marginales, appelées **plans de transport**.

Dualité

Étant un problème de minimisation convexe, $\mathcal{T}_c(\mu, \nu)$ admet comme formulation le problème de maximisation concave suivant

$$\mathcal{T}_c(\mu, \nu) = \sup \left\{ \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y) \text{ t.q. } \varphi(x) + \psi(y) \leq c(x, y) \right\}$$

où φ et ψ sont deux fonctions continues et bornées.



Introduction au Transport Optimal (2/2)

X, Y supposés finis et $\alpha = (\alpha_i) \in \mathbb{R}^X, \beta = (\beta_j) \in \mathbb{R}^Y$ désignent les distributions **initiale** et **finale**.

- **TO Entropique** : il s'agit d'un outil efficace pour calculer une approximation de $\mathcal{T}(\alpha, \beta)$: $\varepsilon > 0$ étant fixé, le problème de TO régularisé est donné par

$$\mathcal{T}^\varepsilon(\alpha, \beta) = \inf \left\{ \langle c | \gamma \rangle - \varepsilon \mathbf{Entropie}(\gamma) \text{ t.q. } \gamma \in \Pi(\alpha, \beta) \right\}$$

dont l'unique solution converge vers la solution de $\mathcal{T}(\alpha, \beta)$ d'entropie maximale.

- **Calcul de la solution de $\mathcal{T}^\varepsilon(\alpha, \beta)$** : les conditions d'optimalité s'écrivent ici

$$\gamma_{ij} = u_i K_{ij} v_j$$

où $u_i = \exp(\varphi_i/\varepsilon)$, $v_j = \exp(\psi_j/\varepsilon)$ et $K = \exp(-c_{ij}/\varepsilon)$. On utilise l'algorithme de Sinkhorn pour approximer la solution de $\mathcal{T}^\varepsilon(\alpha, \beta)$: partant de $u^0 = \mathbb{1}_X$ et $v^0 = \mathbb{1}_Y$, on calcule de manière itérative

$$u^{k+1} = \frac{\alpha}{K \cdot v^k} \text{ et } v^{k+1} = \frac{\beta}{K^T \cdot u^{k+1}}, \text{ for } k \in \mathbb{N}.$$

1 Introduction au Transport Optimal

2 Transport Optimal multi-espèces

B., Journal of Optimization Theory and Applications, 2020.

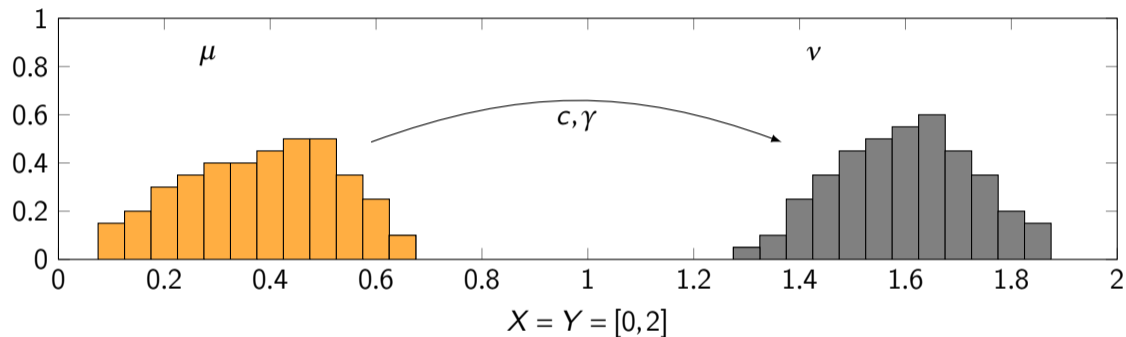
3 Un problème spatial d'échanges économiques à la Pareto

B., G. Carlier, B. Nazaret, accepted for publication in Applied Mathematics and Optimization Journal, 2022.

