



Stage présenté par

Paolo CHEVALIER

Élève-ingénieur de l'INSA Rennes

Spécialité Mathématiques

Appliqués

2023-2024

INITIATION AU TRANSPORT OPTIMAL ENTROPIQUE

Lieu du Stage

Grupo de Física Matemática (Instituto Superior Técnico)

Maître de Stage

Léonard MONSAINGEON

Correspondant pédagogique INSA

Olivier LEY

$\langle \phi | \mu \rangle$ Grupo de
Física Matemática



TÉCNICO
LISBOA

REMERCIEMENTS

Je tiens tout d'abord à remercier Léonard Monsaingeon, mon maître de stage, qui m'a offert l'opportunité de réaliser ce stage à Técnico. Il a su être présent aussi bien sur les aspects mathématiques qu'humains pour des conseils sur la vie à Lisbonne.

Je remercie aussi Olivier Ley, de m'avoir donné des contacts pour ce stage qui s'inscrit dans une volonté de s'ouvrir à l'étranger, à d'autres cultures et horizons tout en développant mon propre parcours.

Pour finir, je n'oublie pas de remercier toutes les autres personnes passionnées que j'ai rencontrées au cours de mon stage et avec qui j'ai eu l'occasion d'échanger.

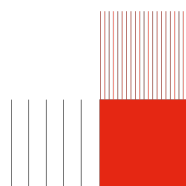


TABLE DES MATIÈRES

Table des matières	1
Introduction	2
1 Notations et notions importantes	2
1.1 Notations	2
1.2 Mesures	3
2 Première formulation du problème (Problème de Monge)	4
2.1 Discret	4
2.2 Push-Foward	4
2.3 Généralisation	5
3 Relaxation (Problème de Monge-Kantorovich)	5
3.1 Intuition/Discret	5
3.2 Généralisation	5
3.3 Distances (Wasserstein)	7
3.4 Problème Dual	7
3.4.1 Discret	7
3.4.2 Généralisation	8
4 Régularisation Entropique	9
4.1 Discret	9
4.2 Généralisation	11
5 Algorithme de Sinkhorn	11
5.1 Algorithme initial	11
5.2 Domaine logarithmique	14
6 Barycentres et interpolations	15
6.1 Formulation à deux plans	15
6.2 Formulation à trois plans	16
Conclusion	20
Annexes	0
Bibliographie	0
Glossaire	0

INTRODUCTION

Étudiant en troisième année à l'INSA de Rennes en Mathématiques Appliquées, j'ai eu l'opportunité à travers mon cursus de réaliser un stage durant l'été. De ce fait, j'ai effectué durant 9 semaines mon stage au Grupo de Física Matemática de l'Instituto Superior Técnico de Lisbonne (Portugal). J'ai, grâce à ce stage, pu découvrir le monde de la recherche et le domaine du TO (Transport Optimal).

Le laboratoire de Recherche dans lequel j'ai effectué mon stage a ses quartiers à l'IST et les chercheurs sont principalement séparés en 4 groupes: systèmes Intégrables et Géométrie, Mécanique et Gravité Quantique, Théorie Spectrale des opérateurs différentiels et enfin celui dont fait partie mon maître de Stage, Analyse Stochastique et TO. Le labo compte dans ses rangs une vingtaine d'enseignants chercheurs et post-docs ainsi qu'une dizaine de doctorants. Le bâtiment de mathématiques de l'IST est particulièrement actif, en raison de la présence de nombreux autres enseignants-chercheurs affiliés à divers laboratoires.

Le plus souvent, j'ai travaillé seul, avec des directives données par mon maître de stage et occasionnellement aussi avec lui.

L'objectif de ce stage était de découvrir la recherche en se familiarisant tout d'abord avec la théorie du TO via l'étude d'ouvrages et d'articles puis par l'implémentation d'algorithmes de résolution numériques pour des problèmes de TO. Les différentes implémentations sur python seront codées uniquement avec numpy sans l'utilisation de packages préexistants.

Quand il évoque le TO, Cédric Villani parle d'un « coup de neuf pour un très vieux problème ». En effet, bien qu'introduit à la fin du XVIII siècle le sujet n'a pas connu un grand intérêt pendant quasiment un siècle et demi jusqu'aux travaux de Kantorovich sur des problèmes de logistique. Vient ensuite un engouement particulier dans les 35 dernières années pour le sujet après les travaux de Brenier et la réalisation du lien important entre le TO et la mécanique des fluides, les équations aux dérivées partielles, la météorologie, la physique statistique et les milieux granulaires pour ne citer que ceux-ci. Toute cette théorie trouve notamment des applications en imagerie (traitement des couleurs/textures), graphisme (manipulation de formes) et machine learning (régression, classification, modèles génératifs).

Ce rapport fera état des connaissances et compétences acquises durant le stage avec aussi bien des aspects théoriques que numériques, les notions seront introduites dans l'ordre le plus intuitif et compréhensible possible. On commencera par la notion de mesure, le problème de Monge en discret et sa généralisation continue grâce aux push-forward. Ce schéma, discret puis généralisation sera suivi pour la relaxation du problème avec le problème de Kantorovich que nous étudierons en détail sans oublier de parler de la distance qu'il induit sur l'espace métrique des mesures de probabilités. On parlera aussi entre autres du problème dual associé, de la régularisation entropique, de résolution numérique grâce à l'algorithme de Sinkhorn (projections alternées) basique ou en domaine logarithmique et enfin de barycentres et interpolations. Des exemples de résolutions numériques seront intégrés tout au long du rapport pour illustrer les concepts théoriques et montrer l'efficacité des algorithmes implémentés.

1 NOTATIONS ET NOTIONS IMPORTANTES

1.1 Notations

On note

- $\Sigma_n := \left\{ a \in \mathbb{R}_+^n : \sum_{i=1}^n a_i = 1 \right\}$
- $\mathbf{1}_n$: vecteur de 1
- Pour $u \in \mathbb{R}^n$, $\text{diag}(u)$: matrice $n \times n$ avec u sur la diagonale et 0 ailleurs
- $\langle \cdot, \cdot \rangle$: le produit scalaire usuel dans la base canonique de \mathbb{R}^d entre deux vecteurs; pour deux matrices de la même taille A et B , $\langle A, B \rangle := \text{tr}(A^T B)$ est la norme de Frobenius.
- $f \oplus g := f \mathbf{1}_m^T + \mathbf{1}_n g^T \in \mathbb{R}^{n \times m}$ pour deux vecteurs $f \in \mathbb{R}^n, g \in \mathbb{R}^m$.

- $\alpha \otimes \beta$ est la mesure produit sur $\mathcal{X} \times \mathcal{Y}$, i.e. $\int_{\mathcal{X} \times \mathcal{Y}} g(x, y) d(\alpha \otimes \beta)(x, y) := \int_{\mathcal{X}} \int_{\mathcal{Y}} g(x, y) d\alpha(x) d\beta(y)$.
- $a \otimes b := ab^T \in \mathbb{R}^{n \times m}$.
- $u \odot v = (u_i v_i) \in \mathbb{R}^n$ pour $u, v \in \mathbb{R}^n$.

1.2 Mesures

Définition 1.2.1 (Mesure discrète): On appellera mesure discrète sur \mathcal{X} de locations $x_1, \dots, x_n \in \mathcal{X}$ et dont les poids respectifs sont donnés par $a = (a_i)_i$

$$\alpha = \sum_{i=1}^n a_i \delta_{x_i}, \text{ i.e. } \alpha(B) = \sum_{i=1}^n a_i \delta_{x_i}(B) = \sum_{i=1}^n a_i \mathbb{1}_B(x_i)$$

Définition 1.2.2 (Espace mesurable): Un espace quelconque et une mesure qu'on lui associe seront notés (\mathcal{X}, α)

Définition 1.2.3 (Mesure): De manière plus générale, définir une mesure nécessite que \mathcal{X} soit muni d'une distance d car une mesure est « définie » par la connaissance de son intégrale par rapport à une fonction continue $f \in C^0(\mathcal{X})$

On peut donc « séparer les mesures en 3 cas »

1. Le cas d'une mesure discrète

$$\int_{\mathcal{X}} f(x) d\alpha(x) = \sum_{i=1}^n a_i f(x_i)$$

2. Le cas d'une mesure sur $\mathcal{X} = \mathbb{R}^d, d \in \mathbb{N}^*$, possédant une densité par rapport à la mesure de Lebesgue, i.e. $d\alpha(x) = \rho_\alpha(x) dx$

$$\int_{\mathcal{X}=\mathbb{R}^d} f(x) d\alpha(x) = \int_{\mathcal{X}} f(x) \rho_\alpha dx$$

3. Le cas d'une mesure complètement arbitraire intégrable par rapport à toute fonction $f \in C^0(\mathcal{X})$

$$\int_{\mathcal{X}} f(x) d\alpha(x) \in \mathbb{R}$$

Attention, si \mathcal{X} n'est pas compact, il faut que f soit à support compact, i.e. $\text{supp}(f) = \overline{\{x \in \mathcal{X} : f(x) \neq 0\}}$ compact ou bien dans une moindre mesure qu'elle ai des limites égales à 0 en l'infini.

