

Soutenance de thèse

Transport Optimal de Mesures Positives : Modèles, Méthodes Numériques, Applications

Lénaïc Chizat

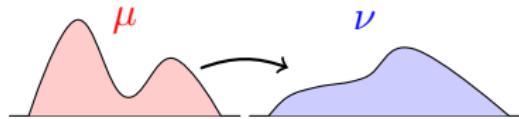
Thèse préparée à l'Université Paris-Dauphine
Encadrée par G. Peyré et F.-X. Vialard

10 Novembre 2017

Objet de la thèse

Problème de Monge (1781)

Comment transporter un tas de terre d'une configuration à une autre à moindre coût ?



⇒ possible seulement si les deux tas ont la même masse

Objet de la thèse

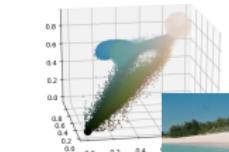
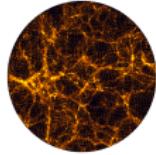
Extension aux tas de masses différentes

Pourquoi s'y intéresser ?

Fort potentiel de modélisation : mesure positive

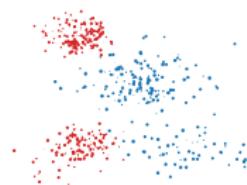
- distribution de probabilité, distribution empirique
- particules pondérées indistinguables
- densité d'un gaz, d'une foule

Origines de l'univers
(Brenier *et al.* '08)



Transfert de couleurs
(Delon '10)

Mouvements de foules
(Roudneff *et al.*, 12')



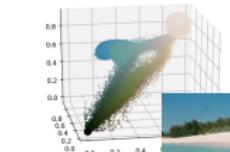
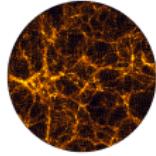
Nuages de points

Pourquoi s'y intéresser ?

Fort potentiel de modélisation : mesure positive

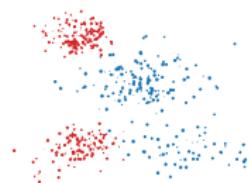
- distribution de probabilité, distribution empirique
- particules pondérées indistinguables
- densité d'un gaz, d'une foule

Origines de l'univers
(Brenier *et al.* '08)



Transfert de couleurs
(Delon '10)

Mouvements de foules
(Roudneff *et al.*, 12')



Nuages de points

Fournit des outils précieux

- distance géodésique utilisant la géométrie de l'espace
- plans de transport, barycentres, flots de gradients...

Stratégie

- préserver les propriétés clés du transport optimal
- combiner les géométries *horizontales* (transport) et *verticales* (linéaire)

Verticale

Horizontale

Mixte

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

4 Une application

Flot de gradient Hele-Shaw

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

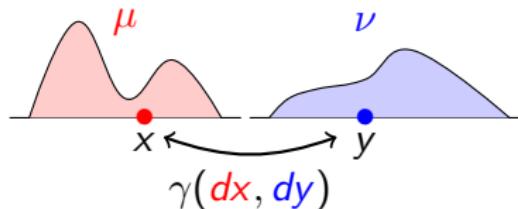
4 Une application

Flot de gradient Hele-Shaw

Théorie classique

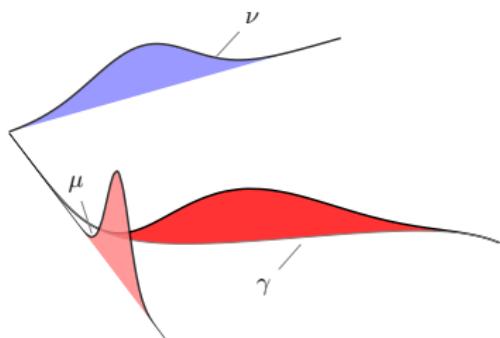
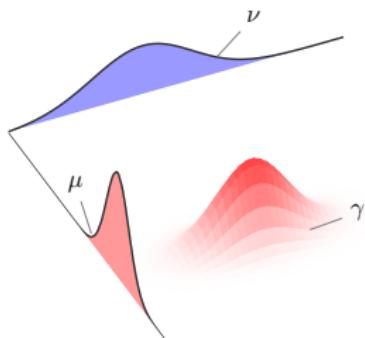
Ingrédients

- Espace ambiant \mathcal{X} (métrique compact dans cet exposé)
- Fonction de coût $c : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$
- Deux mesures de probabilité μ, ν sur \mathcal{X}



Définition (Formulation de Kantorovich)

$$\min \left\{ \int_{\mathcal{X} \times \mathcal{X}} c(x, y) \gamma(dx, dy) ; \gamma \in \Pi(\mu, \nu) \right\}$$



Définition (Couplages)

$$\Pi(\mu, \nu) := \left\{ \gamma \in M_+(\mathcal{X} \times \mathcal{X}) : \pi_1^* \gamma = \mu, \pi_2^* \gamma = \nu \right\}$$

Généralise : applications $\mathcal{X} \rightarrow \mathcal{X}$, permutations

Propriétés : convexe, faiblement compact

Quelques propriétés

Théorème (Distances de Wasserstein)

Si $(\mathcal{X}, \text{dist})$ est un espace géodésique, $p \geq 1$, alors la fonction

$$W_p(\mu, \nu)^p := \min \left\{ \int_{\mathcal{X} \times \mathcal{X}} \text{dist}(x, y)^p \gamma(dx, dy) : \gamma \in \Pi(\mu, \nu) \right\}$$

définit une métrique géodésique sur $P(\mathcal{X})$ qui métrise la convergence faible.

Quelques propriétés

Théorème (Distances de Wasserstein)

Si $(\mathcal{X}, \text{dist})$ est un espace géodésique, $p \geq 1$, alors la fonction

$$W_p(\mu, \nu)^p := \min \left\{ \int_{\mathcal{X} \times \mathcal{X}} \text{dist}(x, y)^p \gamma(dx, dy) : \gamma \in \Pi(\mu, \nu) \right\}$$

définit une métrique géodésique sur $P(\mathcal{X})$ qui métrise la convergence faible.