4 Transport Optimal de matrices SDP et distance de Bures-Wasserstein

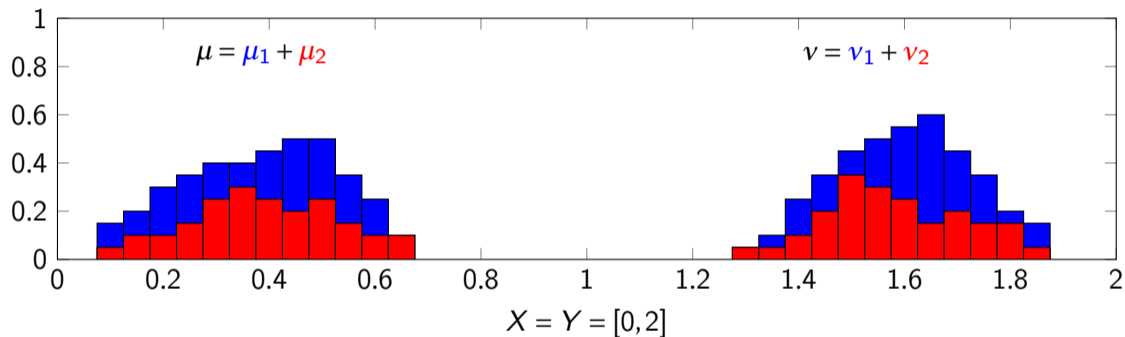
B., F-X. Vialard

Transport Optimal : 1 espèce



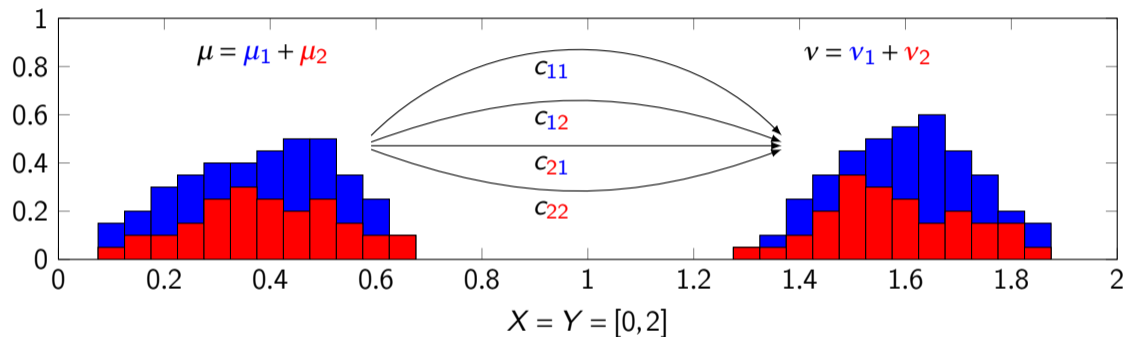
From scalar to Multi-species Optimal Transport

Transport Optimal : 2 espèces



■ $\mu = \mu_1 + \mu_2 \in \mathcal{P}(X)$, $\nu = \nu_1 + \nu_2 \in \mathcal{P}(Y)$.

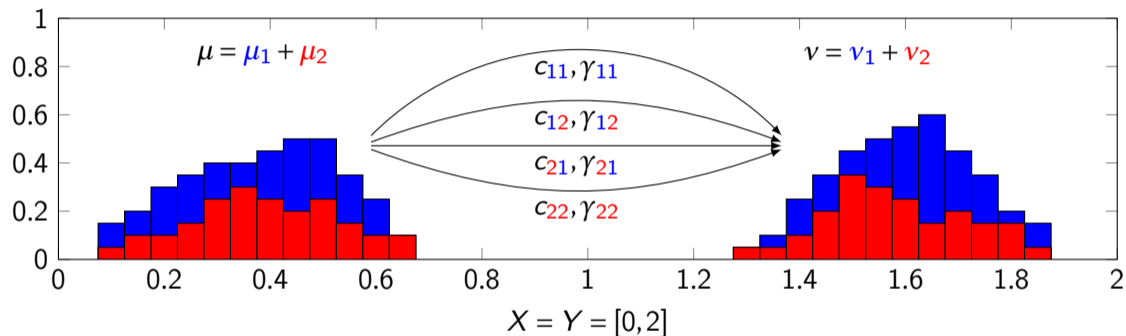
Transport Optimal : 2 espèces



- $\mu = \mu_1 + \mu_2 \in \mathcal{P}(X)$, $\nu = \nu_1 + \nu_2 \in \mathcal{P}(Y)$.
- Quatre coûts de transport différents c_{11} , c_{12} , c_{21} , c_{22} .

From scalar to Multi-species Optimal Transport

Transport Optimal : 2 espèces



- $\mu = \mu_1 + \mu_2 \in \mathcal{P}(X)$, $\nu = \nu_1 + \nu_2 \in \mathcal{P}(Y)$.
- Quatre coûts de transport différents c_{11} , c_{12} , c_{21} , c_{22} .
- ...Et autant de plans de transport γ_{11} , γ_{12} , γ_{21} , γ_{22} .

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	
Distributions finales	
Plans de transport	
Primal : (\mathcal{P}_1)	
Dual : (\mathcal{D}_1)	
Conditions d'optimalité	
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	
Primal : (\mathcal{P}_1)	
Dual : (\mathcal{D}_1)	
Conditions d'optimalité	
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	$\boldsymbol{\gamma} = (\gamma_{ij}) \in \mathcal{M}_+^{n \times m}(X \times Y)$ t.q. $\begin{cases} \pi_1 \# \sum_{k=1}^m \gamma_{ik} = \mu_i \\ \pi_2 \# \sum_{k=1}^m \gamma_{kj} = \nu_j \end{cases}$
Primal : (\mathcal{P}_1)	
Dual : (\mathcal{D}_1)	
Conditions d'optimalité	
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	$\boldsymbol{\gamma} = (\gamma_{ij}) \in \mathcal{M}_+^{n \times m}(X \times Y)$ t.q. $\begin{cases} \pi_1 \# \sum_{k=1}^m \gamma_{ik} = \mu_i \\ \pi_2 \# \sum_{k=1}^m \gamma_{kj} = \nu_j \end{cases}$
Primal : (\mathcal{P}_1)	$\mathcal{H}(\boldsymbol{\mu}, \boldsymbol{\nu}) = \inf \sum_{i,j} \langle c_{ij} \gamma_{ij} \rangle$
Dual : (\mathcal{D}_1)	
Conditions d'optimalité	
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	$\boldsymbol{\gamma} = (\gamma_{ij}) \in \mathcal{M}_+^{n \times m}(X \times Y)$ t.q. $\begin{cases} \pi_1 \# \sum_{k=1}^m \gamma_{ik} = \mu_i \\ \pi_2 \# \sum_{k=1}^m \gamma_{kj} = \nu_j \end{cases}$
Primal : (\mathcal{P}_1)	$\mathcal{K}(\boldsymbol{\mu}, \boldsymbol{\nu}) = \inf \sum_{i,j} \langle c_{ij} \mid \gamma_{ij} \rangle$
Dual : (\mathcal{D}_1)	$\sup \sum_{i=1}^n \langle \varphi_i \mid \mu_i \rangle + \sum_{j=1}^m \langle \psi_j \mid \nu_j \rangle$ t.q. $\varphi_i \oplus \psi_j \leq c_{ij}$
Conditions d'optimalité	
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	$\boldsymbol{\gamma} = (\gamma_{ij}) \in \mathcal{M}_+^{n \times m}(X \times Y)$ t.q. $\begin{cases} \pi_1 \# \sum_{k=1}^m \gamma_{ik} = \mu_i \\ \pi_2 \# \sum_{k=1}^m \gamma_{kj} = \nu_j \end{cases}$
Primal : (\mathcal{P}_1)	$\mathcal{K}(\boldsymbol{\mu}, \boldsymbol{\nu}) = \inf \sum_{i,j} \langle c_{ij} \mid \gamma_{ij} \rangle$
Dual : (\mathcal{D}_1)	$\sup \sum_{i=1}^n \langle \varphi_i \mid \mu_i \rangle + \sum_{j=1}^m \langle \psi_j \mid \nu_j \rangle$ t.q. $\varphi_i \oplus \psi_j \leq c_{ij}$
Conditions d'optimalité	$\varphi_i \oplus \psi_j = c_{ij} \quad \gamma_{ij}$ -p.p.
TO Entropique	