Remarque 1.2.1: Dans les écritures utilisées, tout sera « normalisé », i.e. qu'on utilisera des mesures de probabilités de sorte que toutes les mesures soient positives et $\int_{\mathcal{X}} d\alpha(x) = 1$, i.e. on travaillera avec $\mathcal{P}(\mathcal{X})$

Remarque 1.2.2: 3 peut se voir de la manière suivante:

$$\mathcal{L}_\alpha : C^0(\mathcal{X}) \rightarrow \mathbb{R}, \mathcal{L}_\alpha(f) = \int_{\mathcal{X}} f(x) d\alpha(x)$$

La connaissance de $\mathcal{L}_\alpha(f)$ pour tout $f \in C^0(\mathcal{X})$ permet d'inférer la mesure, c'est en ce sens qu'on définit la mesure par une intégrale.

2 PREMIÈRE FORMULATION DU PROBLÈME (PROBLÈME DE MONGE)

2.1 Discret

Soit deux mesures $\alpha = \sum_{i=1}^n a_i \delta_{x_i}$ et $\beta = \sum_{j=1}^m b_j \delta_{y_j}$

Imaginez que vous êtes un commerçant, possédant n usines, m points de ventes représentés respectivement par des distributions α et β telle que la toute la marchandise soit vendue. Pour transporter votre marchandise d'un point $a \in \alpha$ à un point $b \in \beta$ vous payez un coût $c(x,y)$ (par exemple la distance parcourue). Une question que vous vous posez alors est la suivante, quelle est la meilleure répartition de mes marchandises vers mes points de ventes pour diminuer le coût. Monge propose pour ce problème la formulation suivante.

Le problème de Monge est de chercher une application qui à chaque point x_i associe un unique point y_j et qui transporte toute la masse de α en β , ie

$$T : \{x_1, \dots, x_n\} \rightarrow \{y_1, \dots, y_m\} \text{ tq } \forall j \in \llbracket 1, m \rrbracket b_j = \sum_{\{i: T(x_i)=y_j\}} a_i$$

qu'on notera par la suite $T_{\#}\alpha = \beta$

On doit aussi minimiser un coût défini par $c(x,y) \forall (x,y) \in \alpha \times \beta$

Le problème devient donc

$$\min_T \left\{ \sum_{i=1}^n c(x_i, T(x_i)) : T_{\#}\alpha = \beta \right\}$$

Remarque 2.1.1: Dans le cas $n = m$, $a_i = b_i = \frac{1}{n} \forall i \in \llbracket 1, n \rrbracket$ alors la contrainte de conservation implique que T est une bijection et on retrouve le problème d'affectation suivant ré solvable par exemple par un algorithme hongrois:

$$\min_{\sigma \in \text{Perm}(n)} \sum_{i=1}^n C_{i,\sigma(i)} \text{ en notant } C_{i,j} = c(x_i, y_j)$$

2.2 Push-Foward

Définition 2.2.1: Pour une application continue $T : \mathcal{X} \rightarrow \mathcal{Y}$ on définit l'opérateur push-forward correspondant $T_{\#} : \mathcal{M}(\mathcal{X}) \rightarrow \mathcal{M}(\mathcal{Y})$

Pour $\beta = T_{\#}\alpha \in \mathcal{M}(\mathcal{Y})$, $\alpha \in \mathcal{M}(\mathcal{X})$ vérifie

$$\forall f \in C^0(\mathcal{Y}) \int_{\mathcal{Y}} f(y) d\beta(y) = \int_{\mathcal{X}} f(T(x)) d\alpha(x) \Leftrightarrow \forall B \in \mathcal{Y}, \beta(B) = \alpha(T^{-1}(B))$$

Cet opérateur conserve la masse et la positivité

Remarque 2.2.1: La ou l'application $T : \mathcal{X} \rightarrow \mathcal{Y}$ peut être vue comme le déplacement d'un point de \mathcal{X} vers un point de \mathcal{Y} , l'opérateur $T_{\#}$ lui déplace toute la mesure de proba $\in \mathcal{P}(\mathcal{X})$ vers une autre mesure de proba $\in \mathcal{P}(\mathcal{Y})$ en appliquant T à toute les masses élémentaires de la mesure α de \mathcal{X} pour obtenir une masse élémentaire sur \mathcal{Y}

Remarque 2.2.2: Dans la définition intégrale, si on travaille avec des mesures admettant des densités on à $\int_{\mathcal{Y}} h(y) d\beta(y) = \int_{\mathcal{Y}} h(y) \rho_{\beta}(y) dy \stackrel{y=T(x)}{\equiv} \int_{\mathcal{X}} h(T(x)) \rho_{\beta}(T(x)) |J_T(x)| dx = \int_{\mathcal{X}} h(T(x)) \rho_{\alpha}(x) dx$ d'où $|J_T(x)| = \frac{\rho_{\alpha}(x)}{\rho_{\beta}(T(x))}$ ou J_T est le jacobien de T

2.3 Généralisation

Avec deux espaces mesurables $(\mathcal{X}, \alpha), (\mathcal{Y}, \beta)$ le problème devient:

$$\min_T \left\{ \int_{\mathcal{X}} c(x, T(x)) d\alpha(x) : T_{\#} \alpha = \beta \right\}$$

ou $\alpha \in \mathcal{P}(\mathcal{X}), \beta \in \mathcal{P}(\mathcal{Y})$

3 RELAXATION (PROBLÈME DE MONGE-KANTOROVICH)

3.1 Intuition/Discret

La formulation de Monge, bien que intuitive est restrictive dans le sens ou elle peut ne pas admettre de solutions qui conservent la masse (par exemple si il y a plus de points d'arrivés que de départs, voir Fig. 1).

Kantorovich propose donc de relaxer le problème et d'autoriser le mass-splitting, la masse en un point x_i peut être répartie sur plusieurs y_j . Ainsi, plutôt que de chercher une application T on cherche une matrice dite de « couplage » $P \in \mathcal{M}_{n,m}(\mathbb{R}_+)$ ou $P_{i,j}$ décrit la quantité de masse de x_i qui part en y_j .

L'ensemble des matrices de couplages admissibles est le suivant :

$$\Pi(a, b) = \{P \in \mathcal{M}_{n,m}(\mathbb{R}_+) : P \mathbf{1}_m = b \text{ et } P^T \mathbf{1}_n = a\}$$

ou a et b définissent des mesures discrètes et $P \mathbf{1}_m = \left(\sum_j P_{i,j} \right)_i \in \mathbb{R}^n, P^T \mathbf{1}_n = \left(\sum_i P_{i,j} \right)_j \in \mathbb{R}^m$

Le problème devient:

$$L_C(a, b) = \min_{P \in \Pi(a,b)} \langle C, P \rangle = \min_{P \in \Pi(a,b)} \sum_j \sum_i P_{i,j} C_{i,j} \quad (\mathcal{X}^d)$$

Remarque 3.1.1: On a ici un (PL) mais on ne s'attardera pas sur sa résolution directe car généralement peu efficace et souffre de la malédiction du dimensionnement.

Remarque 3.1.2 (Kantorovich discret): Soient $(\mathcal{X}, \alpha), (\mathcal{Y}, \beta)$ avec des mesures discrètes, on stocke dans C les coûts entre les points des supports de α et β , ie $C_{i,j} = c(x_i, y_j)$. Alors $\mathcal{L}_c(\alpha, \beta) = L_c(a, b)$

3.2 Généralisation

On peut étendre la définition de \mathcal{L}_c à des mesures arbitraires en considérant des mesures de probas $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ plutôt qu'une matrice de couplage, le cas échéant imposant juste $\pi = \sum_i \sum_j P_{i,j} \delta_{(x_i, y_j)}$

Ainsi, l'ensemble des mesures admissibles est

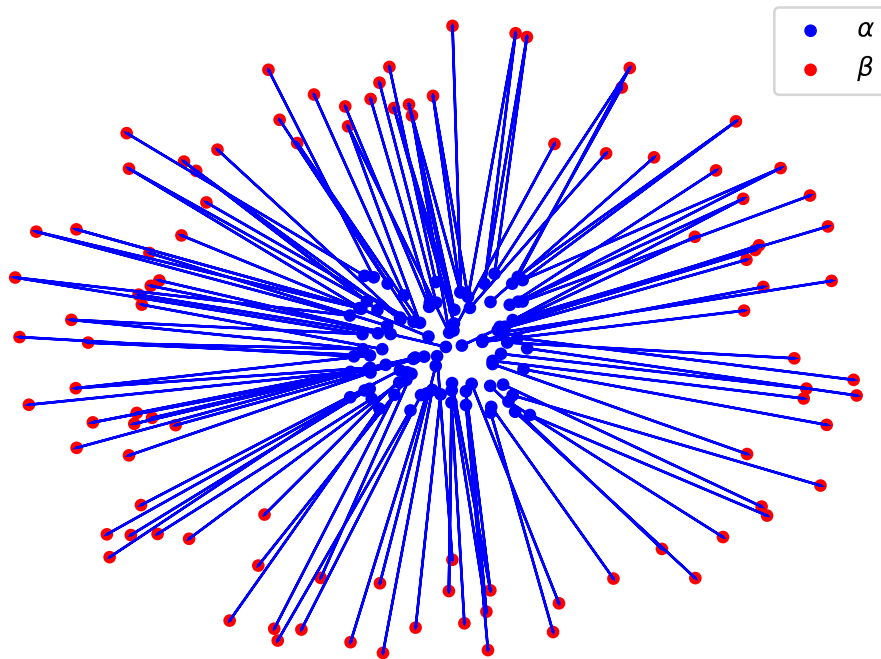
$$\mathcal{U}(\alpha, \beta) = \left\{ \pi \in \mathcal{P}(\mathcal{X}, \mathcal{Y}) : \underbrace{P_{\mathcal{X}\#} \pi}_{=\pi_{\mathcal{X}}} = \alpha \text{ et } \underbrace{P_{\mathcal{Y}\#} \pi}_{=\pi_{\mathcal{Y}}} = \beta \right\}$$