Formulation dynamique (Benamou et Brenier)

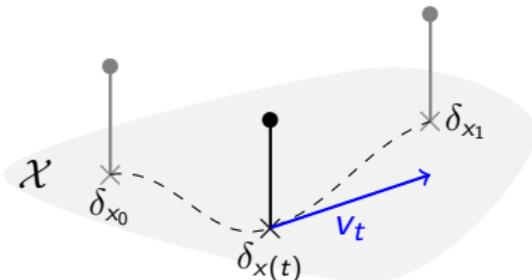
Pour μ, ν mesures de probabilité sur $\Omega \subset \mathbb{R}^d$, on a

$$W_p(\mu, \nu)^p = \min_{(\rho_t, v_t)_{t \in [0, 1]}} \left(\int_0^1 \int_{\mathbb{R}^d} \|v(t, x)\|^p d\rho_t(x) dt \right)$$

sous les contraintes $\partial_t \rho_t = -\text{div}(\rho_t v_t)$ et $(\rho_0, \rho_1) = (\mu, \nu)$.

Plus généralement : Lagrangien convexe $L(v(t, x))$.

Du dynamique au couplage



Coût du plus court chemin

Étant donné un Lagrangien convexe $\mathcal{L} : \mathbb{R}^d \rightarrow \mathbb{R}$, soit

$$c_{\mathcal{L}}(x_0, x_1) := \inf \int_0^1 \mathcal{L}(v(t)) dt$$

où $\delta_{x(t)}$ interpole continûment entre δ_{x_0} et δ_{x_1} et $v_t = x'(t)$.

Équivalence des problèmes (Benamou et Brenier, Jimenez, Villani) :

dynamique avec \mathcal{L} \Leftrightarrow couplage avec $c_{\mathcal{L}}$

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

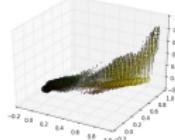
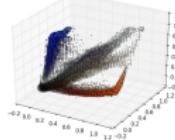
3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

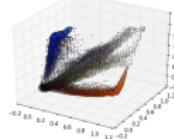
4 Une application

Flot de gradient Hele-Shaw

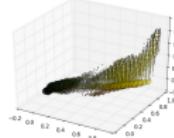


(Delon *et al.*)
(Rabin, Papadakis)

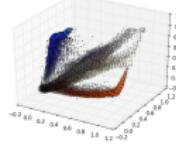
Transfert de couleur



classique



(Delon *et al.*)
(Rabin, Papadakis)



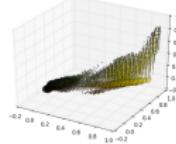
Transfert de couleur



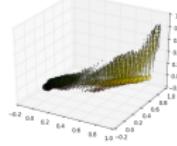
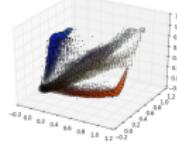
classique



partiel



(Delon *et al.*)
(Rabin, Papadakis)



(Delon *et al.*)
(Rabin, Papadakis)

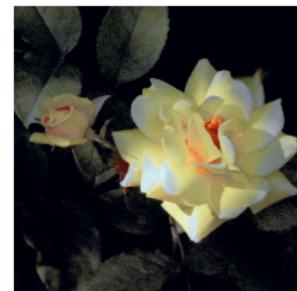
Transfert de couleur



classique



partiel



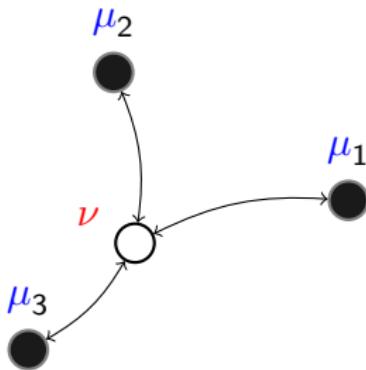
non équilibré (\widehat{W}_2)

Barycentres / moyennes de Fréchet

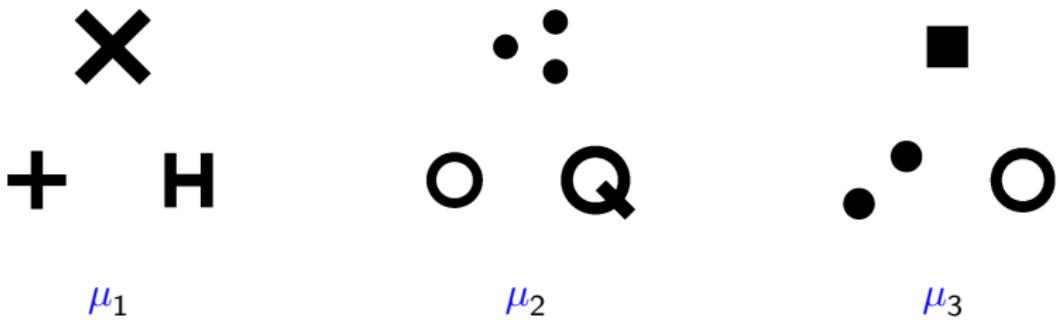
Barycentres de Wasserstein (Aguech et Carlier)

- μ_1, \dots, μ_n des mesures de probabilité
- w_1, \dots, w_n des poids positifs
- Définition du barycentre ν^* :