Tableau récapitulatif

	TO Multi-espèces
Distributions initiales	$\boldsymbol{\mu} = (\mu_i) \in \mathcal{M}_+^n(X)$ t.q. $\sum_{i=1}^n \mu_i \in \mathcal{P}(X)$
Distributions finales	$\boldsymbol{\nu} = (\nu_j) \in \mathcal{M}_+^m(Y)$ t.q. $\sum_{j=1}^m \nu_j \in \mathcal{P}(Y)$
Plans de transport	$\boldsymbol{\gamma} = (\gamma_{ij}) \in \mathcal{M}_+^{n \times m}(X \times Y)$ t.q. $\begin{cases} \pi_1 \# \sum_{k=1}^m \gamma_{ik} = \mu_i \\ \pi_2 \# \sum_{k=1}^m \gamma_{kj} = \nu_j \end{cases}$
Primal : (\mathcal{P}_1)	$\mathcal{H}(\boldsymbol{\mu}, \boldsymbol{\nu}) = \inf \sum_{i,j} \langle c_{ij} \gamma_{ij} \rangle$
Dual : (\mathcal{D}_1)	$\sup \sum_{i=1}^n \langle \varphi_i \mu_i \rangle + \sum_{j=1}^m \langle \psi_j \nu_j \rangle$ t.q. $\varphi_i \oplus \psi_j \leq c_{ij}$
Conditions d'optimalité	$\varphi_i \oplus \psi_j = c_{ij} \quad \gamma_{ij}$ -p.p.
TO Entropique	$\inf \sum_{i,j} \langle c_{ij} \gamma_{ij} \rangle - \varepsilon \text{Entropie}(\gamma_{ij})$

Quelques résultats

Existence d'une solution dans le problème primal (B.)

Si les coûts c_{ij} sont minorés et semi continus inférieurement. Il existe $\gamma = (\gamma_{ij})$ solution de (\mathcal{P}_1) .

Existence d'une solution dans le problème dual (B.)

Si X et Y sont deux espaces métriques compacts et que $c = (c_{ij})$ est continue, il existe (φ, ψ) solution de (\mathcal{D}_1) .

Dualité forte (B.)

Sous les conditions du théorème précédent, il n'y a aucun saut de dualité entre (\mathcal{P}_1) et (\mathcal{D}_1) , autrement dit :

$$(\mathcal{P}_1) = (\mathcal{D}_1).$$

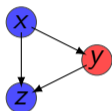
Comment obtenir une distance ?

Comment obtenir une distance ?

- **Transport Optimal scalaire** : si (X, d) est un espace polonais, pour $p \in [1, \infty)$ et $c = d^p$, $\mathcal{T}^{1/p}$ est une distance sur $\mathcal{P}_p(X)$.

Comment obtenir une distance ?

- **Transport Optimal scalaire** : si (X, d) est un espace polonais, pour $p \in [1, \infty)$ et $c = d^p$, $\mathcal{T}^{1/p}$ est une distance sur $\mathcal{P}_p(X)$.

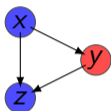


- **Cas multi-espèces** : setting $c_{ij} = d_{ij}^p$, we make the following assumption

$$d_{ik}(x, z) \leq d_{ij}(x, y) + d_{jk}(y, z) \quad (MTI)$$

Comment obtenir une distance ?

- **Transport Optimal scalaire** : si (X, d) est un espace polonais, pour $p \in [1, \infty)$ et $c = d^p$, $\mathcal{T}^{1/p}$ est une distance sur $\mathcal{P}_p(X)$.



- **Cas multi-espèces** : setting $c_{ij} = d_{ij}^p$, we make the following assumption

$$d_{ik}(x, z) \leq d_{ij}(x, y) + d_{jk}(y, z) \quad (MTI)$$

Théorème (B.)

Étant donné n^2 fonctions (d_{ij}) définies sur $X \times X$ et à valeurs dans \mathbb{R}_+ telles que

[1] $\forall (i, j) \in \llbracket 1, n \rrbracket^2$, d_{ij} est symétrique.

[2] (MTI) est vérifiée pour tout $(i, j, k) \in \llbracket 1, n \rrbracket^3$ et $(x, y, z) \in X^3$.

[3] $\forall i \in \llbracket 1, n \rrbracket, \forall x \in X, d_{ii}(x, x) = 0$.

[4] $\forall (i, j) \in \llbracket 1, n \rrbracket^2, i \neq j, \forall (x, y) \in X \times X, d_{ij}(x, y) \neq 0$.

Alors $\mathcal{K}(\boldsymbol{\mu}, \boldsymbol{\nu})^{1/p}$ est une distance sur $\mathcal{P}_p^n(X)$.

