

RAPPORT DE PROJET RECHERCHE

LE PROBLÈME DU TRANSPORT OPTIMAL : FORMULATION DE TRANSPORT OPTIMAL DE LA THÉORIE DE LA FONCTIONNELLE DE LA DENSITÉ ET DE L'ÉQUILIBRE DE COURNOT-NASH

Amine Souiri

Encadrant : M.Sergio PULIDO

19 mars 2021 - 31 mai 2021

Abstract

The first paper describes that one of the challenging scenarios for Kohn-Sham density-functional theory (DFT) in quantum mechanics, that is, when the repulsion between the electrons largely dominates over their kinetic energies dominates, can be reformulated as an optimal transport problem [1], a well-established field of mathematics and economics.

The second article focuses on the Cournot-Nash equilibrium, which is a mathematical notion specific to game theory. It is useful in economics, politics and other fields to model the behavior of individuals, firms or other competing organizations whose goal is to maximize their profit. In particular, it allows to characterize equilibrium situations in which the strategy adopted by each type of agent is optimal.

In this document, we first give the first definitions and study the elementary problems in optimal transport theory to move on the papers study. Then, we reformulate the strong-interaction limit of DFT and the Cournot-Nash equilibrium according to the optimal transport theory. Thus leads us to demonstrate some results about these optimal transport problems.

.....

Résumé

Le premier article décrit que l'un des scénarios les plus complexes de la théorie de la fonctionnelle de la densité (DFT) de Kohn-Sham en mécanique quantique, qui correspond à la situation où la répulsion entre les électrons est largement dominante face à leur énergies cinétiques, peut être reformulé en un problème de transport optimal, un domaine bien établi des mathématiques et de l'économie.

Le second article s'intéresse à l'équilibre de Cournot-Nash qui est une notion mathématique propre à la théorie des jeux. Elle est utile en économie, en politique ou dans d'autres domaines pour modéliser les comportements d'individus, d'entreprises ou d'autres organismes en concurrence qui ont pour but de maximiser leur profit. En particulier, il permet de caractériser des situations d'équilibre dans lesquelles la stratégie adoptée par chaque type d'agent est optimale.

Dans ce document, on donne d'abord les premières définitions et on étudie les problèmes élémentaires dans la théorie du transport optimal [1], pour ensuite passer à l'étude des articles. On reformule ensuite la limite de l'interaction forte en DFT et l'équilibre de Cournot-Nash selon la théorie du transport optimal. Ceci nous conduit ainsi à démontrer certains résultats sur ces problèmes de transport optimal.

Table des matières

1	Introduction à la théorie du transport optimal	2
1.1	Introduction	2
1.2	Le Problème de Monge	3
1.3	Le Problème de Kantorovitch	3
1.4	Dualité	4
1.5	Le Problème de transport multi-marginales	5
2	Premières transports	6
2.1	Le problème de Monge dans le cas discret	6
2.2	Le problème de Kantorovich dans le cas discret	7
2.3	Implémentations numériques dans le cas discret pour $n < m$	8
2.4	Implémentations numériques dans le cas discret pour $n = m$	9
3	La théorie de la fonctionnelle de la densité	10
3.1	Introduction	10
3.2	Interactions forte dans DFT	11
3.3	Résultats physiques	12
3.4	L'apport du transport optimal	15
3.5	Reformulation	15
3.6	Théorèmes énergétiques	16
4	L'équilibre de Cournot-Nash	21
4.1	Le modèle d'équilibre	21
4.2	Recherche d'équilibre	23
4.3	Calculs numérique	23
4.4	Un algorithme de calculs numériques	25
4.5	Équilibre de Cournot-Nash dans le cas fini	26
4.6	Une approche variationnelle des équilibres de Cournot-Nash	27
4.7	Approche numérique	28

1 Introduction à la théorie du transport optimal

1.1 Introduction

On dispose d'une dune de sable et d'un fossé de même volume. On souhaite déplacer tout le sable de la dune vers le fossé afin de le remplir. Cependant, le déplacement de chaque grain de sable (ou chaque portion de la dune selon que l'on considère le problème comme discret ou continu) admet un coût qui dépend de l'emplacement initial du grain et de sa position finale.

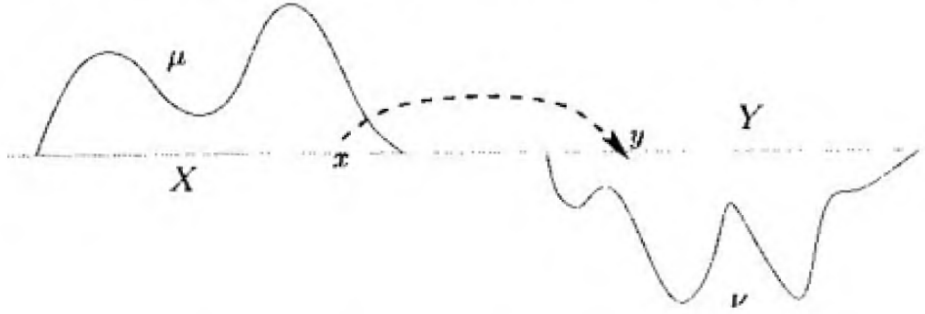


FIGURE 1 – Transport de la dune X distribuée selon μ vers le fossé Y distribué selon ν

L'objectif du problème du transport optimal est de trouver le moyen le moins coûteux de déplacer l'ensemble du sable de la dune dans le fossé.