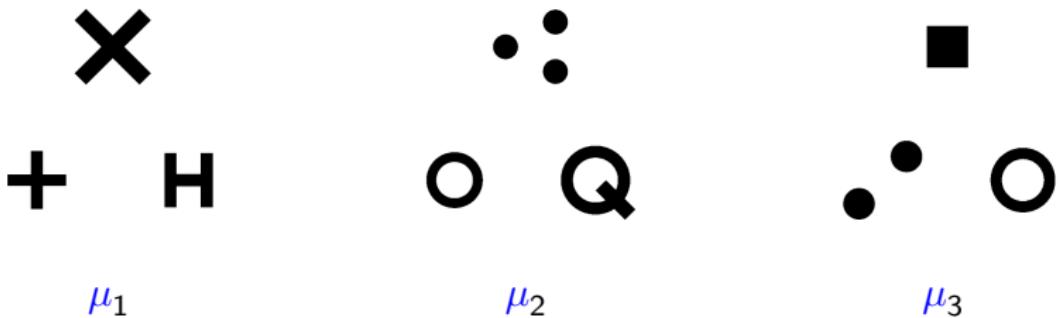
$$\nu^* \in \arg \min_{\nu \in P(\mathcal{X})} \sum_{k=1}^n w_k W_2^2(\mu_k, \nu)$$



Barycentres de mesures

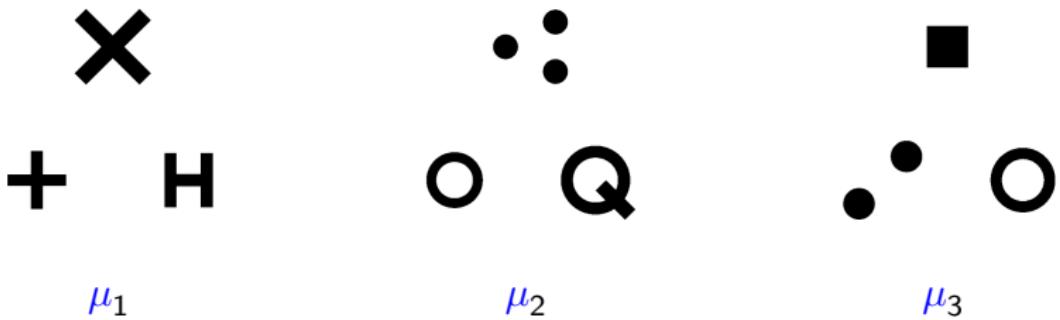


Barycentres de mesures



Iso-barycentre ν^* pour W_2

Barycentres de mesures



Iso-barycentre ν^* pour W_2

Iso-barycentre ν^* pour \widehat{W}_2

Flots de gradient

Filtre de gradient (formel)

Soit E un espace métrique, $F : E \rightarrow \bar{\mathbb{R}}$ et $\mu_0 \in E$,

$$\partial_t \mu_t = -\operatorname{grad}_E F(\mu_t) \quad t > 0$$

métrique	$\operatorname{grad}_E F(\mu)$
L^2	$F'(\mu)$
Hellinger	$\mu F'(\mu)$
W_2	$-\operatorname{div}(\mu \nabla F'(\mu))$

Exemple : l'opérateur Laplacien $-\Delta$ est le gradient de...

- l'énergie de l'énergie de Dirichlet pour L^2 : $F(\mu) = \int |\nabla \mu|_2^2$
- l'entropie pour W_2 : $F(\mu) = \int \mu(\log \mu - 1)$

Flots de gradients Wasserstein

Intérêt

- caractérisent certaines EDP d'évolution

$$\partial_t \mu_t + \operatorname{div}(\mu_t v_t) = 0$$

- théorique : existence, unicité, convergence
- numérique : conservation de masse, positivité intrinsèques

Mouvement de foule
(Roudneff-Chupin, Maury et Santambrogio)

Structure métrique pour les équations avec variations de masse ?

Références et historique

État des lieux en 2015

- re-normalisation
- transport partiel optimal (Kantorovich), (Mc Cann, Cafarelli), (Figalli), (Piccoli, Rossi)
- divers modèles spécifiques (Benamou), (Lombardi, Maître), (Rumpf *et al.*)

Références et historique

État des lieux en 2015

- re-normalisation
- transport partiel optimal (Kantorovich), (Mc Cann, Cafarelli), (Figalli), (Piccoli, Rossi)
- divers modèles spécifiques (Benamou), (Lombardi, Maître), (Rumpf *et al.*)

Depuis 2015

- Modèle quadratique (Liero *et al.*), (Kondratiev *et al.*), (Chayes, Lei), (**C.** *et al.*)
- Théorie générale (Liero *et al.*), (**C.** *et al.*)
- Applications émergentes : imagerie (Feydy *et al.*), apprentissage (Schiebinger *et al.*), (Frogner *et al.*)

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

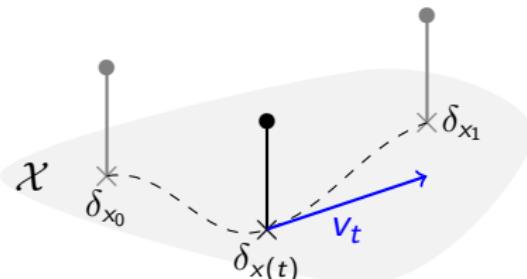
Approche dynamique

Algorithmes de “scaling”