Simulations numériques

$$\mu = \frac{1}{\pi} \mathcal{U}_{B(0,1]}$$

$$v_1 = \frac{1}{2} \delta_{y_1}$$

$$v_2 = \frac{1}{2} \delta_{y_2}$$

$$y_1 = (-t, 0)$$

$$y_2 = (t, 0)$$

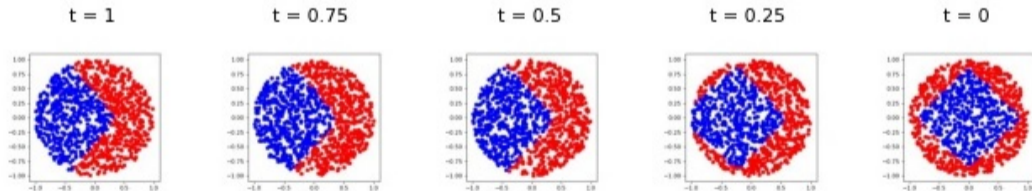
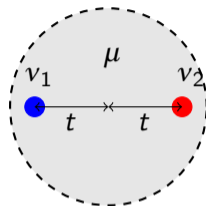


Figure: Partitions de μ pour cinq $t \in [0,1]$ ($p = 1.5, p' = 3$)

1 Introduction au Transport Optimal

2 Transport Optimal multi-espèces

B., Journal of Optimization Theory and Applications, 2020.

3 Un problème spatial d'échanges économiques à la Pareto

B., G. Carlier, B. Nazaret, accepted for publication in Applied Mathematics and Optimization Journal, 2022.

4 Transport Optimal de matrices SDP et distance de Bures-Wasserstein

B., F-X. Vialard

Problème primal

- X est un espace métrique compact et m une mesure de référence *i.e.* $\mu_i \ll m$ pour $i = 1, \dots, N$.
- $\mu = (\mu_1, \dots, \mu_N) \in L^1(X, m)^N$ sont les distributions initiales de biens dans la région X .

Problème primal

$$(\mathcal{P}_2) = \sup \{ \quad \}.$$

Problème primal

- X est un espace métrique compact et m une mesure de référence *i.e.* $\mu_i \ll m$ pour $i = 1, \dots, N$.
- $\mu = (\mu_1, \dots, \mu_N) \in L^1(X, m)^N$ sont les distributions initiales de biens dans la région X .
- $\nu = (\nu_1, \dots, \nu_N) \in \mathcal{M}_+(X)^N$ sont les distributions finales de biens dans la région X .

Problème primal

$$(\mathcal{P}_2) = \sup_{\nu} \left\{ \quad \quad \quad \text{t.q. } \mu_i(X) = \nu_i(X) \right\}.$$

Problème primal

- X est un espace métrique compact et m une mesure de référence *i.e.* $\mu_i \ll m$ pour $i = 1, \dots, N$.
- $\mu = (\mu_1, \dots, \mu_N) \in L^1(X, m)^N$ sont les distributions initiales de biens dans la région X .
- $\mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_N) \in \mathcal{M}_+(X)^N$ sont les distributions finales de biens dans la région X .
- $\mathcal{U} = \mathcal{U}(\mathbf{v})$ est l'utilité moyenne (la somme des préférences de chaque agent se trouvant en x).

Problème primal

$$(\mathcal{P}_2) = \sup_{\mathbf{v}} \left\{ \mathcal{U}(\mathbf{v}) \quad \text{t.q.} \quad \mu_i(X) = \mathbf{v}_i(X) \right\}.$$

Problème primal

- X est un espace métrique compact et m une mesure de référence *i.e.* $\mu_i \ll m$ pour $i = 1, \dots, N$.
- $\mu = (\mu_1, \dots, \mu_N) \in L^1(X, m)^N$ sont les distributions initiales de biens dans la région X .
- $\nu = (\nu_1, \dots, \nu_N) \in \mathcal{M}_+(X)^N$ sont les distributions finales de biens dans la région X .
- $\mathcal{U} = \mathcal{U}(\nu)$ est l'utilité moyenne (la somme des préférences de chaque agent se trouvant en x).
- $\mathcal{T}(\mu, \nu) = \sum_{i=1}^N \mathcal{T}_{c_i}(\mu_i, \nu_i)$ est le coût de transport entre μ et ν .

Problème primal

$$(\mathcal{P}_2) = \sup_{\nu} \left\{ \mathcal{U}(\nu) - \mathcal{T}(\mu, \nu) \text{ t.q. } \mu_i(X) = \nu_i(X) \right\}.$$

Quelques mots sur l'utilité

$U : (x, \beta) \in X \times \mathbb{R}_+^N \mapsto U(x, \beta) \in \mathbb{R} \cup \{-\infty\}$ est la préférence de l'agent se trouvant en x .

- $\forall i = 1, \dots, N, c_i$ est continu, positif et $c_i(x, x) = 0$ pour tout $x \in X$.
- m-p.p. $x \in X, U$ est semi continue supérieurement, concave et monotone.
- $(x, \beta) \mapsto U(x, \beta)$ est sous-linéaire en β uniformément par rapport $x \in X$.

Exemple : pour $a_1, \dots, a_N > 0$ tels que $\sum_{i=1}^N a_i < 1$ et $p = p(x) > 0$ un poids,

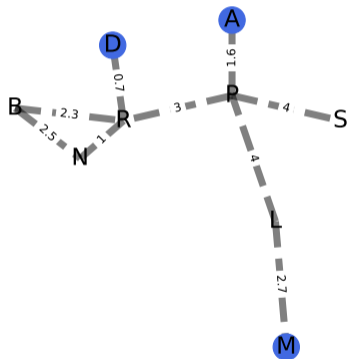
$$U(x, \beta(x)) = p(x) \prod_{i=1}^N \beta_i(x)^{a_i}.$$

(Cobb-Douglas)