Définition 1. Soit X un espace métrique complet séparable. On note $\mathcal{C}(X)$ l'espace des fonctions continues sur X , $\mathcal{C}_0(X)$ l'espace des fonctions continues sur X qui tendent vers 0 à l'infini, $\mathcal{C}_b(X)$ l'espace des fonctions continues bornées sur X . On note $\mathcal{M}(X)$ l'espace des mesures boréliennes finies sur X , et :

$$\begin{aligned}\mathcal{M}^+(X) &= \{\mu \in \mathcal{M}(X), \mu \geq 0\} \\ \mathcal{P}(X) &= \{\mu \in \mathcal{M}^+(X), \mu(X) = 1\}\end{aligned}$$

Définition 2. Une mesure $\mu \in \mathcal{M}(X)$ est dite sans atomes si :

$$\forall x \in X, \mu(\{x\}) = 0$$

Définition 3. Soit X, Y des espaces métriques, $\mu \in \mathcal{M}(X)$ et $T : X \rightarrow Y$ mesurable. La mesure image de μ par T est la mesure $T_{\#}\mu$ sur Y telle que pour tout $B \subseteq Y$:

$$T_{\#}\mu(B) = \mu(T^{-1}(B))$$

Remarque. Si on considère l'exemple de la dune de sable, X représente la surface sur laquelle est disposé le tas de sable et Y la surface au dessous de laquelle s'étend le fossé. μ et ν sont des mesures de volume : Si $A \subset X$, alors $\mu(A)$ représente le volume de sable au dessus de la surface A . Par convention, on normalise le volume de la dune de sable à 1.

On définit dans les sous-sections suivantes des problèmes essentiels dans la théorie du transport optimal.

1.2 Le Problème de Monge

Définition 4. Soit X, Y des espaces métriques, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ deux mesures de probabilités et $T : X \rightarrow Y$ mesurable. On dit que T est une fonction de transport entre μ et ν si et seulement si $T_{\#}\mu = \nu$.

Définition 5. Soit X, Y deux espaces métriques, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ deux mesures de probabilités et une fonction de coût $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$. Le problème de Monge [2] est le problème d'optimisation suivant :

$$(MP) := \inf \left\{ \int_X c(x, T(x)) d\mu(x), T : X \rightarrow Y \text{ et } T_{\#}\mu = \nu \right\}$$

1.3 Le Problème de Kantorovitch

On présente ici une relaxation du problème de Monge énoncée par Kantorovitch et on établit une égalité [3] entre les infimums de ces deux problèmes.

Définition 6. Les mesures marginales d'une mesure γ sur l'espace produit $X \times Y$ sont les mesures $\pi_X \# \gamma$ et $\pi_Y \# \gamma$ où $\pi_X : X \times Y \rightarrow X$ et $\pi_Y : X \times Y \rightarrow Y$ sont les projections telles que pour tout $(x, y) \in X \times Y$, $\pi_X(x, y) = x$ et $\pi_Y(x, y) = y$

Définition 7. Soit X, Y deux espaces métriques, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ deux mesures de probabilités. On appelle plan de transport entre μ et ν toute mesure de probabilité γ sur l'espace produit $X \times Y$ dont les mesures marginales sont μ et ν . De plus, on note l'espace des plans de transport noté $\Pi(\mu, \nu)$ défini par :

$$\Pi(\mu, \nu) = \{\gamma \in \mathcal{P}(X \times Y) | \pi_X \# \gamma = \mu, \pi_Y \# \gamma = \nu\}$$

Définition 8. Soit X, Y deux espaces métriques, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ deux mesures de probabilités et une fonction de coût $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$. Le problème de Kantorovitch [4] est le problème d'optimisation suivant :

$$(KP) := \inf \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y), \gamma \in \Pi(\mu, \nu) \right\}$$

Remarque. Résoudre le problème de Monge est nettement plus difficile que de résoudre le problème de Kantorovitch. Par exemple, il peut ne pas exister de fonction de transport T satisfaisant les deux contraintes du problème de Monge. Dans le problème de Kantorovitch, la mesure $\mu \otimes \nu$ remplit toujours la condition $\mu \otimes \nu \in \Pi(\mu, \nu)$.

De plus, dans le problème de Kantorovitch, nous cherchons une solution sur l'ensemble convexe $\Pi(\mu, \nu)$, alors que dans le problème de Monge, l'ensemble des $\mu, \nu \in \mathcal{P}(X)$ ne satisfait pas la condition $T_{\#}\mu = \nu$, et T n'est pas nécessairement convexe.

Remarque. Si $\gamma = (\text{id}, T)_{\#}\mu$, où $T : X \rightarrow Y$ est une fonction mesurable, and γ est solution au problème de Kantorovitch, alors T est également solution au problème de Monge. Si γ prend la forme précédente, on dit qu'il est *porté par un graphe*.

Théorème. (Théorème d'existence de Kantorovitch) Soit X, Y deux espaces métriques, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ deux mesures de probabilités et une fonction de coût $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$ qui est semi-continue inférieurement et bornée inférieurement. Alors il existe $\gamma \in \Pi(\mu, \nu)$ solution au problème de Kantorovitch.

Démonstration. Cf. preuve du Lemme 1 du théorème 2 pour $N = 2$ et $\mathcal{H} = X_1 \otimes X_2$, où $X_1 = X$ et $X_2 = Y$. \square

Théorème. Soit $X = Y \subset \mathbb{R}^d$ compacts. Supposons que $\mu \in \mathcal{P}(X)$ est sans atomes, $\nu \in \mathcal{P}(Y)$ et $c \in \mathcal{C}(X, Y)$. Alors on a :

$$\inf(MP) = \min(KP).$$

Définition 9. Une suite $(\mu_n)_{n \in \mathbb{N}}$ de mesures de probabilités sur X est dite tendue si :

$$\forall \epsilon > 0, \exists K \text{ compact } \subseteq X, \forall n \in \mathbb{N}, \mu_n(X \setminus K) < \epsilon,$$

Théorème (Théorème de Prokhorov). Soit X un espace métrique complet et séparable et $(\mu_n)_{n \in \mathbb{N}}$ une suite tendue de mesures de probabilités sur X . Il existe alors $\mu \in \mathcal{P}(X)$ et une sous-suite $(\mu_{n_k})_{k \in \mathbb{N}}$ telle que $\mu_{n_k} \rightharpoonup \mu$ (dans la dualité avec $\mathcal{C}_b(X)$). Inversement, toute suite $(\mu_n)_{n \in \mathbb{N}}$ qui vérifie $\mu_n \rightharpoonup \mu$ est nécessairement tendue.

Remarque. Pour des mesures marginales μ et ν , on notera si besoin $\inf(KP) = \mathcal{KP}(\mu, \nu)$ si l'on a besoin de faire apparaître.