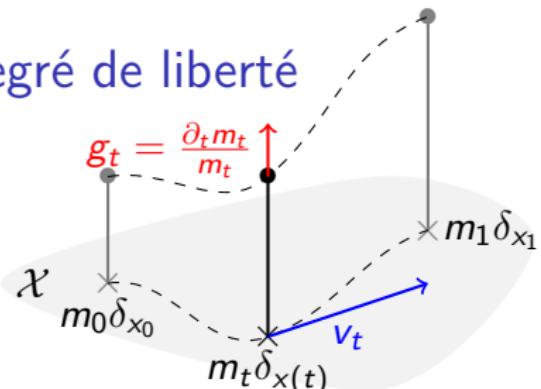
4 Une application

Flot de gradient Hele-Shaw

Ajout d'un degré de liberté

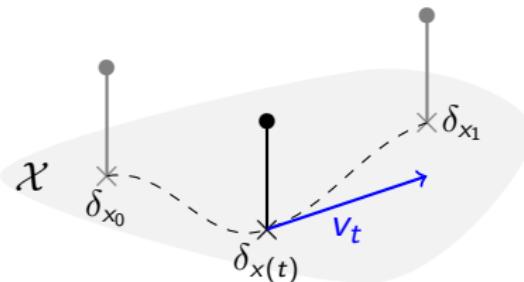


Classique

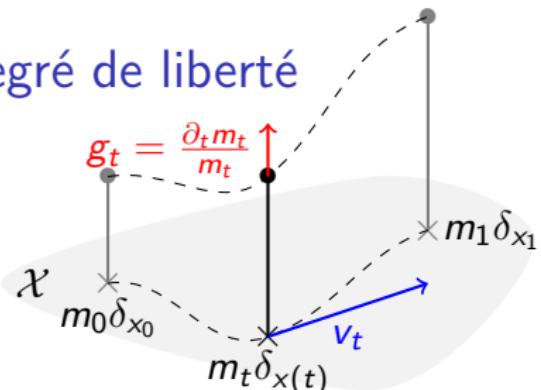


Non-équilibré

Ajout d'un degré de liberté



Classique



Non-équilibré

Formulation dynamique (C. et al.)

Pour $\mu, \nu \in M_+(\Omega)$ mesures de masses *potentiellement* différentes, soit

$$C_L(\mu, \nu) := \min_{(\rho_t, \mathbf{v}_t, \mathbf{g}_t)_{t \in [0, 1]}} \int_0^1 \int_{\mathbb{R}^d} L(\mathbf{v}_t(x), \mathbf{g}_t(x)) d\rho_t(x) dt$$

sous contraintes $\partial_t \rho_t = -\operatorname{div}(\rho_t \mathbf{v}_t) + \rho_t \mathbf{g}_t$ et $(\rho_0, \rho_1) = (\mu, \nu)$.

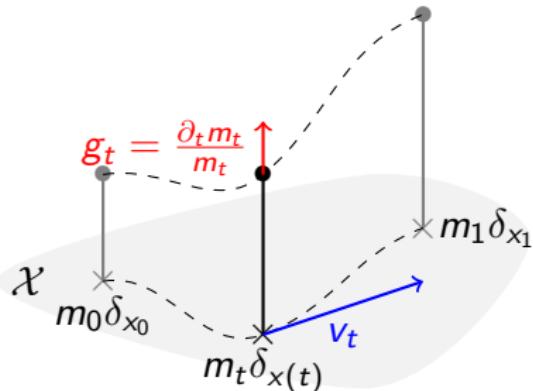
- L Lagrangien convexe, continu, minimal en $L(0, 0) = 0$.

Propriétés

Proposition (C. et al.)

- $C_L(\mu, \nu)$ est finie, existence de minimiseurs
- C_L continue pour la convergence faible
- si L est pair et p -homogène alors $C_L^{1/p}$ définit une métrique sur $M_+(\Omega)$

Du dynamique au couplage (I)



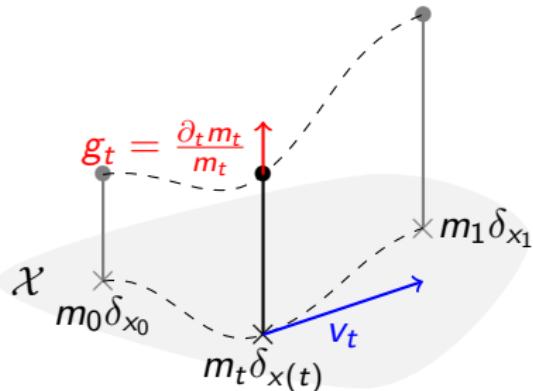
Coût du plus court chemin (cas non-équilibré)

Étant donné un Lagrangien convexe \mathcal{L} ,

$$c_{\mathcal{L}}((x_0, m_0), (x_1, m_1)) = \inf \int_0^1 \mathcal{L}(v(t), g(t)) m(t) dt$$

où $m(t)\delta_{x(t)}$ interpole continûment entre $m_0\delta_{x_0}$ et $m_1\delta_{x_1}$.

Du dynamique au couplage (I)



Coût du plus court chemin (cas non-équilibré)

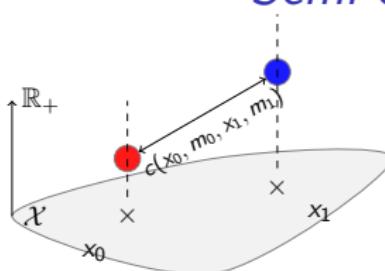
Étant donné un Lagrangien convexe L ,

$$c_L((x_0, m_0), (x_1, m_1)) = \inf \int_0^1 L(v(t), g(t)) m(t) dt$$

où $m(t) \delta_{x(t)}$ interpole continûment entre $m_0 \delta_{x_0}$ et $m_1 \delta_{x_1}$.

- Quelle formulation *couplage* associée ?

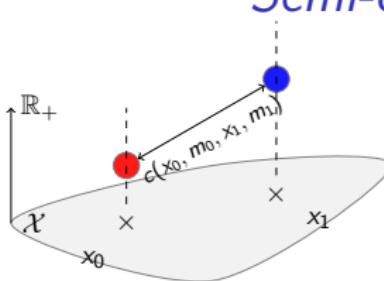
Semi-couplages



Semi-couplages

Mesures $\gamma_0, \gamma_1 \in M_+(\mathcal{X}^2)$ telles que $\pi_#^1 \gamma_0 = \mu$ et $\pi_#^2 \gamma_1 = \nu$.