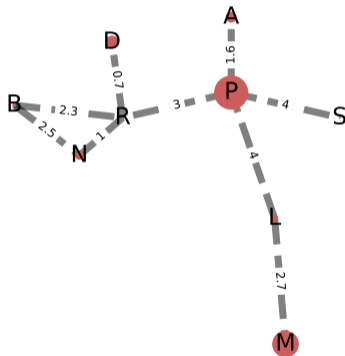
Un exemple avec un seul bien

Échanges économiques entre 9 villes : 1 bien

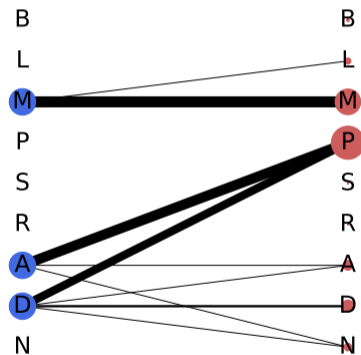
Distribution initiale



Distribution finale



Plan de transport optimal



$$(\mathcal{P}_2) = \sup_{\mathbf{v}} \left\{ \mathcal{U}(\mathbf{v}) - \mathcal{I}(\mu, \mathbf{v}) \text{ t.q. } \mu_i(X) = \mathbf{v}_i(X) \right\}.$$

Proposition (B., Carlier, Nazaret, 2021)

Le problème de maximisation (\mathcal{P}_2) admet au moins une solution dans $L^1(X, m)$.

$$(\mathcal{P}_2) = \sup_{\nu} \left\{ \mathcal{U}(\nu) - \mathcal{I}(\mu, \nu) \text{ t.q. } \mu_i(X) = \nu_i(X) \right\}.$$

Proposition (B., Carlier, Nazaret, 2021)

Le problème de maximisation (\mathcal{P}_2) admet au moins une solution dans $L^1(X, m)$.

Preuve : $N = 1$, en utilisant la méthode directe du calcul des variations, partant d'une suite maximisante $\nu^n = \beta^n \cdot m$, on en extrait une sous-suite (toujours notée ν^n) qui admet la limite suivante

$$\nu^n \xrightarrow{*} \nu$$

$$(\mathcal{P}_2) = \sup_{\mathbf{v}} \left\{ \mathcal{U}(\mathbf{v}) - \mathcal{I}(\mu, \mathbf{v}) \text{ t.q. } \mu_i(X) = \mathbf{v}_i(X) \right\}.$$

Proposition (B., Carlier, Nazaret, 2021)

Le problème de maximisation (\mathcal{P}_2) admet au moins une solution dans $L^1(X, m)$.

Preuve : $N = 1$, en utilisant la méthode directe du calcul des variations, partant d'une suite maximisante $\mathbf{v}^n = \beta^n \cdot m$, on en extrait une sous-suite (toujours notée \mathbf{v}^n) qui admet la limite suivante

$$\mathbf{v}^n \xrightarrow{*} \mathbf{v}$$

⚠ \mathbf{v} peut ne pas appartenir à $L^1(m)$ ⚠

Existence of a minimizer

Pour venir à bout de cette difficulté, partant de ν , on construit un autre candidat admissible $\tilde{\nu}$ qui fait augmenter \mathcal{U} et diminuer $\mathcal{T}(\mu, \cdot)$. Soit $\beta^a \in L^1$ et $\nu^s \in \mathcal{M}_+(X)$ tels que $\nu^s \perp m$ et

$$\nu = \beta^a \cdot m + \nu^s \quad (\text{Radon-Nikodym})$$

Soit $\gamma \in \Pi(\mu, \nu)$ optimal entre μ et ν . On le décompose en

$$\gamma = \underbrace{\gamma|_{X \times A}}_{\gamma^a} + \underbrace{\gamma|_{X \times A^c}}_{\gamma^s}$$

où A vérifie $\nu^s(A) = m(A^c) = 0$. En posant

$$\tilde{\gamma} = \gamma^a + (\text{Id}, \text{Id}) \# [\text{proj}_1 \# \gamma^s] \cdot m,$$

la seconde marginale de $\tilde{\gamma}$ est solution de (\mathcal{P}) si $c \geq 0$ et $c(x, x) = 0$ pour tout $x \in X$.

Le problème (\mathcal{P}_2) apparait naturellement comme étant le problème dual d'un problème de minimisation convexe :

$$(\mathcal{D}_2) = \inf \left\{ \mathcal{K}(\varphi) + \mathcal{V}(\varphi) \text{ t.q. } \varphi \in C(X, \mathbb{R})^N \right\}$$

où

$$\mathcal{K}(\varphi) = - \sum_{i=1}^N \int_X \varphi_i^{c_i} d\mu_i, \quad \text{and} \quad \mathcal{V}(\varphi) = \int_X V(x, \varphi_1(x), \dots, \varphi_N(x)) dm(x)$$

et $V(x, \varphi) = \sup \{ U(x, \beta) - \sum_{i=1}^N \beta_i \varphi_i : \beta \in \mathbb{R}_+^N \}$.

Théorème (Dualité (B., Carlier, Nazaret))

Il n'y a aucun saut de dualité entre (\mathcal{P}_2) et (\mathcal{D}_2) , i.e.

$$(\mathcal{D}_2) = (\mathcal{P}_2).$$

Problème entropique et algorithme

- X est fini et m est la mesure de comptage sur X .

$$(\mathcal{D}_2) = \inf_{\varphi \in \mathbb{R}^{X \times N}} \sum_{y \in X} V(y, \varphi(y)) + \sum_{i=1}^N \sum_{x \in X} \alpha_i(x) \overbrace{\max_{y \in X} \{\varphi_i(y) - c_i(x, y)\}}^{= -\varphi^c(x)}.$$

$$(\mathcal{D}_\varepsilon) = \inf_{\varphi \in \mathbb{R}^{X \times N}} \sum_{y \in X} V(y, \varphi(y)) + \varepsilon \sum_{i=1}^N \sum_{x \in X} \alpha_i(x) \log \left(\sum_{y \in X} e^{\frac{\varphi_i(y) - c_i(x, y)}{\varepsilon}} \right).$$

$(\mathcal{D}_\varepsilon)$ peut être réécrite en considérant la formulation convexe suivante

$$(\tilde{\mathcal{D}}_\varepsilon) = \inf_{\varphi, \psi} \Phi_\varepsilon(\varphi, \psi)$$

$$\text{où } \Phi_\varepsilon(\varphi, \psi) = \sum_{y \in X} V(y, \varphi(y)) - \sum_{i=1}^N \sum_{x \in X} \alpha_i(x) \psi_i(x) + \varepsilon \sum_{i=1}^N \sum_{(x, y) \in X^2} e^{\frac{\psi_i(x) + \varphi_i(y) - c_i(x, y)}{\varepsilon}}$$

dont la solution est calculée par une descente de gradient coordonnée par coordonnée.