1.4 Dualité

Afin de trouver une solution au problème de Kantorovich, on souhaite créer un problème dual et trouver une solution de ce dernier. On pose :

$$\sup_{\varphi, \psi} \left\{ \int_X \varphi d\mu + \int_Y \psi d\nu - \int_{X \times Y} (\varphi(x) + \psi(y)) d\gamma(x, y) \right\} = \begin{cases} 0 & \text{si } \gamma \in \Pi(\mu, \nu) \\ +\infty & \text{sinon} \end{cases}$$

On étudie alors l'expression :

$$\inf_{\gamma \in \mathcal{M}^+(X \times Y)} \left\{ \int_{X \times Y} c d\gamma + \sup_{\varphi, \psi} \left\{ \int_X \varphi d\mu + \int_Y \psi d\nu - \int_{X \times Y} (\varphi(x) + \psi(y)) d\gamma(x, y) \right\} \right\}$$

Par interversion de la borne sup et la borne inf, l'étude revient à l'expression suivante :

$$\sup_{\varphi, \psi} \left\{ \int_X \varphi d\mu + \int_Y \psi d\nu + \inf_{\gamma \in \mathcal{M}^+(X \times Y)} \left\{ \int_{X \times Y} (c(x, y) - \varphi(x) - \psi(y)) d\gamma(x, y) \right\} \right\}$$

, où l'on a :

$$\inf_{\gamma \in \mathcal{M}^+(X \times Y)} \left\{ \int_{X \times Y} (c - \varphi(x) - \psi(y)) d\gamma \right\} = \begin{cases} 0 & \text{si } \varphi(x) + \psi(y) \leq c(x, y) \quad \forall x, y \\ -\infty & \text{sinon} \end{cases}$$

On peut alors définir le problème dual :

Définition 10. Soit $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ et une fonction de coût $c \in \mathcal{C}(X \times Y)$. Le problème dual est le problème d'optimisation suivant :

$$(DP) := \sup \left\{ \int_X \phi d\mu + \int_Y \psi d\nu \mid \phi \in \mathcal{C}_b(X), \psi \in \mathcal{C}_b(Y), \phi \oplus \psi \leq c \right\}$$

Remarque. Les fonctions ϕ et ψ solutions au problème dual, et plus généralement, toute fonction impliquée dans la minimisation solution du problème dual est appelé potentiel de Kantorovitch.

Définition 11. Soit une fonction $\Phi : X \longrightarrow \overline{\mathbb{R}}$. On définit la c -transformée de Φ comme la fonction $\Phi^c : Y \longrightarrow \overline{\mathbb{R}}$ qui vérifie :

$$\Phi^c(y) = \inf_{x \in X} \{c(x, y) - \Phi(x)\}.$$

De même, la \bar{c} -transformée de $\Psi : Y \longrightarrow \overline{\mathbb{R}}$ est la fonction $\Psi^{\bar{c}} : X \longrightarrow \overline{\mathbb{R}}$ qui vérifie :

$$\Psi^{\bar{c}}(x) = \inf_{y \in Y} \{c(x, y) - \Psi(y)\}.$$

Une fonction φ est dite c -concave si il existe Φ telle que $\varphi = \Phi^c$ une fonction ψ est dite \bar{c} -concave si il existe Ψ telle que $\psi = \Psi^{\bar{c}}$.

Théorème (Théorème de dualité). *Soit X, Y des espaces compacts, $c \in \mathcal{C}(X, Y)$. Alors il existe une fonction φ c -concave telle que (φ, φ^c) est solution au problème dual, c'est à dire :*

$$\max(DP) = \max_{\varphi \text{ } c\text{-concave}} \left\{ \int_X \varphi d\mu + \int_Y \varphi^c d\nu \right\} \quad (1.1)$$

Remarque. Dans ce cas, on peut supposer que l'hypothèse de dualité forte est vérifiée, c'est à dire que $\min(KP) = \max(DP)$, et donc :

$$\min(KP) = \max_{\varphi \text{ } c\text{-concave}} \left\{ \int_X \varphi d\mu + \int_Y \varphi^c d\nu \right\}. \quad (1.2)$$

On pourra se référer aux premiers chapitres des livres de [5] Ambrosio et [6] Villani pour avoir une introduction exhaustive de la théorie du transport optimal et de leurs applications dans les discussions du livre de [1] Santambrogio. On démontrera un certain nombre de résultats dans la deuxième partie pour le cas du problème de transport multi-marginales présenté dans la sous-section suivante.

1.5 Le Problème de transport multi-marginales

Soit N un entier strictement supérieur à 1.

Définition 12. Les mesures marginales d'une mesure γ sur l'espace produit $\mathcal{H} = X_1 \times X_2 \times \dots \times X_N$ sont les mesures $\pi_{i\#}\gamma$ pour $i \in \{1, 2, \dots, N\}$ où $\pi_i : \mathcal{H} \rightarrow X_i$ sont les projections telles que pour tout $(x_1, x_2, \dots, x_N) \in \mathcal{H}$, $\pi_i(x_1, x_2, \dots, x_N) = x_i$

Définition 13. Soit X_1, X_2, \dots, X_N des espaces métriques et l'espace produit $\mathcal{H} = X_1 \times X_2 \times \dots \times X_N$ et $i \in \{1, 2, \dots, N\}$. Soit $\mu_i \in \mathcal{P}(X_i)$ des mesures de probabilités. On appelle plan de transport du N -uplet de mesure $\mu = (\mu_i)_{i \in \{1, 2, \dots, N\}}$ toute mesure de probabilité γ sur l'espace produit \mathcal{H} dont les mesures marginales sont les μ_i . De plus, on note l'espace des plans de transport noté $\Pi(\mu)$ défini par :

$$\Pi(\mu) = \{\gamma \in \mathcal{P}(\mathcal{H}) | \forall i \in \{1, 2, \dots, N\}, \pi_{i\#}\gamma = \mu_i\}$$