Semi-couplages



Semi-couplages

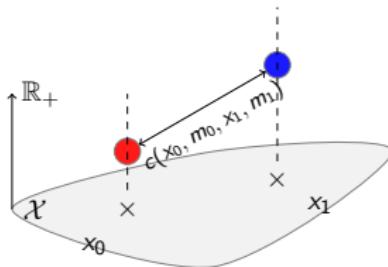
Mesures $\gamma_0, \gamma_1 \in M_+(\mathcal{X}^2)$ telles que $\pi_\#^1 \gamma_0 = \mu$ et $\pi_\#^2 \gamma_1 = \nu$.

Définition (Formulation semi-couplages (C. et al.))

Soit $c((x_0, m_0), (x_1, m_1))$ une fonction de coût *sous-linéaire* en (m_0, m_1) . On définit en minimisant sur l'ensemble des semi-couplages (γ_0, γ_1)

$$C_c(\mu, \nu) := \inf \int_{\mathcal{X}^2} c\left((x_0, \frac{d\gamma_0}{d\lambda}), (x_1, \frac{d\gamma_1}{d\lambda})\right) d\lambda \quad \gamma_0, \gamma_1 \ll \lambda.$$

Propriétés



Théorème (Propriété métrique, $p \geq 1$)

Si $c^{1/p}$ est une métrique sur $\text{Cone}(\mathcal{X}) := (\mathcal{X} \times \mathbb{R}_+)/\sim$ alors $C_c^{1/p}$ est une métrique sur $M_+(\mathcal{X})$.

- équivalence avec le problème de transport *lifté* sur $\mathcal{X} \times \mathbb{R}_+$ de (Liero *et al.*)
- formulation duale, continuité faible

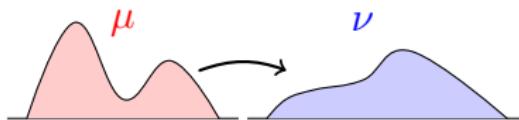
Du dynamique au couplage (II)

Théorème (C. et al.)

Soit $\Omega \subset \mathbb{R}^d$ compact régulier, L un Lagrangien et c_L la régularisation convexe du coût associé. Alors pour toutes mesures positives $\mu, \nu \in M_+(\Omega)$,

$$C_L(\mu, \nu) = C_c(\mu, \nu).$$

Troisième approche

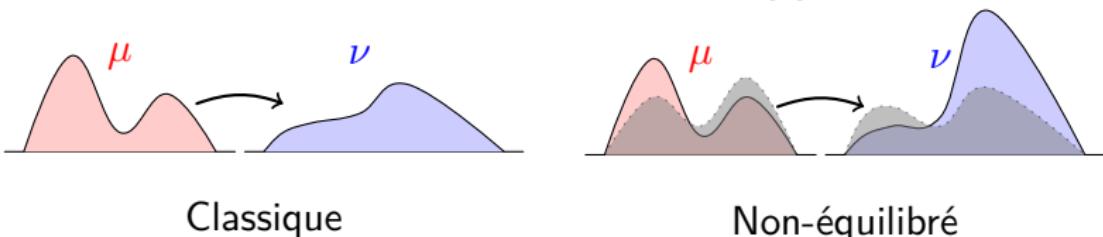


Classique



Non-équilibré

Troisième approche

Problèmes entropie-transport (Liero *et al.*)

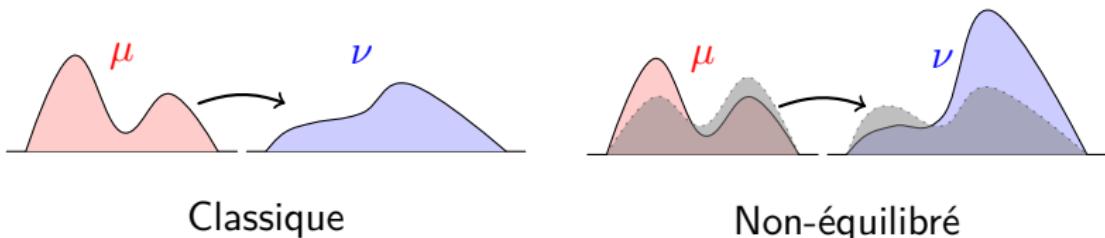
$$C_{c,f}(\mu, \nu) := \min_{\gamma \in M_+(\mathcal{X}^2)} D_f(\pi_\#^1 \gamma | \mu) + D_f(\pi_\#^2 \gamma | \nu) + \int_{\mathcal{X}^2} c \, d\gamma$$

f-divergence (Csiszár)

$$D_f(\mu | \nu) := \int_{\mathcal{X}} f\left(\frac{d\mu}{d\nu}\right) d\nu + \mu^\perp(\mathcal{X}) f'(\infty) \quad \mu, \nu \in M_+(\mathcal{X})$$

avec $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ convexe, minimale en $f(1) = 0$.

Troisième approche



Problèmes entropie-transport (Liero et al.)

$$C_{c,f}(\mu, \nu) := \min_{\gamma \in M_+(\mathcal{X}^2)} D_f(\pi_\#^1 \gamma | \mu) + D_f(\pi_\#^2 \gamma | \nu) + \int_{\mathcal{X}^2} c \, d\gamma$$

Théorème (Liero et al.)