Soient (β, φ) optimaux dans (\mathcal{P}_2) et (\mathcal{D}_2) . Alors on a un équilibre économique pour la dotation initiale $w = \langle \varphi^c \mid \alpha \rangle + \langle \varphi \mid \beta \rangle (= \mathcal{T}_c(\alpha, \beta))$ dans le sens où:

- Les **vendeurs** en x maximisent leurs profits en exportant leurs biens $\alpha(x)$:

$$\begin{aligned} \text{profit unitaire}_i(x) &= \max_y \varphi_i(y) - c_i(x, y) \quad (= -\varphi_i^{c_i}(x)) \\ \text{profits totaux}(x) &= \langle \text{profit unitaire}(x) \mid \alpha(x) \rangle \end{aligned}$$

- Les **acheteurs** en y ont une dotation initiale $w(y)$ et achètent $\beta_i(y)$ de la manière suivante :

$$\beta_i(y) = \operatorname{argmax}_{\beta} \left\{ \mathbf{U}(y, \beta) \text{ t.q. } \langle \varphi \mid \beta \rangle \leq \underbrace{\langle -\varphi^c(y) \mid \alpha(y) \rangle + w(y)}_{\text{revenu total}} \right\}$$

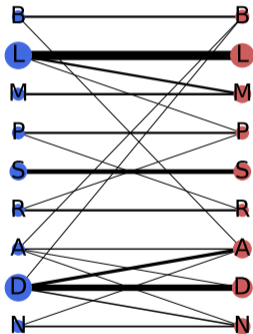
profits d'exportation

- **Libre marché** : il existe un plan de transport optimal pour toutes les distributions finales.

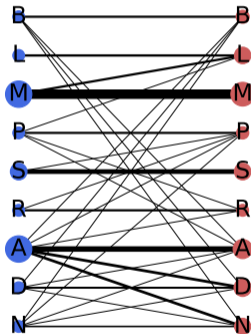
Échanges économiques entre 9 villes : 2 biens



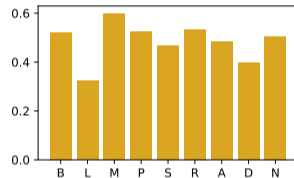
Plan de transport optimal #1



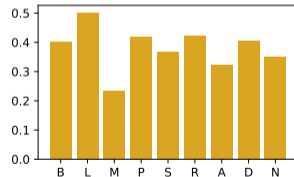
Plan de transport optimal #2



Prix #1



Prix #2



- Prendre comme coût de transport des biens un coût multi-espèces :

$$\sum_{i=1}^N \mathcal{I}_{c_i}(\mu_i, \nu_i) \leftarrow \mathcal{K}(\boldsymbol{\mu}, \boldsymbol{\nu}).$$

- Retirer la contrainte de préservation de la masse et intégrer un Transport Optimal non équilibré (*unbalanced*).
- Dans le cas où l'espace X est un compact de \mathbb{R}^d , proposer un modèle dynamique.

1 Introduction au Transport Optimal

2 Transport Optimal multi-espèces

B., Journal of Optimization Theory and Applications, 2020.

3 Un problème spatial d'échanges économiques à la Pareto

B., G. Carlier, B. Nazaret, accepted for publication in Applied Mathematics and Optimization Journal, 2022.

4 Transport Optimal de matrices SDP et distance de Bures-Wasserstein

B., F-X. Vialard

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
non convexe de d_{BW}

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
convexe de d_{BW}

Géodésiques

$(S, U) \rightsquigarrow (S, R = SU)$

Formulation variationnelle
non convexe de d_{BW}

Formulation variationnelle (non) convexe

Monsaingeon & Vorotnikov (2021) ont proposé le modèle suivant : étant donné $S_0, S_1 \in \mathcal{S}_+$, ils considèrent la formulation *non convexe*

$$\inf \left\{ \frac{1}{2} \int_0^1 \text{Tr}(U_t^* S_t U_t) dt \text{ t.q. } \partial_t S_t = (S_t U_t)^{\text{sym}}, S_{t=0} = S_0 \text{ et } S_{t=1} = S_1 \right\}.$$

- **Changement de variables** : $(S, U) \longleftarrow (S, R = SU)$ et ainsi le problème ci-dessus se transforme en le problème *convexe* suivant

$$\inf \left\{ \int_0^1 A(S_t, R_t) dt \text{ t.q. } \partial_t S_t = R_t^{\text{sym}}, S_{t=0} = S_0 \text{ and } S_{t=1} = S_1 \right\} \quad (P_3)$$

où

$$A(S, R) = \begin{cases} \frac{1}{2} \text{Tr}(R^* S^\dagger R) & \text{if } S \in \mathcal{S}_+ \text{ and } \text{im}(R) \subset \text{im}(S) \\ +\infty & \text{sinon.} \end{cases}$$

$$\sup_{Q \in C^1([0,1], \mathcal{S})} \left\{ \text{Tr}(S_1 Q_1) - \text{Tr}(S_0 Q_0) \text{ tel que } \partial_t Q + \frac{1}{2} Q^2 \leq 0 \right\} \quad (D_3)$$

$$(P_3) = \sup_{Q \in C^1([0,1], \mathcal{S})} \left\{ \text{Tr}(S_1 Q_1) - \text{Tr}(S_0 Q_0) \text{ tel que } \partial_t Q + \frac{1}{2} Q^2 \leq 0 \right\} \quad (D_3)$$

$$(P_3) = \sup_{Q \in C^1([0,1], \mathcal{S})} \left\{ \text{Tr}(S_1 Q_1) - \text{Tr}(S_0 Q_0) \text{ tel que } \partial_t Q + \frac{1}{2} Q^2 \leq 0 \right\} \quad (D_3)$$