Définition 14. Soit X_1, X_2, \dots, X_N des espaces métriques et l'espace produit $\mathcal{H} = X_1 \times X_2 \times \dots \times X_N$ et $i \in \{1, 2, \dots, N\}$. Soit $\mu_i \in \mathcal{P}(X_i)$ des mesures de probabilités. et une fonction de coût $c : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Le problème de transport multi-marginales est le problème d'optimisation suivant :

$$(MMP) := \inf \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) d\gamma(x_1, x_2, \dots, x_N), \gamma \in \Pi(\mu) \right\}$$

Définition 15. Soit $\mu_i \in \mathcal{P}(X_i)$ des mesures de probabilités et une fonction de coût $c : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Le problème dual version multi-marginles est le problème d'optimisation suivant :

$$(DMMP) := \sup \left\{ \sum_{i=1}^N \int_{X_i} \phi_i d\mu_i \mid \phi_i : X_i \rightarrow \mathbb{R}, \phi_i \in \mathcal{C}_b(X_i), \sum_{i=1}^N \phi_i(x_i) \leq c(x_1, x_2, \dots, x_N) \right\}$$

Remarque. On notera que le problème (MP) version multi-marginales n'est pas présenté ici car c'est tout l'enjeu de la sous-section 3.4 dans le cas $X = X_1 = X_2 = \dots = X_N$.

2 Premières transports

On fixe $n, m \in \mathbb{N}$

2.1 Le problème de Monge dans le cas discret

Définition 16. Soit $X = \{x_1, \dots, x_n\} \in \mathbb{R}^n, Y = \{y_1, \dots, y_m\} \in \mathbb{R}^m$.

Soient $a_1, \dots, a_n \in \mathbb{R}_+^*, b_1, \dots, b_m \in \mathbb{R}_+^*$ tels que :

$$\mu = \sum_{i=1}^n a_i \delta_{x_i} \quad , \quad \nu = \sum_{j=1}^m b_j \delta_{y_j}$$

où δ_x est la mesure de dirac au point $x \in \mathbb{R}^d$.

Exemple :

On peut imaginer la situation suivante. En France, il y a n centres de stockage de vaccins et m vaccinodromes. Chaque centre de stockage x_i dispose de a_i vaccins et chaque vaccinodrome y_j a besoin de b_j doses qui doivent toutes lui être fournies par un même centre. On peut imaginer que le coût c représente le coût de transport entre un centre de stockage et un vaccinodrome.

Le problème de Monge revient à trouver une application $T : \{x_1, \dots, x_n\} \rightarrow \{y_1, \dots, y_m\}$ qui vérifie $T_{\#}\mu = \nu$ soit ici :

$$\forall j \in \{1, \dots, m\} \quad , \quad b_j = \sum_{i, T(x_i)=y_j} a_i$$

et telle que la quantité :

$$\sum_{i=1}^n a_i c(x_i, T(x_i))$$

soit minimale.

Remarque. Comme T associe à chaque x_i un unique y_j , on a obligatoirement $n \geq m$. Mais dans le cas d'égalité $n = m$ et $\forall i, j \in \llbracket 1; n \rrbracket, a_i = b_j = \frac{1}{n}$ de par la condition d'égalité des n termes et la condition de masse totale égale à 1. Dans ce cas, T est une permutation de $\{1, \dots, n\}$. On peut associer à la fonction de coût c la matrice de coût $(C_{i,j})_{i,j \in \{1, \dots, n\}}$ et il s'agit de trouver :

$$\min_{\sigma \in S_n} \sum_{i=1}^n C_{i, \sigma(i)}$$

2.2 Le problème de Kantorovich dans le cas discret

Généralement, on cherche une matrice $P \in M_{n,m}(\mathbb{R}^+)$ où $P_{i,j}$ représente la quantité de matière allant de x_i à y_j , P doit donc vérifier :

$$\forall i \in \llbracket 1; n \rrbracket, \sum_{j=1}^m P_{i,j} = a_i \quad \text{et} \quad \forall j \in \llbracket 1; m \rrbracket, \sum_{i=1}^n P_{i,j} = b_j$$

Dans ce cas :

$$P \in \mathcal{U}(a, b) = \{P \in M_{n,m}(\mathbb{R}^+) , \forall i \in \llbracket 1; n \rrbracket, \sum_{j=1}^m P_{i,j} = a_i \quad \text{et} \quad \forall j \in \llbracket 1; m \rrbracket, \sum_{i=1}^n P_{i,j} = b_j\}$$

Remarque. Ces 2 conditions signifient respectivement que la somme des colonnes de P et la somme des lignes de P sont respectivement les vecteurs $a = (a_1, \dots, a_n)$ et $b = (b_1, \dots, b_m)$. Ces 2 conditions peuvent donc se réécrire :

$$P^T 1^n = a \quad \text{et} \quad P 1^m = b$$

, où $1^n \in \mathbb{R}^n$ et $1^m \in \mathbb{R}^m$ sont des vecteurs uniquement composés de 1. Cette réécriture sera utile dans la prochaine sous-section.

Notons $C \in M_{n,m}(\mathbb{R}^+)$ la matrice telle que :

$$\forall i, j \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket, C_{i,j} = c(x_i, y_j)$$

Le problème de Kantorovich dans le cas discret consiste à trouver une matrice P qui minimise le produit scalaire :

$$\mathcal{T}_C(a, b) = \min_{P \in \mathcal{U}(a, b)} \langle P, C \rangle \quad \text{où} \quad \langle P, C \rangle = \sum_{i,j} P_{i,j} C_{i,j}$$

Remarque : Une propriété intéressante du problème du transport optimal est qu'il permet de définir une distance sur $\mathcal{P}(X)$.

Proposition 1. *On suppose que $X = Y$ est muni de la distance d et que $c(x, y) = d(x, y)^p$ pour $p \in \mathbb{N}^*$. Si on pose :*

$$\mathcal{W}_p : (\mu, \nu) \in \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow \mathcal{T}_{d^p}(\mu, \nu)^{\frac{1}{p}}$$

Alors $\mathcal{W}_p(\mu, \nu)$ définit une distance sur $\mathcal{P}(X)$.