$C_{c,f}$ est équivalent au problème de semi-couplage avec le coût

$$\tilde{c}((x_0, m_0), (x_1, m_1)) := C_{c,f}(m_0 \delta_{x_0}, m_1 \delta_{x_1})$$

Résumé

Introduction

Théorie
classique

Théorie classique

Applications et
limites

Théorie
nouvelle

Formulations

Cas particuliers

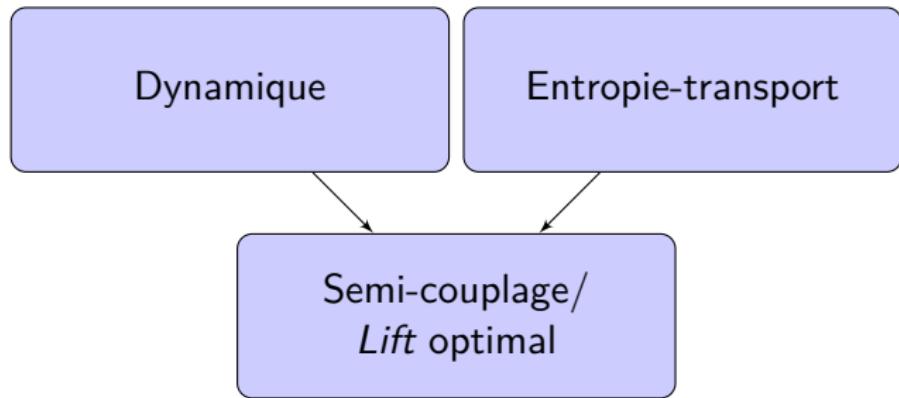
Numérique

Approche dynamique

Algorithmes de
“scaling”

Application

Flot de gradient
Hele-Shaw



Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

4 Une application

Flot de gradient Hele-Shaw

Analogue de W_p

On pose pour $p \geq 1$ et le paramètre d'échelle $\alpha > 0$,

$$\textcolor{blue}{L}(\textcolor{blue}{v}, \textcolor{red}{g}) = (|\textcolor{blue}{v}|/\alpha)^p + (|\textcolor{red}{g}|/p)^p.$$

Définition (Distance $\widehat{W}_{p,\alpha}$)

Pour $\mu, \nu \in M_+(\Omega)$, on définit

$$\widehat{W}_{p,\alpha}(\mu, \nu)^p := \min_{(\rho_t, \textcolor{blue}{v}_t, \textcolor{red}{g}_t)_{t \in [0,1]}} \int_0^1 \int_{\mathbb{R}^d} \textcolor{blue}{L}(\textcolor{blue}{v}_t(x), \textcolor{red}{g}_t(x)) d\rho_t(x) dt$$

sous les contraintes $\partial_t \rho_t = -\operatorname{div}(\rho_t \textcolor{blue}{v}_t) + \rho_t \textcolor{red}{g}_t$ et ρ_0, ρ_1 fixés.

Analogue de W_p

On pose pour $p \geq 1$ et le paramètre d'échelle $\alpha > 0$,

$$\textcolor{blue}{L}(\textcolor{blue}{v}, \textcolor{red}{g}) = (|\textcolor{blue}{v}|/\alpha)^p + (|\textcolor{red}{g}|/p)^p.$$

Définition (Distance $\widehat{W}_{p,\alpha}$)

Pour $\mu, \nu \in M_+(\Omega)$, on définit

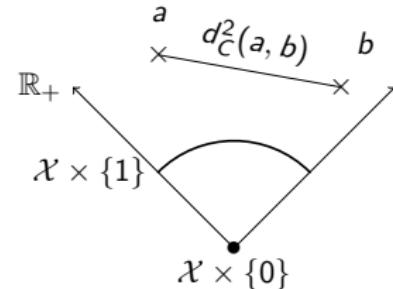
$$\widehat{W}_{p,\alpha}(\mu, \nu)^p := \min_{(\rho_t, \textcolor{blue}{v}_t, \textcolor{red}{g}_t)_{t \in [0,1]}} \int_0^1 \int_{\mathbb{R}^d} \textcolor{blue}{L}(\textcolor{blue}{v}_t(x), \textcolor{red}{g}_t(x)) d\rho_t(x) dt$$

sous les contraintes $\partial_t \rho_t = -\operatorname{div}(\rho_t \textcolor{blue}{v}_t) + \rho_t \textcolor{red}{g}_t$ et ρ_0, ρ_1 fixés.

Propriétés (C. et al.)

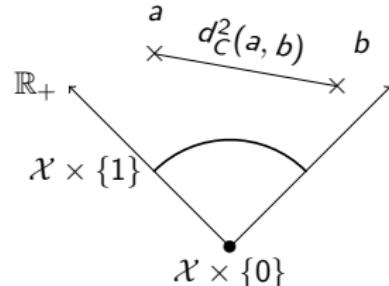
- $\widehat{W}_{p,\alpha}$ est une métrique géodésique, convergence faible
- modèles limites (Γ -convergence) :
 - $\alpha \rightarrow \infty$: transport optimal W_p (généralisé)
 - $\alpha \rightarrow 0$: famille de métriques “verticales” (Hellinger, TV)

Cas quadratique (i)



$$L(v, g) = |\textcolor{blue}{v}|^2 + \frac{1}{4}|\textcolor{red}{g}|^2$$

Cas quadratique (i)



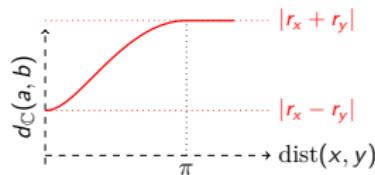
$$L(v, g) = |\textcolor{blue}{v}|^2 + \frac{1}{4}|\textcolor{red}{g}|^2$$

Proposition (Coût de chemin minimal)

Le coût de chemin minimal est donné par la métrique de cône

$$c_{\textcolor{green}{L}}((x_0, m_0), (x_1, m_1)) = d_C((x_0, \sqrt{m_0}), (x_1, \sqrt{m_1}))^2$$

où $d_C((x_0, r_0), (x_1, r_1))^2 = r_0^2 + r_1^2 - 2r_0r_1 \cos(\min\{\text{dist}(x_0, x_1), \pi\})$.



Cas quadratique (ii)

Théorème (Liero et al., C. et al.)