Conditions d'optimalité : $(S, R, Q) \in C^1([0, 1], \mathcal{S} \times \mathbb{R}^{n \times n} \times \mathcal{S})$ sont optimaux dans les problèmes (P_3) et (D_3) si et seulement si

$$\begin{cases} S_{t=0} = S_0, S_{t=1} = S_1 \\ \partial_t S = (SQ)^{\text{sym}} \\ \partial_t Q + \frac{1}{2} Q^2 \leq 0 \\ \text{Tr} S (\partial_t Q + \frac{1}{2} Q^2) = 0 \end{cases}$$

Rang un : pour deux matrices SDP de rang 1, notées $S_0 = s_0 s_0^*$ et $S_1 = s_1 s_1^*$, une géodésique pour la distance de Bures-Wasserstein est donnée par

$$S_t \stackrel{\text{def.}}{=} s_t s_t^* \quad \text{où} \quad s_t \stackrel{\text{def.}}{=} (1-t)s_0 + ts_1.$$

Matrices de rangs égaux

- À toute $S \in \mathcal{S}_{++}$ est associée une mesure gaussienne centrée de variance S :

$$\mathcal{S}_{++} \ni S \longleftrightarrow \mu_S \in \mathcal{P}(\mathbb{R}^n)$$

et il y a une isométrie entre $(\mathcal{S}_{++}, d_{BW})$ et (Gaussiennes, \mathcal{T}_2) [McCann 1996, Takatsu 2008].

- La géodésique entre μ_{S_0} et μ_{S_1} est donnée par $\mu_t = [(1-t)\text{Id} + t\Lambda] \# \mu_{S_0}$ où

$$\Lambda \in \mathcal{S}_{++} \text{ et } \Lambda S_0 \Lambda = S_1.$$

(Éq. de Riccati)

Matrices de rangs égaux

- À toute $S \in \mathcal{S}_{++}$ est associée une mesure gaussienne centrée de variance S :

$$\mathcal{S}_{++} \ni S \longleftrightarrow \mu_S \in \mathcal{P}(\mathbb{R}^n)$$

et il y a une isométrie entre $(\mathcal{S}_{++}, d_{BW})$ et (Gaussiennes, \mathcal{T}_2) [McCann 1996, Takatsu 2008].

- La géodésique entre μ_{S_0} et μ_{S_1} est donnée par $\mu_t = [(1-t)\text{Id} + t\Lambda] \# \mu_{S_0}$ où

$$\Lambda \in \mathcal{S}_{++} \text{ et } \Lambda S_0 \Lambda = S_1. \quad (\text{Éq. de Riccati})$$

- Si on suppose que $\ker(S_0) \subset \ker(S_1)$, l'équation de Riccati admet une solution SDP [Fujii, 2018].

Matrices de rangs égaux

- À toute $S \in \mathcal{S}_{++}$ est associée une mesure gaussienne centrée de variance S :

$$\mathcal{S}_{++} \ni S \longleftrightarrow \mu_S \in \mathcal{P}(\mathbb{R}^n)$$

et il y a une isométrie entre $(\mathcal{S}_{++}, d_{BW})$ et (Gaussiennes, \mathcal{T}_2) [McCann 1996, Takatsu 2008].

- La géodésique entre μ_{S_0} et μ_{S_1} est donnée par $\mu_t = [(1-t)\text{Id} + t\Lambda] \# \mu_{S_0}$ où

$$\Lambda \in \mathcal{S}_{++} \text{ et } \Lambda S_0 \Lambda = S_1. \quad (\text{Éq. de Riccati})$$

- Si on suppose que $\ker(S_0) \subset \ker(S_1)$, l'équation de Riccati admet une solution SDP [Fujii, 2018].
- (B., Vialard) Si de plus, $\text{rank}(S_0) = \text{rank}(S_1) = d$, à toute solution SDP de l'équation de Riccati est associée une géodésique entre S_0 et S_1 donnée par

$$S_t = [(1-t)\text{Id} + t\Lambda] S_0 [(1-t)\text{Id} + t\Lambda]^*$$

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
convexe de d_{BW}

Géodésiques

$(S, U) \rightsquigarrow (S, R = SU)$

Formulation variationnelle
non convexe de d_{BW}

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
convexe de d_{BW}

Géodésiques

$(S, U) \rightsquigarrow (S, R = SU)$

Formulation variationnelle
non convexe de d_{BW}

TO non équilibré scalaire

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
convexe de d_{BW}

Géodésiques

$(S, U) \rightsquigarrow (S, R = SU)$

Formulation variationnelle
non convexe de d_{BW}

TO non équilibré SDP

$\rho \leftarrow \mathcal{F}$

TO non équilibré scalaire

TO non équilibré et distance de Bures-Wasserstein

- **Transport Optimal non équilibré** : étant donné ρ_0 et ρ_1 deux mesures finies sur Ω un domaine de \mathbb{R}^p , on s'intéresse à la minimisation de la fonctionnelle

$$\text{Action}(\rho, \omega, \zeta) = \frac{1}{2} \int_0^1 \left(\int_{\Omega} \frac{|\omega_t(x)|^2}{\rho_t(x)} dx + \int_{\Omega} \frac{|\zeta_t(x)|^2}{\rho_t(x)} dx \right) dt.$$

sous la contrainte de l'équation de continuité avec terme source suivante

$$\partial_t \rho + \text{div}_x \omega = \zeta \quad \text{et} \quad \rho_{t=0} = \rho_0, \rho_{t=1} = \rho_1.$$