Démonstration. Puisqu'on est ici dans le cas discret, on a nécessairement que $n = m$. On a pour $p \in \mathbb{N}^*$, $C = D^p$ où $D = (D)_{i,j}$ est une matrice distance qui vérifie :

- $D \in M_n(\mathbb{R}_+)$ est symétrique
- $D_{i,j} = 0 \Leftrightarrow i = j$
- $\forall i, j, k \in \llbracket 1; n \rrbracket, D_{i,k} \leq D_{i,j} + D_{j,k}$

On a :

$$\Sigma_n \rightarrow \Sigma_n, \quad a, b \rightarrow \mathcal{W}_p(a, b) = \mathcal{L}_{d^p}(a, b)$$

définit une distance sur $\mathcal{P}(X)$. □

2.3 Implémentations numériques dans le cas discret pour $n < m$

On considère μ et ν définies comme en Définition 16 en sous-section 2.1 pour $n < m$. On peut simuler X et Y sous la forme de nuages de points comme en Figure 2 où la taille de chaque point x_i est proportionnel à son poids a_i et de même avec chaque point y_j qui est proportionnel à son poids b_j .

Il est possible de résoudre ce problème de transport dans le cas discret décrit en sous-section 2.2 en utilisant le module python cvxpy qui est un module d'optimisation convexe. Ainsi on peut trouver une matrice P solution au problème de transport discret que l'on retrouve en Figure 3 où chaque coordonnée (i, j) en violet correspond au cas où $P_{i,j} = 0$ et où chaque coordonnée d'une couleur différente au cas où $P_{i,j} > 0$ de telle sorte que plus cette couleur s'éloigne du violet pour se rapprocher du jaune et plus la valeur de $P_{i,j}$ est élevée.

On trouve alors en Figure 4 les connexions entre les points des nuages définies par la matrice solution où chaque connexion entre deux points est représentée par une arête entre ces deux mêmes points.

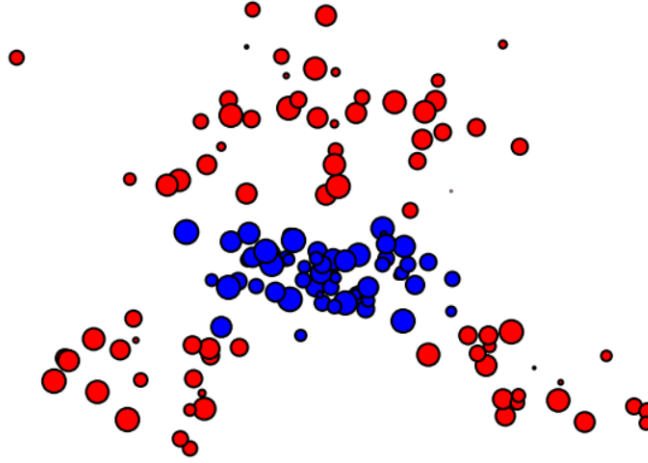


FIGURE 2 – Nuages de points X (en bleu) et Y (en rouge) pour $n = 60$ et $m = 80$

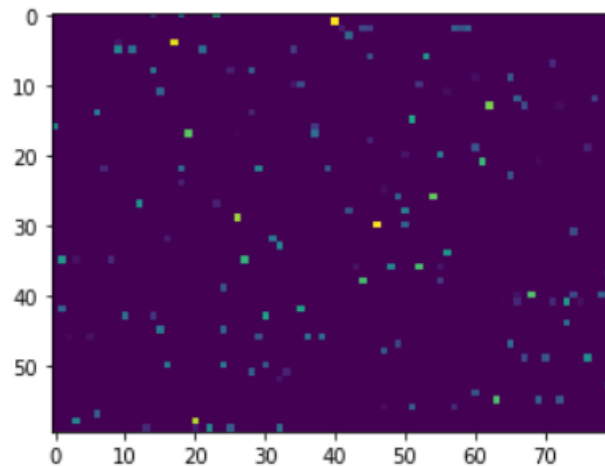


FIGURE 3 – Matrice P solution au problème de transport discret

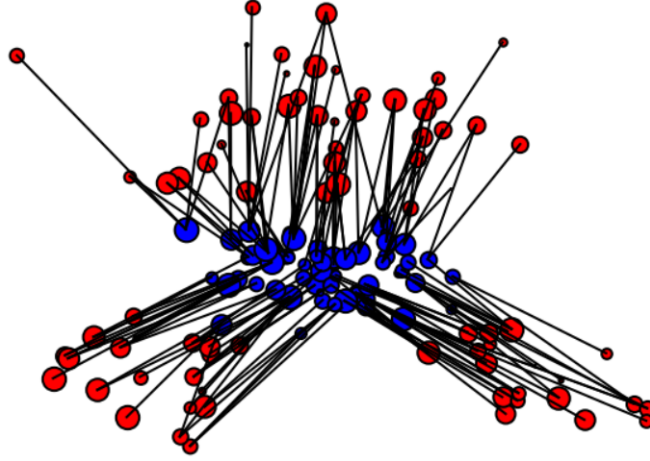


FIGURE 4 – Connexions définies entre les nuages de point par la matrice solution P

2.4 Implémentations numériques dans le cas discret pour $n = m$

De manière analogue à la sous-section précédente, on est capable d'exhiber une matrice de permutation P solution de par la remarque en en sous-section 2.2 exhibant une certaine permutation T de $\{1, \dots, n\}$ pour laquelle P est la matrice. On effectue alors les trois mêmes figures et on constate en Figure 6 que la matrice solution n'est bien constituée que de points violet et jaune, et en Figure 7 que chaque point n'est impliqué que dans une et une seule connexion selon la matrice de transport solution.