Ou $P_{\mathcal{X}\#}, P_{\mathcal{Y}\#}$ sont les push forward des projections sur \mathcal{X} et \mathcal{Y} respectivement

Le problème devient donc:

$$\mathcal{L}_c(\alpha, \beta) = \min_{\pi \in \mathcal{U}(\alpha, \beta)} \iint_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dx, dy) \quad (\mathcal{X})$$

Optimal transport map



Optimal transport map

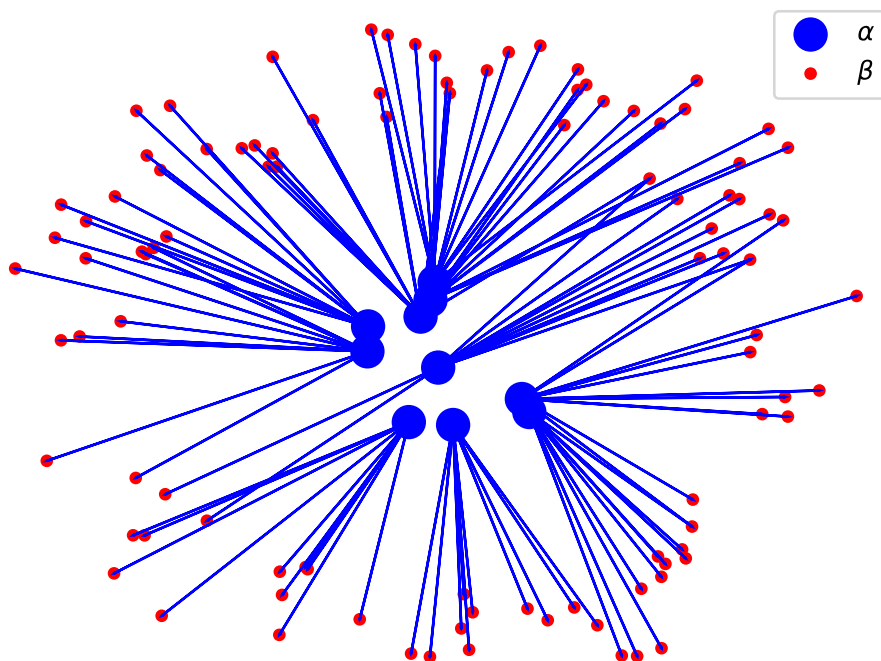


Fig. 1. – En haut, un exemple dans lequel la formulation de Monge permet de résoudre le problème entre deux mesures α, β uniformes. En bas un exemple dans lequel la formulation de Monge n'est pas valable et où du mass-splitting à lieu.

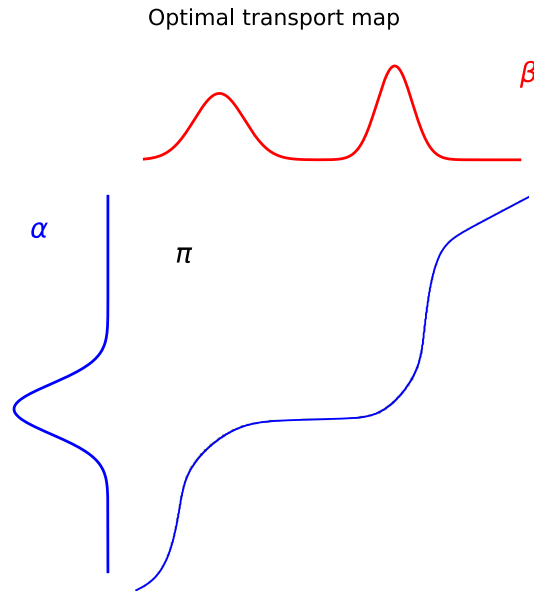


Fig. 2. – Exemple de TO obtenu entre deux gaussiennes avec une discrétisation de $N=256$

3.3 Distances (Wasserstein)

En fait, le TO permet de définir une distance entre les mesures. Dans le cas discret, la proposition est la suivante:

Proposition 3.3.1 (Distance de Wasserstein discrète): On suppose $n = m$ et $p \geq 1$, $C = D^p = (D_{i,j}^p)_{i,j} \in \mathcal{M}_n(\mathbb{R})$ ou $D \in \mathcal{M}_n(\mathbb{R})$ est une distance sur \mathcal{X} Alors $W_p(a, b) := L_{D^p}(a, b)^{\frac{1}{p}}$ définit une distance sur Σ_n

Preuve: Découle de Proposition 3.3.2 □

Ce qui se généralise assez naturellement aux mesures arbitraires:

Proposition 3.3.2 (Distance de Wasserstein pour des mesures arbitraires): On suppose $\mathcal{X} = \mathcal{Y}$ et $p \geq 1$, $c(x, y) = d(x, y)^p$ ou d est une distance sur \mathcal{X} Alors $W_p(\alpha, \beta) := \mathcal{L}_{d^p}(\alpha, \beta)^{\frac{1}{p}}$ définit une distance sur $\mathcal{P}(\mathcal{X})$

Preuve: Une preuve est proposée dans (Villani, 2003) □

Remarque 3.3.1: La distance de Wasserstein permet entre autres de caractériser la convergence en loi d'une v.a

Remarque 3.3.2: Le cas $p = +\infty$ est omis

3.4 Problème Dual

3.4.1 Discret

Le problème (\mathcal{X}^d) discret peut se réécrire à l'aide de son Lagrangien en introduisant $(\varphi, \psi) \in \mathbb{R}^n \times \mathbb{R}^m$

$$\begin{aligned} \min_{P \geq 0} \max_{(\varphi, \psi)} \mathcal{L}(P, \varphi, \psi) &= \min_{P \geq 0} \max_{(\varphi, \psi)} \langle C, P \rangle + \langle \varphi, a - P \mathbf{1}_m \rangle - \langle \psi, b - P^T \mathbf{1}_n \rangle \\ &= \min_{P \geq 0} \max_{(\varphi, \psi)} \langle \varphi, a \rangle + \langle \psi, b \rangle + \langle P, C - \varphi \mathbf{1}_m^T - \mathbf{1}_n \psi^T \rangle \end{aligned}$$

Le problème dual est le suivant

$$\begin{aligned} & \max_{(\varphi, \psi)} \min_{P \geq 0} \langle \varphi, a \rangle + \langle \psi, b \rangle + \langle P, C - \varphi \mathbb{1}_m^T - \mathbb{1}_n \psi^T \rangle \\ &= \max_{(\varphi, \psi)} \left(\langle \varphi, a \rangle + \langle \psi, b \rangle + \min_{P \geq 0} \langle P, C - \varphi \mathbb{1}_m^T - \mathbb{1}_n \psi^T \rangle \right) \\ &= \max_{(\varphi, \psi) \in R(C)} \langle \varphi, a \rangle + \langle \psi, b \rangle \end{aligned} \quad (\mathcal{DK}^d)$$

car $\min_{P \geq 0} \langle P, M \rangle = \begin{cases} 0 & \text{si } M \geq 0 \\ -\infty & \text{sinon} \end{cases}$ où $R(C)$ correspond à la contrainte

$$C - \varphi \mathbb{1}_m^T - \mathbb{1}_n \psi^T \geq 0$$

$$\text{donc } R(C) = \left\{ (\varphi, \psi) \in \mathbb{R}^n \times \mathbb{R}^m : \underbrace{\varphi \oplus \psi \leq C}_{\varphi_i + \psi_j \leq C_{i,j}} \right\}$$

Théorème 3.4.1.1 (Equivalence Kantorovich et Dual Kantorovich): Il y a dualité forte entre (\mathcal{K}^d) et (\mathcal{DK}^d) , ie « $\min(K) = \max(DK)$ ». La solution de (\mathcal{K}^d) est exactement celle de (\mathcal{DK}^d) et on peut retrouver P grâce à φ et ψ

Preuve: Une preuve est donnée dans (Peyré et Cuturi, 2020) □

Remarque 3.4.1.1 (Signification du problème dual): Ce nouveau problème admet une signification particulière, il représente la situation ou plutôt que de transporter vous même vos marchandises, vous décidez de faire appel à un transporteur extérieur qui vous charge φ_i pour la prise en charge des marchandises et ψ_j pour la décharge des marchandises, indépendamment du chemin. Ce dernier vous garanti que quoi qu'il arrive vous payerez au plus le même prix que si vous transportiez par vous même. Alors Théorème 3.4.1.1 signifie que la solution qui arrange le marchand est aussi celle qui arrange le transporteur.