On a une formulation semi-couplage explicite, et la formulation entropie-transport

$$\widehat{W}_2(\mu, \nu)^2 = \min_{\gamma \in M_+(\mathcal{X}^2)} \mathcal{H}(\pi_\#^1 \gamma | \mu) + \mathcal{H}(\pi_\#^2 \gamma | \nu) + \int_{\mathcal{X}^2} c_\ell d\gamma.$$

- $c_\ell(x, y) := -\log \cos^2(\min\{\text{dist}(x, y), \pi/2\})$.
- \mathcal{H} is the relative entropy (or Kullback-Leibler divergence) :

$$\mathcal{H}(\mu | \nu) := \int_{\mathcal{X}} \log(d\mu/d\nu) d\mu - \mu(\mathcal{X}) + \nu(\mathcal{X})$$

- distance nommée *Wasserstein-Fisher-Rao* (C. et al.), ou *Hellinger-Kantorovich* (Liero et al.)
- l'équivalent de W_2 sur $M_+(\mathcal{X})$

Cas $p = 1$

$$L(v, g) = |\textcolor{blue}{v}| + |\textcolor{red}{g}|$$

Proposition

La métrique \widehat{W}_1 est la métrique *bounded Lipschitz* sur $M_+(\mathcal{X})$:

$$\widehat{W}_1(\mu, \nu) = \sup_{\varphi: \mathcal{X} \rightarrow \mathbb{R}} \left\{ \int_{\mathcal{X}} \varphi d(\mu - \nu) ; \|\varphi\|_{\infty} \leq 1, \text{Lip}(\varphi) \leq 1 \right\}$$

Transport partiel optimal

Théorème (Piccoli et Rossi, C. et al.)

Pour un Lagrangien de la forme

$$L(\mathbf{v}, \mathbf{g}) = \tilde{L}(\mathbf{v}) + |\mathbf{g}|,$$

le problème statique équivalent est une formulation du transport partiel optimal.

Transport partiel optimal

Théorème (Piccoli et Rossi, C. et al.)

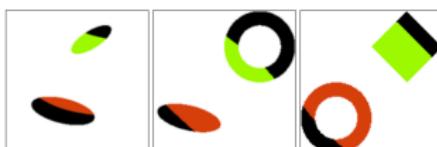
Pour un Lagrangien de la forme

$$L(\mathbf{v}, \mathbf{g}) = \tilde{L}(\mathbf{v}) + |\mathbf{g}|,$$

le problème statique équivalent est une formulation du transport partiel optimal.

Transport partiel optimal

Soient $\mu, \nu \in M_+(\mathcal{X})$ telles que $m \leq \mu(\mathcal{X}), \nu(\mathcal{X})$. Trouver la méthode optimale pour déplacer une quantité de masse m parmi la masse disponible.



(Benamou et al.)

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

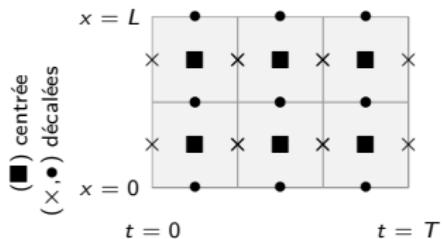
4 Une application

Flot de gradient Hele-Shaw

Discrétisation

$$\min_{\rho, \mathbf{v}, \mathbf{g}} \left\{ \int_0^1 \int_{\Omega} \mathcal{L}(\mathbf{v}, \mathbf{g}) d\rho ; \partial_t \rho + \operatorname{div}(\rho \mathbf{v}) = \mathbf{g} \rho \text{ et } (\rho_0, \rho_1) = (\mu, \nu) \right\}$$

- en les variables $(\rho, \omega, \zeta) = (\rho, \mathbf{v}\rho, \mathbf{g}\rho)$
 \Rightarrow fonctionnelle convexe, contraintes linéaires
- X, \tilde{X} densités discrétisées sur grilles centrées et décalées



Problème convexe non lisse de dimension finie

$$\min_{X, \tilde{X}} \left\{ F(X) ; A(\tilde{X}) = b \text{ et } X = \operatorname{interp}(\tilde{X}) \right\}$$

Approche dynamique

- résolution avec algorithmes *proximaux*, adaptation de (Papadakis *et al.*)
- opérateurs proximaux calculables en complexité quasi-linéaire

FIGURE – Geodesics for densities on \mathbb{R}^2

Hellinger

W_2

partiel

\widehat{W}_2

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

4 Une application

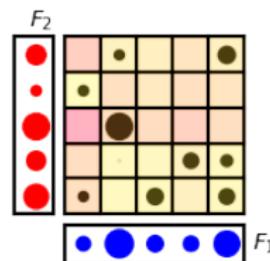
Flot de gradient Hele-Shaw

Formulation générique

- fonction de coût $c : \mathcal{X}^2 \rightarrow \mathbb{R}$
- fonctions des marginales F_1, F_2 convexes

Résoudre

$$\min_{\gamma \in M_+(\mathcal{X}^2)} \int_{\mathcal{X}^2} c \cdot d\gamma + F_1(\pi_1^\# \gamma) + F_2(\pi_2^\# \gamma)$$



- transport optimal classique
- problèmes entropie-transport
- barycentres, flots de gradient...

Régularisation entropique

Suivant (Cuturi),

$$\min_{\gamma \in M_+(\mathcal{X}^2)} \int_{\mathcal{X}^2} c \cdot d\gamma + F_1(\pi_\#^1 \gamma) + F_2(\pi_\#^2 \gamma) + \epsilon \mathcal{H}(\gamma)$$

où $\epsilon > 0$ est l'*importance* de la régularisation.

Algorithmes de *scaling*

Théorème (Optimalité)

Soit le noyau $K(x, y) = \exp(-c(x, y)/\epsilon)$. Sous hypothèses, il existe $a : \mathcal{X} \rightarrow \mathbb{R}_+$ et $b : \mathcal{X} \rightarrow \mathbb{R}_+$ tels qu'à optimalité

$$\gamma_{\text{opt}} = a(x)K(x, y)b(y)$$

Algorithmes de scaling

Théorème (Optimalité)

Soit le noyau $K(x, y) = \exp(-c(x, y)/\epsilon)$. Sous hypothèses, il existe $a : \mathcal{X} \rightarrow \mathbb{R}_+$ et $b : \mathcal{X} \rightarrow \mathbb{R}_+$ tels qu'à optimalité

$$\gamma_{\text{opt}} = a(x)K(x, y)b(y)$$

Algorithme de scaling (C. et al.)