- De \mathbb{R}_+ à \mathcal{S}_+ : formellement on remplace les trois variables

$$\rho \longleftarrow \mathcal{S}, \quad \omega \longleftarrow \mathcal{M}, \quad \zeta \longleftarrow \mathcal{R},$$

$$\mathcal{A}(\mathcal{S}, \mathcal{M}, \mathcal{R}) = \frac{1}{2} \int_0^1 \left(\int_{\Omega} \text{Tr} \left(\mathcal{M}_t^*(x) \mathcal{S}_t^\dagger(x) \mathcal{M}_t(x) \right) + \text{Tr} \left(\mathcal{R}_t^*(x) \mathcal{S}_t^\dagger(x) \mathcal{R}_t(x) \right) dx \right) dt$$

$$\partial_t \mathcal{S} + \text{div}_x \mathcal{M}^{\text{sym}} = \mathcal{R}^{\text{sym}} \quad \text{et} \quad \mathcal{S}_{t=0} = \mathcal{S}_0, \mathcal{S}_{t=1} = \mathcal{S}_1.$$

TO non équilibré et distance de Bures-Wasserstein

- **Transport Optimal non équilibré** : étant donné ρ_0 et ρ_1 deux mesures finies sur Ω un domaine de \mathbb{R}^p , on s'intéresse à la minimisation de la fonctionnelle

$$\text{Action}(\rho, \omega, \zeta) = \frac{1}{2} \int_0^1 \left(\int_{\Omega} \frac{|\omega_t(x)|^2}{\rho_t(x)} dx + \int_{\Omega} \frac{|\zeta_t(x)|^2}{\rho_t(x)} dx \right) dt.$$

sous la contrainte de l'équation de continuité avec terme source suivante

$$\partial_t \rho + \text{div}_x \omega = \zeta \quad \text{et} \quad \rho_{t=0} = \rho_0, \rho_{t=1} = \rho_1.$$

- De \mathbb{R}_+ à \mathcal{S}_+ : et si Ω est réduit à un singleton (ou s'il n'y a aucun mouvement) :

$$\mathcal{A}(\mathcal{I}, \mathcal{M}, \mathcal{R}) = \frac{1}{2} \int_0^1 \left(\int_{\Omega} \text{Tr} \left(\mathcal{M}_t^*(x) \mathcal{I}_t^\dagger(x) \mathcal{M}_t(x) \right) + \text{Tr} \left(\mathcal{R}_t^*(x) \mathcal{I}_t^\dagger(x) \mathcal{R}_t(x) \right) dx \right) dt$$

$$\partial_t \mathcal{I} + \text{div}_x \mathcal{M}^{\text{sym}} = \mathcal{R}^{\text{sym}} \quad \text{et} \quad \mathcal{I}_{t=0} = \mathcal{I}_0, \mathcal{I}_{t=1} = \mathcal{I}_1.$$

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2 \text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+

Géodésiques

Formulation variationnelle
convexe de d_{BW}

Géodésiques

$(S, U) \rightsquigarrow (S, R = SU)$

Formulation variationnelle
non convexe de d_{BW}

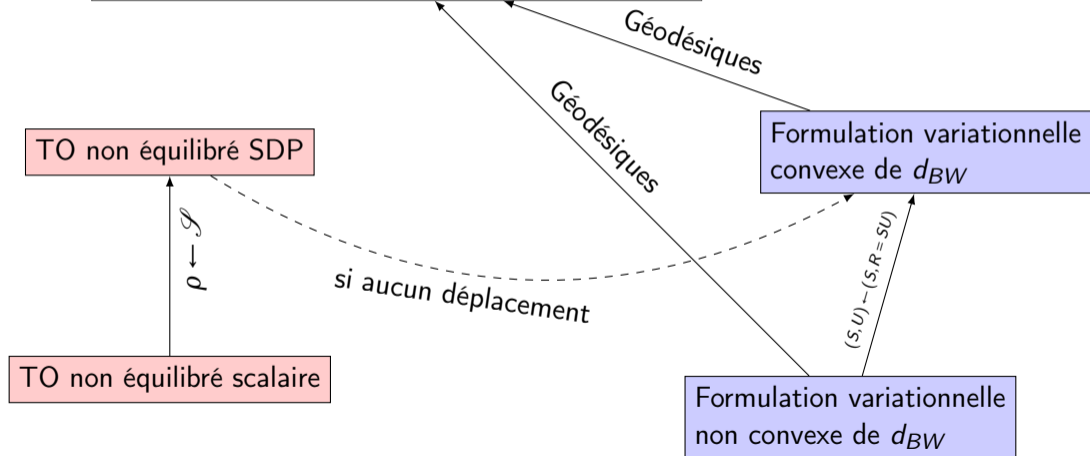
TO non équilibré SDP

$\rho \leftarrow \mathcal{F}$

TO non équilibré scalaire

$$d_{BW}^2(S_0, S_1) = \text{Tr } S_0 + \text{Tr } S_1 - 2\text{Tr} \left(S_0^{1/2} S_1 S_0^{1/2} \right)^{1/2}$$

Bures-Wasserstein sur \mathcal{S}_+



- Étendre à l'ensemble des matrices SDP les résultats exploitant l'équation de Riccati.
- Étudier numériquement le problème dynamique de transport de matrices SDP.
- Existe-t-il une formulation *statique* (à la Kantorovich) du problème dynamique ?

Merci pour votre attention.