3.4.2 Généralisation

Encore une fois, on généralise ce principe à des mesures arbitraires

Le problème dual devient:

$$\mathcal{L}_c(\alpha, \beta) = \sup_{(\varphi, \psi) \in \mathcal{R}(C)} \int_{\mathcal{X}} \varphi(x) d\alpha(x) + \int_{\mathcal{Y}} \psi(y) d\beta(y) \quad (\mathcal{DK})$$

$$\text{ou } \mathcal{R}(C) = \{(\varphi, \psi) \in C^0(\mathcal{X}) \times C^0(\mathcal{Y}) : \forall (x, y) \in (\mathcal{X}, \mathcal{Y}), \varphi(x) + \psi(y) \leq c(x, y)\}$$

Remarque 3.4.2.1: φ et ψ sont appelés les potentiels duals

Théorème 3.4.2.1 (de Brenier): Dans le cas $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ et $c(x, y) = \frac{\|x - y\|^2}{2}$, si au moins une des deux mesures (notée α) possède une densité ρ_α par rapport à la mesure de Lebesgue alors le π optimal dans (\mathcal{K}) est unique et supporté par le graphe $(x, T(x))$ d'une application de Monge $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$, cela signifie que $\pi = (\text{Id}, T)_\# \alpha$, ie $\forall f \in C^0(\mathcal{X} \times \mathcal{Y}), \iint_{\mathcal{X} \times \mathcal{Y}} f(x, y) d\pi(x, y) = \int_{\mathcal{X}} f(x, T(x)) d\alpha(x)$

De plus, cette application T est définie comme le gradient d'une fonction convexe, $T(x) = \nabla \left(\underbrace{\frac{\|x\|^2}{2} - \varphi(x)}_{f(x)} \right)$
 ou φ est le potentiel dual et f l'unique fonction convexe (à constante près) qui vérifie $(\nabla f)_\# \alpha = \beta$

Preuve: Une preuve de ce théorème (assez délicate) est donnée dans (Villani, 2003) □

4 RÉGULARISATION ENTROPIQUE

4.1 Discret

Définition 4.1.1 (Entropie): On définit l'entropie d'une matrice de couplage par

$$H(P) = \sum_i \sum_j P_{i,j} (\log(P_{i,j}) - 1)$$

avec une définition similaire pour les vecteurs sous la convention que $H(a) = +\infty$ si une des entrées a_j est ≥ 0

Remarque 4.1.1: La fonction H est strictement convexe car sa Hessienne est $\text{diag}\left(\frac{1}{P_{i,j}}\right) > 1$

L'idée est donc de régulariser (\mathcal{K}) qui est un problème convexe mais « limite » dans le sens où on le rend strictement convexe par l'ajout d'un terme de régularisation $\varepsilon H(P)$.

On a donc le problème suivant:

$$L_c^\varepsilon(a, b) = \min_{P \in \Pi(a, b)} \langle C, P \rangle + \varepsilon H(P) \quad (\mathcal{K}_\varepsilon^d)$$

Remarque 4.1.2: La régularisation permet de flouter/bruiter/ajouter du désordre à la solution, voir Fig. 3

Proposition 4.1.1: La solution de $(\mathcal{K}_\varepsilon^d)$ converge vers la solution optimale d'entropie minimale parmi les solutions de (\mathcal{K}^d) , ie

$$P_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \text{argmin}\{H(P) : P \in \Pi(a, b) \text{ et } \langle C, P \rangle = L_c(a, b)\}$$

Preuve: Soit $(\varepsilon_k)_{k \geq 0}$ tq $\varepsilon_k \geq 0$ et $\varepsilon_k \xrightarrow{k \rightarrow \infty} 0$. Pour $\varepsilon = \varepsilon_k$ fixé on note P_k la solution de $(\mathcal{K}_\varepsilon^d)$ et P une solution de (\mathcal{K}^d) . Comme $\Pi(a, b)$ est un fermé borné, P_k admet comme valeur d'adhérence $P^* \in \Pi(a, b)$.

Par optimalité de P et P_k (la sous suite qui admet une valeur d'adhérence qu'on ne renommera pas) pour leurs problèmes respectifs on à

$$\begin{cases} \langle C, P \rangle \leq \langle C, P_k \rangle \\ \langle C, P_k \rangle + \varepsilon_k H(P_k) \leq \langle C, P \rangle + \varepsilon_k H(P) \end{cases} \Leftrightarrow 0 \leq \langle C, P_k \rangle - \langle C, P \rangle \leq \varepsilon_k (H(P) - H(P_k))$$

Ainsi, en passant à la limite quand $k \rightarrow \infty$ par continuité de H on obtient $\langle C, P^* \rangle = \langle C, P \rangle = L_c(a, b)$.

En divisant par ε_k et en passant encore à la limite on a que $H(P^*) \leq H(P)$ donc P^* est une solution au problème.

Par stricte convexité de H la solution du problème est unique et la suite P_k converge vers P^* □

Remarque 4.1.3: On a donc en particulier $L_c^\varepsilon(a, b) \xrightarrow{\varepsilon \rightarrow 0} L_c(a, b)$

Définition 4.1.2 (Divergence de Kullback-Leibler): On note

$$\text{KL}(P|K) = \sum_i \sum_j P_{i,j} \log\left(\frac{P_{i,j}}{K_{i,j}}\right) - P_{i,j} + K_{i,j}$$

ou $K_{i,j} = e^{-\frac{c_{i,j}}{\varepsilon}}$

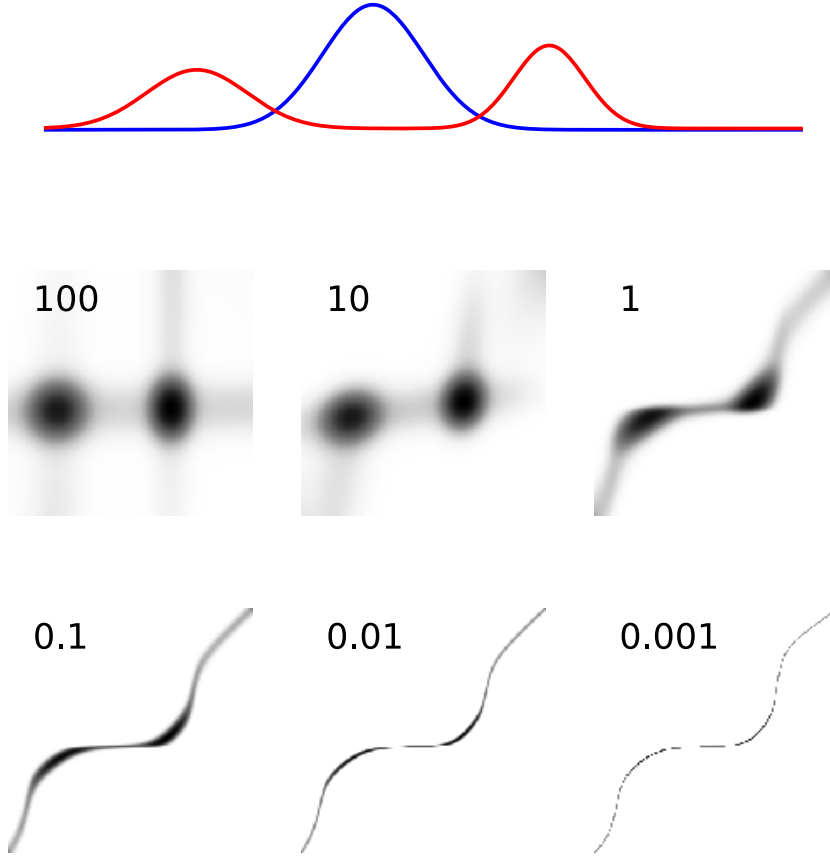


Fig. 3. – En haut: les densités α et β utilisées (Gaussiennes discrétisées avec $N=256$). En bas: Influence de ε sur l'application π , le résultat de Proposition 4.1.1 apparaît.

Proposition 4.1.2: L'unique solution de $(\mathcal{K}_\varepsilon^d)$ est une projection de K sur $\Pi(a, b)$, ie $P_\varepsilon = \text{proj}_{\Pi(a,b)}^{\text{KL}}(K) = \text{argmin}_{P \in \Pi(a,b)} \text{KL}(P|K)$

Preuve: L'unicité vient du fait de la stricte convexité de la fonction objectif.

Ensuite,

$$\begin{aligned}
 \langle C, P \rangle + \varepsilon H(P) &= \varepsilon \sum_{i,j} P_{i,j} \left(\frac{C_{i,j}}{\varepsilon} + \log(P_{i,j}) - 1 \right) \\
 &= \varepsilon \sum_{i,j} P_{i,j} \left(\log \left(e^{\frac{C_{i,j}}{\varepsilon}} \right) + \log(P_{i,j}) - 1 \right) \\
 &= \varepsilon \sum_{i,j} P_{i,j} \left(\log \left(\frac{P_{i,j}}{K_{i,j}} \right) - 1 \right) \\
 &= \varepsilon \sum_{i,j} P_{i,j} \log \left(\frac{P_{i,j}}{K_{i,j}} \right) - P_{i,j} + K_{i,j} - \varepsilon \sum_{i,j} K_{i,j} \\
 &= \varepsilon \text{KL}(P|K) + \text{cst}
 \end{aligned}$$

D'où le résultat. □

Remarque 4.1.4: Pour des mesures discrètes, on note $\mathcal{L}_\varepsilon^\varepsilon(\alpha, \beta) = L_\varepsilon^\varepsilon(a, b)$

4.2 Généralisation

Pour des mesures arbitraires, on note

$$\mathcal{L}_c^\varepsilon(\alpha, \beta) = \min_{\pi \in \mathcal{U}(\alpha, \beta)} \iint_{x \times y} c(x, y) d\pi(x, y) + \varepsilon \text{KL}(\pi | \alpha \otimes \beta) \quad (\mathcal{K}_\varepsilon)$$

ou $d\alpha \otimes d\beta(x, y) = d\alpha(x)d\beta(y)$ et

$$\text{KL}(\pi, \xi) := \begin{cases} \iint_{x \times y} \log\left(\frac{d\pi}{d\xi}(x, y)\right) d\pi(x, y) + \iint_{x \times y} (d\xi(x, y) - d\pi(x, y)) & \text{si } \pi \ll \xi \\ +\infty & \text{sinon} \end{cases}$$