- ① initialiser $b = 1_m$ et répéter jusqu'à convergence
 - ① $a \leftarrow \text{prox}_{F_1}^{\mathcal{H}}(Kb) \oslash (Kb)$
 - ② $b \leftarrow \text{prox}_{F_2}^{\mathcal{H}}(K^T a) \oslash (K^T a)$
- ② retourner $\gamma_{\text{opt}} = (a_i K_{i,j} b_j)_{i,j}$.

$$\text{prox}_F^{\mathcal{H}}(\bar{s}) := \arg \min \{ F(s) + \epsilon \mathcal{H}(s|\bar{s}) \}$$

Convergence linéaire

Métrique de Hilbert sur L_{++}^{∞}

$$d_H(a, b) := \log(\sup(a/b) \sup(b/a))$$

**Proposition (Franklin,
Lorenz)**

Convergence linéaire des
itérées de Sinkhorn pour d_H .

Convergence linéaire

Métrique de Hilbert sur L_{++}^{∞}

$$d_H(a, b) := \log(\sup(a/b) \sup(b/a))$$

Proposition (Franklin, Lorenz)

Convergence linéaire des
itérées de Sinkhorn pour d_H .

Métrique de Thompson sur L_{++}^{∞}

$$d_T(a, b) := \log \max(\sup(a/b), \sup(b/a))$$

Proposition (C. et al.)

$s \mapsto \text{prox}_F^{\mathcal{H}}(s)/s$
est non expansive pour d_T et
contractante si $F = \mathcal{H}$.

Convergence linéaire

Métrique de Hilbert sur L_{++}^{∞}

$$d_H(a, b) := \log(\sup(a/b) \sup(b/a))$$

Proposition (Franklin, Lorenz)

Convergence linéaire des itérées de Sinkhorn pour d_H .

Corollaire

Convergence globale des itérées pour d_T pour la résolution des problèmes liés à \widehat{W}_2 . Vrai en dimension infinie, sous conditions de positivité.

Métrique de Thompson sur L_{++}^{∞}

$$d_T(a, b) := \log \max(\sup(a/b), \sup(b/a))$$

Proposition (C. et al.)

$s \mapsto \text{prox}_F^{\mathcal{H}}(s)/s$ est non expansive pour d_T et contractante si $F = \mathcal{H}$.

Organisation de l'exposé

1 Transport optimal classique

Théorie classique

Applications et limites

2 Modèles non-équilibrés et propriétés

Formulations

Cas particuliers

3 Méthodes numériques

Approche dynamique

Algorithmes de “scaling”

4 Une application

Flot de gradient Hele-Shaw

Flots de gradients (formel)

Rappel : flot de gradient W_2

Pour $F : P(\Omega) \rightarrow \bar{\mathbb{R}}$ et $\rho_0 \in P(\Omega)$,

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla F'(\rho_t))$$

Flots de gradients (formel)

Rappel : flot de gradient W_2

Pour $F : P(\Omega) \rightarrow \bar{\mathbb{R}}$ et $\rho_0 \in P(\Omega)$,

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla F'(\rho_t))$$

Flot de gradient \widehat{W}_2

Pour $F : M_+(\Omega) \rightarrow \bar{\mathbb{R}}$ et $\rho_0 \in M_+(\Omega)$,

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla F'(\rho_t)) - 4\rho_t F'(\rho_t)$$

Flots de gradients (formel)

Rappel : flot de gradient W_2

Pour $F : P(\Omega) \rightarrow \bar{\mathbb{R}}$ et $\rho_0 \in P(\Omega)$,

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla F'(\rho_t))$$

Flot de gradient \widehat{W}_2

Pour $F : M_+(\Omega) \rightarrow \bar{\mathbb{R}}$ et $\rho_0 \in M_+(\Omega)$,

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla F'(\rho_t)) - 4\rho_t F'(\rho_t)$$

- schéma de *splitting* (Gallouet, Monsaingeon, Laborde)
- avec Di Marino, étude d'un problème particulier :
 - EDP dégénérée de type Hele-Shaw
 - vraie structure de flot de gradient
 - similarité avec (Roudneff-Chupin, Santambrogio, Maury)

Flot de gradient dans $M_+(\Omega)$

Théorème (Di Marino et C.)

*Le modèle de croissance de tumeur de type Hele-Shaw
(Perthame et al.)*

$$\begin{aligned}\partial_t \rho_t + \operatorname{div}(\rho_t \nabla p_t) &= 4(1 - p_t)\rho_t \\ p_t(1 - \rho_t) &= 0 \quad \text{and} \quad \rho_t \leq 1\end{aligned}$$

charactérise les flots de gradients pour la métrique \widehat{W}_2 de la fonction

$$G(\rho_t) = \begin{cases} -\rho_t(\Omega) & \text{si } \rho_t \leq 1 \\ \infty & \text{sinon.} \end{cases}$$

- existence de solutions faibles par construction JKO
- unicité par caractérisation EVI (pour Ω convexe)

Soutenance de thèse

Lénaïc Chizat

Introduction

Théorie
classique

Théorie classique

Applications et
limites



Théorie
nouvelle

Formulations

Cas particuliers

Numérique

Approche dynamique

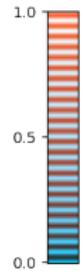
Algorithmes de
"scaling"

Application

Flot de gradient
Hele-Shaw

$$\rho_t$$

$$p_t$$



Conclusion

Transport optimal de mesures positives :

- extension motivée par plusieurs applications
- théorie unifiée incluant modèles nouveaux et existants
- complexité numérique similaire au transport classique

Perspectives :

- transport de mesures à valeurs vectorielles (matrices)