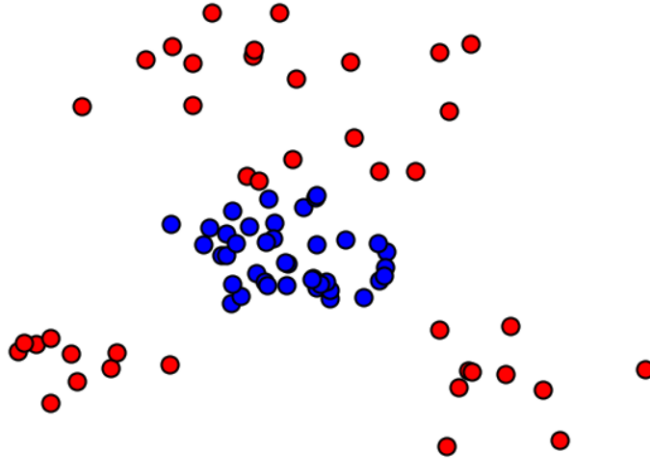


FIGURE 5 – Nuages de points X (en bleu) et Y (en rouge) pour $n = 40$ et $m = 40$

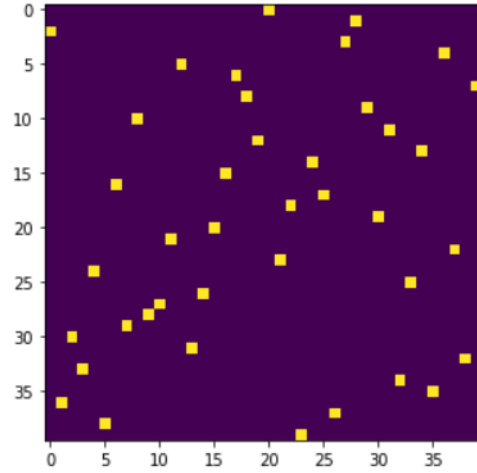


FIGURE 6 – Matrice de permutation P solution au problème de transport discret

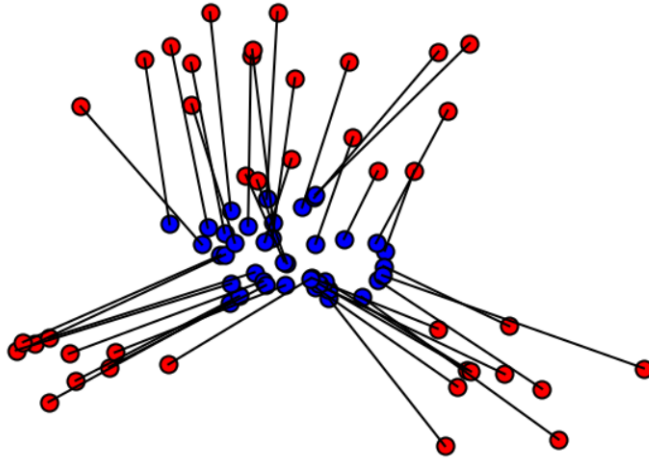


FIGURE 7 – Connexions définies entre les nuages de point par la matrice solution P

3 La théorie de la fonctionnelle de la densité

3.1 Introduction

Soit $d \in \mathbb{N}^*$ et N un entier strictement supérieur à 1 qui représentent respectivement la dimension de l'espace et le nombre d'électrons.

Les électrons se déplacent relativement lentement pour s'éviter autant que possible à cause de leur répulsion et on montre que résoudre le problème de trouver l'énergie minimale de répulsion interne pour N électrons dans une densité donnée $\rho(r)$ équivaut à trouver la manière optimale de transporter $N-1$ fois la densité ρ en elle-même, avec la fonction de coût donnée par la répulsion de coulomb.

On s'attelle, d'abord, dans cette section à introduire les notions essentielles nécessaires au traitement du cas limite de l'interaction forte dans la théorie de la fonctionnelle de la densité (DFT) [7], pour ensuite exposer certains résultats physiques que l'on reformulera dans le contexte de la théorie du transport optimal afin de les démontrer en dernière partie.

3.2 Interactions forte dans DFT

On introduit ici les notions essentielles nécessaire au traitement du cas limite de l'interaction forte dans DFT.

Définition 17. Un fermion est une particule de spin $\frac{1}{2}$.

Proposition 2. Toute fonction d'onde de fermion est antisymétrique.

Proposition 3. Un électron est un fermion.

Corollaire 1. Toute fonction d'onde électronique est antisymétrique.

Définition 18. Dans un cadre général d-dimensionnel, on définit \mathcal{A}_N l'ensemble des fonctions d'onde électronique normalisées de gradient L^2 pour une structure électronique à N électrons :

$$\mathcal{A}_N = \{\Psi \in L^2(\mathbb{R}^{Nd}) | \nabla \Psi \in L^2(\mathbb{R}^{Nd}), \Psi \text{ antisymétrique}, \|\Psi\|_{L^2} = 1\}$$

Remarque.

On notera par abus : $\mathbb{R}^{Nd} = (\mathbb{R}^d)^N$ dans le cas d-dimensionnel.

Dans le cadre de la mécanique quantique, \mathcal{A}_N est défini usuellement [8] dans le cas tridimensionnel, c'est à dire $d = 3$, mais on peut également retrouver cette définition prenant en compte le spin de chaque électron en plus de leur position.

$$\mathcal{A}_N = \{\Psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}^2)^N) | \nabla \Psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}^2)^N), \Psi \text{ antisymétrique}, \|\Psi\|_{L^2} = 1\}$$

Théorème (1er théorème de Hohenberg et Kohn (HK)). *Les propriétés électroniques à l'état fondamental sont calculés en minimisant la fonctionnelle d'énergie $E[\rho]$ par rapport à la densité $\rho(r)$ avec :*

$$E[\rho] = F[\rho] + \int v_{ext}(r)\rho(r)dr \quad (3.1)$$

, où $v_{ext}(r)$ est le potentiel externe et $F[\rho]$ la fonctionnelle universelle de la densité, définie comme la valeur moyenne de l'énergie interne, qui est la somme de l'énergie cinétique $\hat{T} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2$ et de l'énergie d'interaction électron-électron $\hat{V}_{ee} = \sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|}$, de la minimisation de la fonction d'onde générant la densité $\rho(r)$:

$$F[\rho] = \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \quad (3.2)$$

Le problème de minimisation de $E[\rho]$ devient alors celui de $F[\rho]$.