Définition 4.2.1 (Absolue continuité): On dit que π est absolument continue par rapport à μ , noté $\pi \ll \mu$ si π est de la forme (unique) $\pi(dx) = \rho(x)\mu(dx)$ et on note dans ce cas $\rho'(x) = \frac{d\pi}{d\mu}(x)$

5 ALGORITHME DE SINKHORN

5.1 Algorithme initial

Théorème 5.1.1: La solution de $(\mathcal{K}_\varepsilon^d)$ est de la forme $P_{i,j} = u_i K_{i,j} v_j$ avec u et v des vecteurs inconnus de \mathbb{R}_+^n et \mathbb{R}_+^m

Preuve: Le Lagrangien de $(\mathcal{K}_\varepsilon^d)$ s'écrit

$$\mathcal{L}(P, \varphi, \psi) = \langle P, C \rangle + \varepsilon H(P) - \langle \varphi, P \mathbb{1}_m - a \rangle - \langle \psi, P^T \mathbb{1}_n - b \rangle$$

Comme c'est un problème strictement convexe, la condition d'optimalité du 1er ordre sur le Lagrangien nous donne

$$\frac{\partial \mathcal{L}(P_{i,j}, \varphi_i, \psi_j)}{\partial P_{i,j}} = 0 = C_{i,j} + \varepsilon \log(P_{i,j}) - \varphi_i - \psi_j$$

D'où

$$P_{i,j} = \exp\left(\frac{\varphi_i}{\varepsilon}\right) \exp\left(-\frac{C_{i,j}}{\varepsilon}\right) \exp\left(\frac{\psi_j}{\varepsilon}\right) = u_i K_{i,j} v_j$$

□

Cette condition peut se réécrire sous la forme $P = \text{diag}(u)K\text{diag}(v)$

Or P doit appartenir à $U(a, b)$ donc

$$\begin{aligned} \text{diag}(u)K\text{diag}(v)\mathbb{1}_m &= a \\ \text{diag}(u)Kv &= a \\ u \odot Kv &= a \end{aligned}$$

et similairement $v \odot K^T u = b$

On obtient donc l'algorithme de Sinkhorn qui consiste à choisir un v initial puis à itérer progressivement sur les équations ci dessus.

Proposition 5.1.2 (Sinkhorn): $v^{(0)} = \mathbb{1}_m$ et $\forall l \geq 0$ $u^{(l+1)} = \frac{a}{Kv^{(l)}}$, $v^{(l+1)} = \frac{b}{K^T u^{(l+1)}}$ ou les divisions sont à comprendre éléments par éléments. Cet algorithme converge.

Preuve: La preuve est basée sur l'algorithme des projections alternées de Bregman. Voir (Peyré et Cuturi, 2020) □

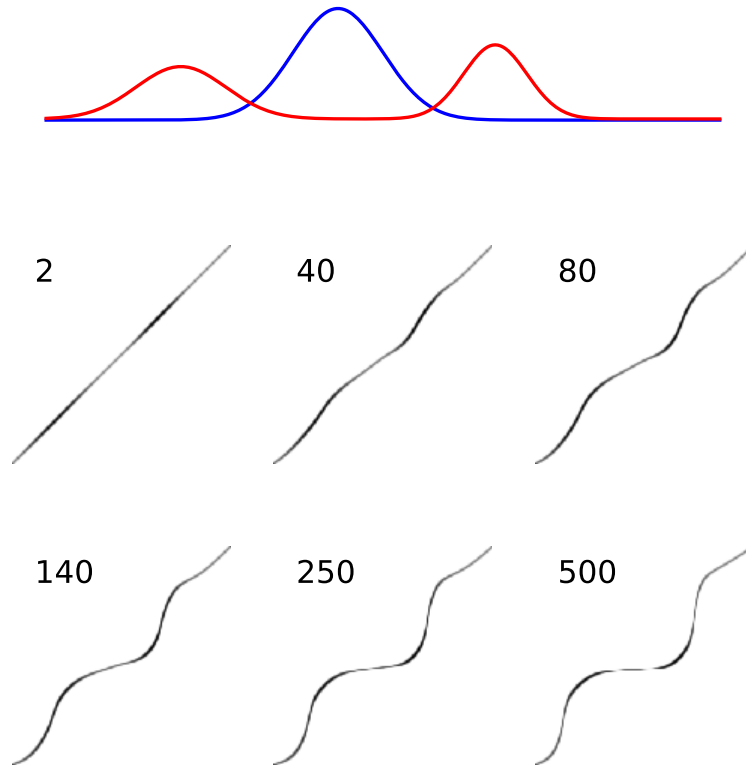


Fig. 4. – En haut: les densités α et β utilisées (Gaussiennes discrétisées avec $N=256$). En bas: Évolution de l'application π avec le nombre d'itérations, les itérations déplacent progressivement la masse de la diagonale.

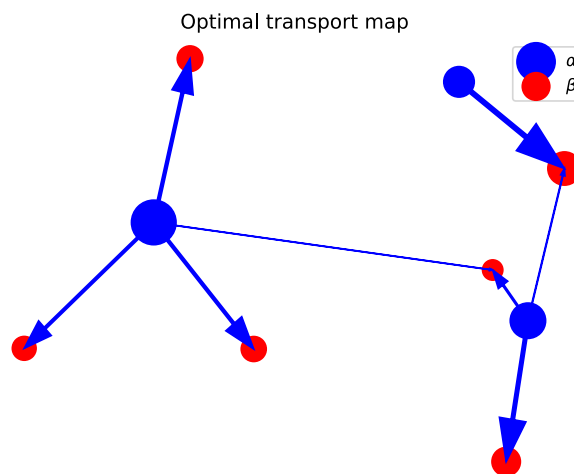


Fig. 5. – TO entre deux mesures discrètes aléatoires α et β avec $\varepsilon = 10^{-5}$

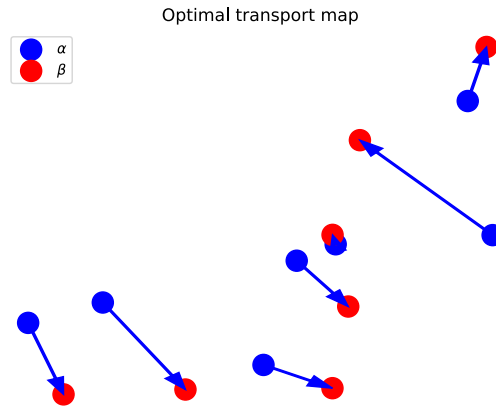


Fig. 6. – TO entre deux mesures discrètes aléatoires α et β **uniformes**, on retrouve bien une bijection entre α et β avec $\varepsilon = 10^{-4}$

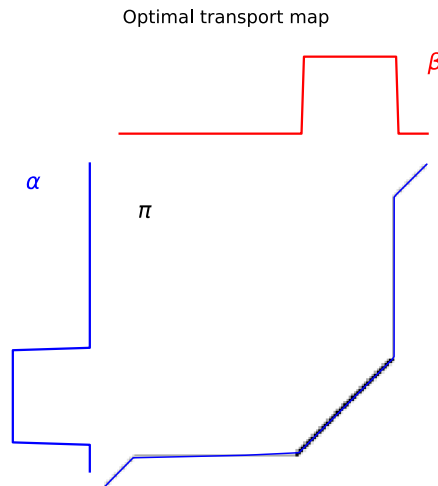


Fig. 7. – TO entre deux « portes » discrétisées avec $N = 128$, α et β . En bleu sur l'application la projection barycentrique de π avec $\varepsilon = 10^{-2}$. On observe bien le résultat de Théorème 3.4.2.1

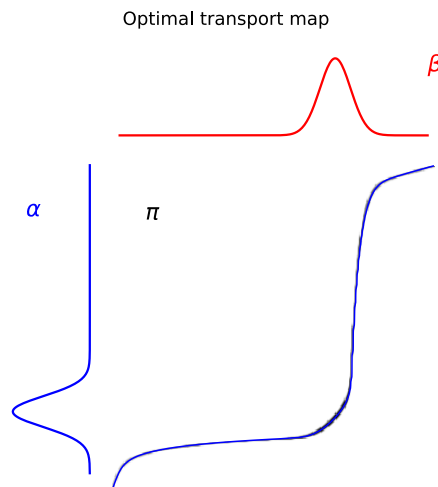


Fig. 8. – TO entre deux gaussiennes centrées réduites discrétisées avec $N = 256$. En bleu sur l'application la projection barycentrique de π avec $\varepsilon = 10^{-2}$. On observe bien encore le résultat de Théorème 3.4.2.1

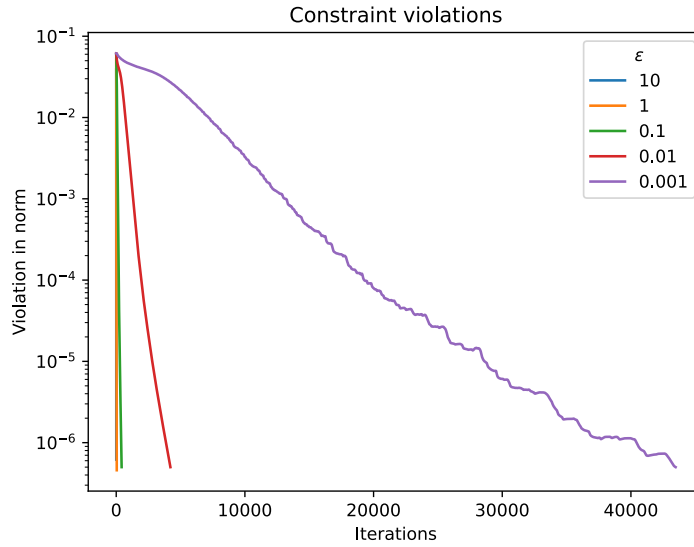


Fig. 9. – Influence du paramètre ε sur la vitesse de convergence pour un TO entre deux gaussiennes discrétisées avec $N=128$. Plus ε est petit moins l’algorithme converge vite, c’est intuitivement logique étant donné que le problème devient de « moins en moins » strictement convexe

5.2 Domaine logarithmique

Bien qu’en théorie l’algorithme de Sinkhorn converge pour tout $\varepsilon > 0$, en pratique on observe assez vite des problèmes de stabilité (underflows et overflows) lors des itérations de Sinkhorn.

On passe alors au domaine logarithmique pour éviter les divisions par le 0 machine. En fait, faire ce changement revient à itérer sur les potentiels duals plutôt que u et v . Ces derniers sont reliés par la relation $(u, v) = \left(e^{\frac{\varphi_i}{\varepsilon}}, e^{\frac{\psi_j}{\varepsilon}} \right)$. Passer aux logarithme nous donne donc les itérations suivantes:

$$\begin{aligned}\varphi^{(l+1)} &= \varepsilon \log(a) - \varepsilon \log\left(K e^{\frac{\psi^{(l)}}{\varepsilon}} \right) \\ \psi^{(l+1)} &= \varepsilon \log(b) - \varepsilon \log\left(K^T e^{\frac{\varphi^{(l+1)}}{\varepsilon}} \right)\end{aligned}$$

Définition 5.2.1 (Softmin): On note $\min_\varepsilon(z) = -\varepsilon \log \sum_i e^{-\frac{z_i}{\varepsilon}}$. C’est une approximation asymptotique du min, en effet $\min_\varepsilon(z) \xrightarrow{\varepsilon \rightarrow 0^+} \min(z)$

Preuve (convergence): On note $\bar{z} = \min(z)$, $\min_\varepsilon(z - \bar{z}) = \min_\varepsilon(z) - \bar{z}$ donc $\min_\varepsilon(z) = \bar{z} + \min_\varepsilon(z - \bar{z})$. Sans perte de généralité, supposons que le minimum est z_1 . Alors

$$\begin{aligned}\min_\varepsilon(z - \bar{z}) &= -\varepsilon \log \sum_i e^{-\frac{z_i - \bar{z}}{\varepsilon}} = -\varepsilon \log \left(1 + \underbrace{\sum_{i \neq 1} e^{-\frac{z_i - \bar{z}}{\varepsilon}}}_{\xrightarrow{\varepsilon \rightarrow 0} 0} \right) \\ &= -\varepsilon \left[\sum_{i \neq 1} e^{-\frac{z_i - \bar{z}}{\varepsilon}} + o\left(\sum_{i \neq 1} e^{-\frac{z_i - \bar{z}}{\varepsilon}} \right) \right] \xrightarrow{\varepsilon \rightarrow 0} 0\end{aligned}$$

d’où le résultat. □

Remarque 5.2.1: Dans nos calculs numériques, pour éviter les underflows et assurer la stabilité, on calculera notre softmin $\min_\varepsilon(z) = \bar{z} + \min_\varepsilon(z - \bar{z})$ car numpy à le bon goût de bien calculer l'exponentielle pour des valeurs petites et la stabilisation du softmin réalisé permet de rester dans ce cas de figure, on appelle cela une stabilisation *log-sum-exp*.

$$\begin{aligned}
\varphi_i^{(l+1)} &= \varepsilon \log(a_i) - \varepsilon \log\left(K e^{\frac{\psi^{(l)}}{\varepsilon}}\right)_i \\
&= \varepsilon \log(a_i) - \varepsilon \log\left(\sum_j K_{i,j} \exp\left(\frac{\psi_j^{(l)}}{\varepsilon}\right)\right) \\
&= \varepsilon \log(a_i) - \varepsilon \log\left(\sum_j \exp\left(-\frac{C_{i,j} - \psi_j^{(l)}}{\varepsilon}\right)\right) \\
&= \varepsilon \log(a_i) + \underbrace{\min_\varepsilon\left(C_{i,j} - \psi_j^{(l)}\right)}_j \\
&\quad \text{minimum sur toutes les colonnes de la matrice } (C - \mathbb{1}_n g^{(l)T})
\end{aligned}$$

Remarque 5.2.2: On note $\min_\varepsilon^{\text{row}}$ et $\min_\varepsilon^{\text{col}}$ respectivement les softmins sur les lignes et colonnes d'une matrice. On obtient donc finalement les itérations suivantes:

$$\begin{aligned}
\varphi^{(l+1)} &= \min_\varepsilon^{\text{row}}\left(C - \mathbb{1}_n \psi^{(l)T}\right) + \varepsilon \log(a) \\
\psi^{(l+1)} &= \min_\varepsilon^{\text{col}}\left(C - \varphi^{(l+1)} \mathbb{1}_n^T\right) + \varepsilon \log(b)
\end{aligned}$$

6 BARYCENTRES ET INTERPOLATIONS

6.1 Formulation à deux plans

On présentera ici le cas du calcul d'un barycentre entre deux densités par rapport à la mesure de Lebesgue. Un cas général est présenté dans (Benamou *et al.*, 2014) et (Nenna, 2016)

Soit deux mesures μ_0 et μ_1 de Σ_n , $t \in [0, 1]$. On cherche $\rho_t = \min\{(1-t)L_c(\mu_0, \pi) + tL_c(\mu_1, \pi), \pi \in \Sigma_n\}$. Ce problème est très compliqué à résoudre, on décide donc de s'occuper de sa version régularisée et de résoudre $\rho_t^\varepsilon = \min\{(1-t)L_c^\varepsilon(\mu_0, \pi) + tL_c^\varepsilon(\mu_1, \pi), \pi \in \Sigma_n\}$ qui converge quand $\varepsilon \rightarrow 0^+$ vers ρ_t .

Ce problème se réécrit, cf (Nenna, 2016) avec $\bar{\pi} = (\pi_0, \pi_1) \in \mathcal{P}(\mathbb{R}^2)^2$

$$\min_{\bar{\pi}} \left\{ (1-t)\text{KL}(\pi_0|K_\varepsilon) + t\text{KL}(\pi_1|K_\varepsilon) : \underbrace{\pi_0 \mathbb{1}_n = \mu_0}_{\pi_0^x = \mu_0}, \underbrace{\pi_1 \mathbb{1}_n = \mu_1}_{\pi_1^x = \mu_1}, \underbrace{\pi_0^T \mathbb{1}_n = \pi_1^T \mathbb{1}_n}_{\pi_0^y = \pi_1^y} \right\} \quad (\mathcal{B}_\varepsilon)$$

Et se résout avec un algorithme très similaire à celui de Sinkhorn obtenu grâce à l'algorithme de projections alternées de Bregman et à la forme particulière des matrices de couplage, cf Chapitre 5 et (Benamou *et al.*, 2014).

$\bar{\pi}^{(n)} = (\pi_0^{(n)}, \pi_1^{(n)})$ ou

$$\begin{aligned}
\pi_0^{(n)} &= \text{diag}\left(u_0^{(n)}\right) K_\varepsilon \text{diag}\left(v_0^{(n)}\right) \\
\pi_1^{(n)} &= \text{diag}\left(u_1^{(n)}\right) K_\varepsilon \text{diag}\left(v_1^{(n)}\right)
\end{aligned}$$

On itère donc la manière suivante: $v_0^{(0)} = \mathbb{1}_n = v_1^{(0)}$ puis

$$u_0^{(n)} = \frac{\nu^{(n)}}{K_\varepsilon v_0^{(n)}}, u_1^{(n)} = \frac{\nu^{(n)}}{K_\varepsilon v_1^{(n)}}$$

$$v_0^{(n+1)} = \frac{\mu_0}{K_\varepsilon^T u_0^{(n)}}, v_1^{(n+1)} = \frac{\mu_1}{K_\varepsilon^T u_1^{(n)}}$$

ou à chaque itération

$$\nu^{(n)} = \left(u_0^{(n)} \odot \left(K_\varepsilon v_0^{(n)} \right) \right)^{1-t} \odot \left(u_1^{(n)} \odot \left(K_\varepsilon v_1^{(n)} \right) \right)^t$$

est l'estimation du barycentre, toutes les opérations sont à considérer éléments par éléments.

6.2 Formulation à trois plans

Une autre formulation du problème est possible.