Définition 19. Le minimum de la valeur moyenne de \hat{T} seule sur toutes les fonctions d'onde de fermions générant la densité ρ est donnée par la fonctionnelle $T_s[\rho]$ définie par :

$$T_s[\rho] = \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle \quad (3.3)$$

Définition 20. La partie restante de la fonctionnelle d'énergie exacte $E_{Hxc}[\rho]$, qui est approximée, est définie par :

$$E_{Hxc}[\rho] = F[\rho] - T_s[\rho] \quad (3.4)$$

Ainsi, on a :

$$F[\rho] = T_s[\rho] + E_{Hxc}[\rho]$$

c'est à dire, trivialement :

$$F[\rho] = T_s[\rho] + F[\rho] - T_s[\rho]$$

On a alors :

$$\inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle = \inf_{\Psi' \in \mathcal{A}_N, \Psi' \rightarrow \rho} \langle \Psi' | \hat{T} | \Psi' \rangle + \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle - \inf_{\Psi' \in \mathcal{A}_N, \Psi' \rightarrow \rho} \langle \Psi' | \hat{T} | \Psi' \rangle$$

Néanmoins, rien ne garantit que les minimiseurs Ψ et Ψ' soient les mêmes, c'est pourquoi scinder le problème de minimisation en 2 minimisations n'est qu'une approximation de ce dernier.

De la même manière, le problème de minimisation d'une somme n'est pas exactement le problème de minimisation des termes de cette somme puisque le minimiseur d'une somme n'est pas forcément le même pour chacun des termes de la somme, c'est pourquoi on a :

$$F[\rho] = \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \geq \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle + \inf_{\Psi' \in \mathcal{A}_N, \Psi' \rightarrow \rho} \langle \Psi' | \hat{V}_{ee} | \Psi' \rangle$$

La première quantité de la somme de droite étant $T_s[\rho]$, c'est la deuxième quantité que l'on essaie alors d'approximer afin d'approcher inférieurement $F[\rho]$ avec une bonne précision car $T_s[\rho]$ définit un système de référence n'interagissant pas avec le système d'interaction de la densité électronique.

Définition 21. La fonctionnelle classique de Hartree $U[\rho]$ est définie par :

$$U[\rho] = \frac{1}{2} \int dr \int dr' \frac{\rho(r)\rho(r')}{|r - r'|} \quad (3.5)$$

Proposition 4. On peut séparer $E_{Hxc}[\rho]$ en $E_{Hxc}[\rho] = U[\rho] + E_{xc}[\rho]$, où $E_{xc}[\rho]$ est l'énergie d'échange-corrélation, qui est la quantité cruciale qui est approximée.

On s'est donc ramené au problème de minimisation de la quantité définie maintenant :

Définition 22. Le minimum de la valeur moyenne de \hat{V}_{ee} seule sur toutes les fonctions d'onde générant la densité ρ est donnée par la fonctionnelle $V_{ee}^{SCE}[\rho]$ définie par :

$$V_{ee}^{SCE}[\rho] = \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{V}_{ee} | \Psi \rangle \quad (3.6)$$

Il est possible dans certains cas de construire $V_{ee}^{SCE}[\rho]$ par exemple pour les densités à symétrie sphérique. Reformuler $V_{ee}^{SCE}[\rho]$ en un problème de transport optimal permet de paramétrer la construction explicite de cette fonctionnelle et d'importer des algorithmes d'un autre domaine de recherche bien établi.

3.3 Résultats physiques

On s'intéresse, à présent, au problème de minimisation de $V_{ee}^{SCE}[\rho]$ posé en équation (6) et on anticipe ici les résultats des sous-sections suivantes.

Définition 23. Les fonctions de co-mouvement $f_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$ pour $i \in \{1, 2, \dots, N\}$ sont des fonctionnelles implicites de la densité, déterminés par un ensemble d'équations différentielles, qui assurent l'invariance de la densité par la transformation $s \rightarrow f_i(s)$,

$$\forall i \in \{1, 2, \dots, N\}, \rho(f_i(s)) df_i(s) = \rho(s) ds \quad (3.7)$$

Pour tout $s \in \mathbb{R}^d$ et $i \in \{1, 2, \dots, N\}$, $f_i(s)$ détermine la position du i -ème électron par rapport à s , où s est distribué selon la densité normalisée $\frac{\rho(s)}{N}$. Ainsi, si s détermine la position d'un électron arbitraire que l'on numérotera "1" alors $f_i(s)$ détermine la position du i -ème électron par rapport au électron 1 et possède la même densité par invariance selon f_i . Dans ce cas, on a nécessairement $f_1(s) = s$.

Définition 24. Les configurations admissibles pour N électrons en dimension d sont restreintes à un sous-espace d -dimensionnel Ω_0 de l'espace de configuration Nd -dimensionnel total, où Ω_0 est défini par :

$$\Omega_0 = \{(f_1(s), \dots, f_N(s)) | s \in \mathbb{R}^d \text{ et } \forall i \in \{1, 2, \dots, N\}, f_i \text{ est une fonction de co-mouvement}\}$$

Proposition 5. La fonctionnelle $V_{ee}^{SCE}[\rho]$ est donnée par :

$$V_{ee}^{SCE}[\rho] = \int_X ds \frac{\rho(s)}{N} \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|f_i(s) - f_j(s)|} \quad (3.8)$$

Proposition 6. La fonctionnelle $V_{ee}^{SCE}[\rho]$ correspond au cas limite $\lambda \rightarrow +\infty$ de la connexion adiabatique traditionnelle de la DFT qui est la fonctionnelle $F_\lambda[\rho]$, dépendant de λ , définie par :

$$F_\lambda[\rho] = \inf_{\Psi \in \mathcal{A}_N, \Psi \rightarrow \rho} \langle \Psi | \hat{T} + \lambda \hat{V}_{ee} | \Psi \rangle \quad (3.9)$$

Proposition 7. Si $\Psi_\lambda[\rho]$ est le minimiseur de $F_\lambda[\rho]$ (9) et $W_\lambda[\rho]$ le minimum correspondant :

$$W_\lambda[\rho] = \langle \Psi_\lambda[\rho] | \hat{V}_{ee} | \Psi_\lambda[\rho] \rangle - U[\rho], \quad (3.10)$$

alors, sous des hypothèses légères, on a la formule exacte pour la fonctionnelle d'échange-corrélation de la DFT KS :

$$E_{xc}[\rho] = \int_0^1 W_\lambda[\rho] d\lambda \quad (3.11)$$

Proposition 8. Dans le cas limite $\lambda \rightarrow +\infty$:

$$\lim_{\lambda \rightarrow \infty} W_\lambda[\rho] = V_{ee}^{SCE}[\rho] - U[\rho] \quad (3.12)$$

Remarque. L'existence de cette limite nous permet alors un calcul explicite de $E_{xc}[\rho]$ par un théorème de convergence dominée en passant à la limite lorsque $\lambda \rightarrow +\infty$.