On prend i,j,k des tuples, c'est à dire qu'on peut avoir par exemple $i = (i_1, i_2, i_3)$, $P = P(dx, dy, dz) = P_{i,j,k}$, $C^\lambda(x, y, z) = (1 - \lambda)c(x, y) + \lambda c(y, z)$, $K^\lambda = e^{-\frac{C^\lambda}{\varepsilon}}$

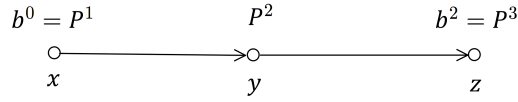


Fig. 10. – Schéma représentant l'intuition de la formulation avec 3 plans. On ajoute un 3ème plan de sorte que la condition sur le barycentre soit implicite via la matrice de coût pour pouvoir utiliser un algorithme de Sinkhorn plus simplement. $C^\lambda(x, y, z)$ correspond à la distance de x à z en passant par y coefficientés par $1 - \lambda, \lambda$. Au lieu de devoir « résoudre deux problèmes » on ne s'occupe que d'un seul légèrement plus complexe

On cherche donc

$$\min_P \{ \langle C^\lambda, P \rangle + \varepsilon H(P \mid dx dy dz) : P^1 = b^0 \text{ et } P^3 = b^2 \} \quad (\mathcal{B}_\varepsilon^{3P})$$

Le lagrangien de la fonction objectif est

$$\begin{aligned}
 \mathcal{L}(P, f, g) &= \langle C^\lambda, P \rangle + \varepsilon H(P) - \langle f, P^1 - b^0 \rangle - \langle g, P^3 - b^2 \rangle \\
 &= \sum_{i,j,k} C_{i,j,k}^\lambda P_{i,j,k} + \varepsilon \sum_{i,j,k} P_{i,j,k} (\log(P_{i,j,k}) - 1) \\
 &\quad - \sum_i f_i \sum_{j,k} (P_{i,j,k} - b_i^0) - \sum_k g_k \sum_{i,j} (P_{i,j,k} - b_k^2)
 \end{aligned}$$

Les conditions d'optimalité du premier ordre nous donnent:

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial P_{i,j,k}} = 0 &\Leftrightarrow C_{i,j,k}^\lambda + \varepsilon \log(P_{i,j,k}) - f_i - g_k = 0 \\
 &\Leftrightarrow P_{i,j,k} = e^{\frac{f_i}{\varepsilon}} e^{-\frac{C_{i,j,k}}{\varepsilon}} e^{\frac{g_k}{\varepsilon}} = u_i K_{i,j,k} v_k
 \end{aligned}$$

En injectant ce résultat dans les conditions $P^1 = b^0$ et $P^3 = b^2$ on obtient:

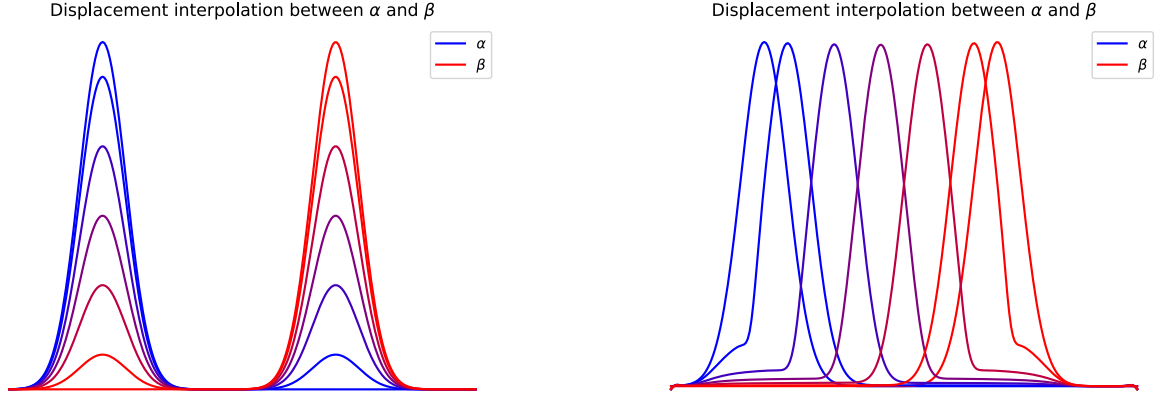


Fig. 11. – Comparaison avec une autre approche possible. A gauche : interpolation obtenue en effectuant des simples pondérations. A droite: interpolation obtenue grâce aux barycentres de Wasserstein, on observe bien ici la différence entre un déplacement vertical obtenu naïvement avec des pondérations et un déplacement horizontal.

$$\begin{aligned}
\forall i, \sum_{j,k} P_{i,j,k} = b_i^0 &\Leftrightarrow \sum_{j,k} u_i K_{i,j,k} v_k = b_i^0 \\
&\Leftrightarrow u_i \sum_k \left(\sum_j K_{i,j,k} \right) v_k = b_i^0 \\
&\Leftrightarrow u_i \underbrace{\sum_k K_{i,k}^{1,3}}_{(K^{1,3}v)_i} v_k = b_i^0 \\
&\Leftrightarrow u \odot (K^{1,3}v) = b^0 \\
&\Leftrightarrow u = \frac{b^0}{K^{1,3}v}
\end{aligned}$$

et de manière analogue $v = \frac{b^2}{(K^{1,3})^T u}$ ce qui nous permet d'itérer comme dans Chapitre 5 pour venir récupérer P^2 qui correspond au barycentre.

Remarque 6.2.1: Dans un souci d'optimisation, on ne calculera pas les tenseurs C et K . En effet il est possible de se passer numériquement de ces calculs pour gagner en rapidité en calculant uniquement $K^{1,3}$ via la matrice de coût et deux tenseurs intermédiaires.

$$K_{i,k}^{1,3} = \sum_j e^{-\frac{(1-\lambda)c(i,j)}{\varepsilon}} e^{-\frac{\lambda c(j,k)}{\varepsilon}} = \sum_j K_{i,j}^{1-\lambda} K_{j,k}^\lambda \Rightarrow K^{1,3} = K^{1-\lambda} K^\lambda$$

Pour récupérer ν on réutilise les deux matrices ayant servies dans le calcul de $K^{1,3}$

$$\begin{aligned}
\nu_j = P_j^2 &= \sum_{i,k} P_{i,j,k} = \sum_{i,k} u_i K_{i,j,k} v_k = \sum_{i,k} u_i K_{i,j}^{1-\lambda} K_{j,k}^\lambda v_k = \sum_i u_i K_{i,j}^{1-\lambda} \sum_k K_{j,k}^\lambda v_k \\
&\Rightarrow \nu = \left[(K^{1-\lambda})^T u \right] \odot [K^\lambda v]
\end{aligned}$$

Cette manière de calculer les barycentres est en moyenne 3 fois plus rapide que celle présentée au début du chapitre pour des mesures en 1D.

Remarque 6.2.2: Pour les calculs de barycentres en 2D on utilisera plutôt la première formulation car elle permet une implémentation sans tenseurs et uniquement avec des matrices pour les multiplications par le noyau K_ε . En effet, dans ce cas (i et j sont des doubles indices) $K_{i,j} = e^{-\|x_i - y_j\|^2 / \varepsilon}$ peut être séparé, et appliquer le noyau revient à appliquer un noyau intermédiaire deux fois sous la forme d'une matrice $K_1 = e^{(x_{i_1} - y_{j_1})^2 / \varepsilon}$ par multiplication sur

les lignes puis les colonnes, ie $K(A) = K_1AK_1$. Cette approche allège considérablement les besoins en RAM et calculs.

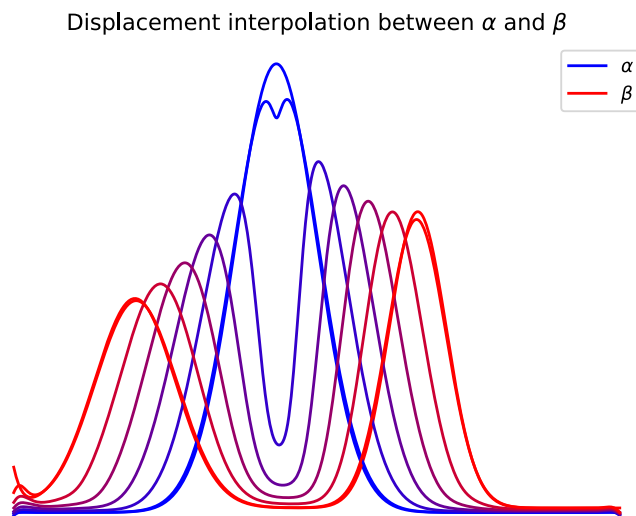


Fig. 12. – Exemple d’interpolation obtenue entre des courbes gaussiennes (bleu vers rouge) avec $N = 256$

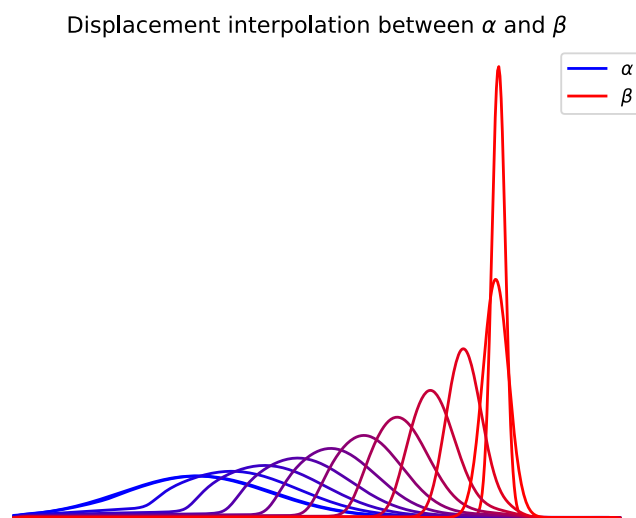


Fig. 13. – Exemple d’interpolation obtenue entre des courbes gaussiennes (bleu vers rouge) avec $N = 256$
 Les images 2D sont représentées par des mesures uniformes sur un carré de taille 500×500 .



Fig. 14. – Interpolation entre une croix et deux disques



Fig. 15. – Interpolation entre une spirale et un chat



Fig. 16. – Interpolation entre un trèfle et un coq



Fig. 17. – Interpolation entre une croix et un triskel



Fig. 18. – Interpolation entre un chat et un triskel

On peut comparer un des résultats obtenu par rapport à celui d'une simple moyenne pondérée pour bien apprécier les résultats obtenus.

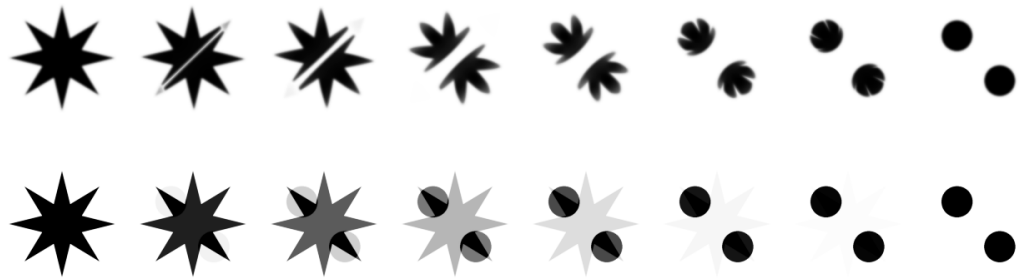


Fig. 19. – Comparaison avec une autre approche possible. En haut: interpolation obtenue grâce aux barycentres de Wasserstein. En bas : interpolation obtenue en effectuant des simples pondérations.

CONCLUSION

L'objectif de ce stage était de se familiariser avec la théorie du TO et d'implémenter des algorithmes de résolutions numériques. Ce sont choses faites. J'ai pu travailler sur la théorie du TO par la lecture d'ouvrages et articles scientifiques. En particulier sur le problème relaxé de Kantorovich, sa régularisation et la distance qu'il induit. Pour l'aspect numérique j'ai pu travailler sur l'algorithme de Sinkhorn pour le calcul de TO et sa stabilisation grâce au domaine log. J'ai aussi pu travailler sur les barycentres de Wasserstein pour réaliser des interpolations, avec deux méthodes différentes dont une qui bien que non inédite non trouvable dans des articles.

En conclusion du travail effectué, je pense qu'il est important de noter la différence entre une théorie qui marche parfaitement et les implémentations que l'on en fait. En effet bien que l'algorithme de Sinkhorn converge quand $\varepsilon \rightarrow 0$ en pratique, ce n'est pas le cas car le 0 machine est atteint beaucoup trop vite et on observe l'apparition d'overflows et d'underflows. D'où l'intérêt de la stabilisation à l'aide du logarithme. Il est aussi important de noter l'importance de coder de manière efficace, en limitant au maximum l'utilisation de boucles pour privilégier les calculs vectorisés qui sont beaucoup plus rapides.

Il reste cependant certains points pouvant être améliorés ou restant à faire, j'aurais peut-être pu optimiser d'autant plus mon code, bien qu'il est déjà assez rapide. Il aurait aussi été intéressant d'implémenter les calculs de barycentres sous domaine logarithmique, mais c'est plus délicat que pour le calcul de TO et mes algorithmes fonctionnaient déjà très bien sans.

Ce travail pourra possiblement être utilisé par d'autres étudiants du même niveau (L3) comme support ou bien à poursuivre. Il y a certaines perspectives que je n'ai pas trop eu l'occasion d'explorer. En particulier le lien avec les EDP qui est le principal axe de recherche dans le domaine avec les applications en machine learning. Une poursuite de mon travail pourrait s'appuyer plus sur ce lien, à condition d'avoir déjà travaillé avec des EDP.

En terme de bilan personnel et d'apports du stage, ce stage a été une expérience passionnante et enrichissante pour une simple découverte du monde de la recherche. J'ai eu l'opportunité de rencontrer de nombreuses personnes toutes aussi passionnées les unes que les autres. De plus, le sujet traité, non étudié à l'INSA et dont les applications sont nombreuses, pourra être un atout sur un CV.

Être initié à la recherche m'a permis de continuer à développer un certain état d'esprit analytique et critique, demandant une réflexion approfondie. Cette expérience a non seulement enrichi mes compétences techniques mais aussi amélioré ma capacité à aborder les problèmes et surtout la difficulté.

Dans le futur, après cette expérience, j'aimerais aussi tenter également le secteur industriel. Réaliser un stage en industrie me permettrait de comparer les deux milieux et mieux choisir la direction vers laquelle m'orienter. Cette dualité entre recherche académique et industrie pourrait me guider et m'aider à faire des choix pour mon avenir.

ANNEXES

Le code pour la partie sur Sinkhorn est disponible [ici \(nbviewer\)](#).

Le code pour la partie sur les barycentres est disponible [ici \(nbviewer\)](#).

BIBLIOGRAPHIE

Benamou, J.-D. *et al.* (2014) *Iterative Bregman Projections for Regularized Transportation Problems*. Disponible sur: <https://arxiv.org/abs/1412.5154>

Nenna, L. (2016) *Numerical Methods for Multi-Marginal Optimal Transportation*. Disponible sur: <http://www.theses.fr/2016PSLED017/document>

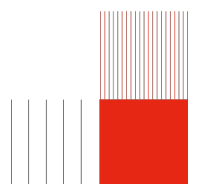
Peyré, G. et Cuturi, M. (2020) *Computational Optimal Transport*. Disponible sur: <https://arxiv.org/abs/1803.00567>

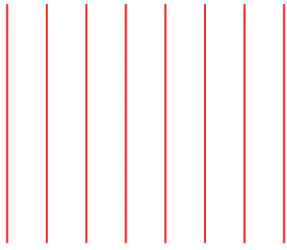
Villani, C. (2003) « Topics in Optimal Transportation Theory », 58, p. . Disponible sur: <https://doi.org/10.1090/gsm/058>

GLOSSAIRE

TO Transport Optimal

IST Instituto Superior Técnico



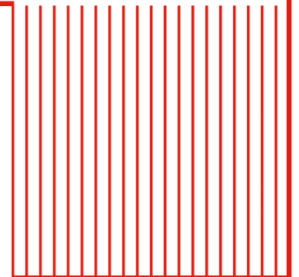


La théorie du TO (Transport Optimal) est premièrement introduite par Monge dans un mémoire en 1781. L'idée est la suivante; un ouvrier souhaite déplacer un tas de sable afin d'en construire un château à un autre endroit. L'ouvrier essaye naturellement de minimiser son effort, que l'on pourra quantifier de différentes manières (par exemple la distance parcourue). Mathématiquement le problème est en fait celui de la comparaison de deux mesures de probabilités, deux tas de sable de même masse. On cherche alors les moyens optimaux de transporter chaque élément de masse élémentaire de la première mesure vers la deuxième en fonction d'un certain coût. On s'intéresse alors aux propriétés mathématiques de ces transports et à des algorithmes qui permettent de résoudre ces problèmes efficacement. Le coût minimal définit alors une distance entre les mesures de probabilités contenant de l'information sur l'espace sous-jacent, permettant par exemple de réaliser des interpolations.

Ce rapport aura comme vocation de faire état du travail que j'ai réalisé durant mon stage en insistant sur les bases théoriques et fondations du TO. D'abord avec la notion de mesure, le problème de Monge en discret et sa généralisation continue grâce aux push-forward. Ce schéma, discret puis généralisation sera suivi pour la relaxation du problème avec le problème de Kantorovich que nous étudierons en détail sans oublier de parler de la distance qu'il induit sur l'espace métrique des mesures de probabilités. On parlera aussi entre autres du problème dual associé, de la régularisation entropique, de résolution numérique grâce à l'algorithme de Sinkhorn (projections alternées) basique ou en domaine logarithmique et enfin de barycentres et interpolations. Des exemples de résolutions numériques seront intégrés tout au long du rapport pour illustrer les concepts théoriques, montrer l'efficacité des algorithmes implémentés et on essayera dans la mesure du possible de démontrer un maximum de résultats.

The theory of Optimal Transport (OT) was first introduced by Monge in a memoir in 1781. The idea is the following: a worker wishes to move a pile of sand in such a way that he constructs a sandcastle at another location. Naturally, the worker tries to minimize his effort, which can be quantified in various ways (for example, the distance traveled). Mathematically, the problem is actually about comparing two probability measures, two piles of sand of the same mass. We look for the optimal ways to transport each elementary mass element from the first measure to the second according to a certain cost. We are interested by the mathematical properties of these transports and algorithms used to solve these problems efficiently. The minimal cost then defines a distance between the probability measures, containing information about the underlying space, allowing for example to do interpolations.

This report aims to present the work I did during my internship, emphasizing on the theoretical foundations and basics of Optimal Transport (OT), starting with the concept of measure, the Monge problem in discrete form, and its generalization via push-forwards. This pattern of discrete analysis followed by generalization will be applied to the relaxation of the problem, focusing on the Kantorovich problem, which we will study in detail, including the distance it induces on the probability metric space. We will also discuss the associated dual problem, entropic regularization, numerical resolution using the Sinkhorn algorithm (alternate projections) basic or in the logarithmic domain and finally barycenters and interpolations. Numerical resolution examples will be integrated throughout the report to illustrate the theoretical concepts, demonstrate the effectiveness of the implemented algorithms and we will try when possible to prove results.



INSA Rennes

20 avenue des Buttes de Coësmes
CS 70839

35708 Rennes cedex 7

Tél : + 33 (0)2 23 23 82 00

www.insa-rennes.fr



INSA | INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
RENNES


**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**
*Liberté
Égalité
Fraternité*