Proposition 9. Si $\rho(r)$ est v -représentable à l'état fondamental, alors pour tout $\lambda \in \mathbb{R}$, $\Psi_\lambda[\rho]$ est l'état fondamental de l'hamiltonien électronique $\hat{H}_\lambda[\rho]$, défini par :

$$\hat{H}_\lambda[\rho] = \hat{T} + \lambda \hat{V}_{ee} + \hat{V}_\lambda[\rho] \quad (3.13)$$

, où $\hat{V}_\lambda[\rho]$ est le potentiel local à un corps, qui maintient la densité égale à la densité physique $\rho(r)$ pour tout λ , défini par :

$$\hat{V}_\lambda[\rho] = \sum_{i=1}^N v_\lambda[\rho](r_i) \quad (3.14)$$

Proposition 10. On a dans le cas limite $\lambda \rightarrow +\infty$:

$$\lim_{\lambda \rightarrow \infty} \frac{v_\lambda[\rho](r)}{\lambda} = v_{SCE}[\rho](r) \quad (3.15)$$

, où $v_{SCE}[\rho](r)$ est la contrepartie du potentiel KS dans la limite de l'interaction forte, liée aux fonctions de co-mouvement par l'équation d'équilibre classique :

$$\nabla v_{SCE}[\rho](r) = \sum_{i=2}^N \frac{r - f_i(r)}{|r - f_i(r)|^3} \quad (3.16)$$

Remarque. La proposition précédente n'est valable que si et seulement si on a l'existence des fonctions de co-mouvement, c'est à dire des fonctions de transports.

Proposition 11. *Plus généralement, on a :*

$$\frac{\partial V_{ee}^{SCE}[\rho]}{\partial \rho(r)} = -v_{SCE}[\rho](r) \quad (3.17)$$

Proposition 12. *On peut relier le potentiel de kantorovitch $u(r)$ et $v_{SCE}[\rho](r)$ par la relation :*

$$u(r) = -v_{SCE}[\rho](r) + C \quad (3.18)$$

où C est la constante définie par :

$$C = -\lim_{|r| \rightarrow \infty} u(r)$$

puisque le potentiel est nul à l'infini lorsque $|r| \rightarrow +\infty$, où $|r|$ est la distance au centre de la charge :

$$\lim_{|r| \rightarrow \infty} v_{SCE}[\rho](r) = 0$$

Dans les théorèmes des sous-sections suivantes, on prouve, sous des hypothèses très légères sur $\rho(r)$, que ce potentiel existe, qu'il est borné, et qu'il est différentiable presque partout et ceci indépendamment de l'existence des fonctions de co-mouvement.

Le théorème 4 alors prouvera que la valeur de $V_{ee}^{SCE}[\rho]$ est donné exactement par le maximum du problème dual de kantorovitch,

$$V_{ee}^{SCE}[\rho] = \sup_u \left\{ \int u(r) \rho(r) dr \mid \sum_{i=1}^N u(r_i) \leq \sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|} \right\} \quad (3.19)$$

Remarque. La condition $\sum_{i=1}^N u(r_i) \leq \sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|}$ a un sens physique simple : elle signifie, qu'à l'optimalité, le sous-espace atteignable Ω_0 de l'espace total de configuration Nd est le minimum de l'énergie potentielle classique. En effet, on peut réécrire la condition :

$$0 \leq \sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|} - \sum_{i=1}^N u(r_i) = \sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|} - \sum_{i=1}^N (-v_{SCE}[\rho](r) + C)$$

D'où :

$$\sum_{1 \leq i < j \leq N} \frac{1}{|r_i - r_j|} + \sum_{i=1}^N v_{SCE}[\rho](r) \geq E_{SCE} \quad (3.20)$$

satisfaite uniquement pour les configurations appartenant à Ω_0 et E_{SCE} l'énergie totale dans la limite SCE :

$$E_{SCE} = \lim_{\lambda \rightarrow \infty} \frac{E_\lambda}{\lambda}$$

, où E_λ est l'énergie de l'état fondamental dans (13).

3.4 L'apport du transport optimal

Soit l'espace métrique $X = \mathbb{R}^d$ qui est bien complet et séparable. On réintroduit donc ici les notions de la théorie du transport optimal selon les notations adoptées. Dans le cas de 2 électrons, le problème est celui du transport d'une distribution de masse initiale en une autre distribution de masse finale. On note respectivement la distribution de masse initiale $p_1 = \rho_1 dx$ et la distribution de masse finale $p_2 = \rho_2 dx$ et la condition de transport de p_1 à p_2 par la fonction de transport f , notée $f_{\#}p_1 = p_2$, signifie que :

$$\forall S \text{ mesurable } \subseteq X, p_2(S) = p_1(f^{-1}(S)) \quad (3.21)$$

On cherche alors à minimiser la quantité ,dépendant de la fonction de coût c , suivante :

$$\int_X c(x, f(x)) dp_1(x) \quad (3.22)$$

Le problème de Monge s'écrit donc :

$$(MP) := \inf \left\{ \int_X c(x, f(x)) dp_1(x), f : X \rightarrow X \text{ et } f_{\#}p_1 = p_2 \right\}$$

Ici les mesures p_1 et p_2 sont les mesures marginales. On les suppose normalisées toutes les deux afin d'assurer qu'elles soient de masse égale. Ainsi p_1 et p_2 sont bien des probabilités sur X .

On peut donc se ramener, par relaxation, au problème de Kantorovitch en adoptant pour $P \in \mathcal{P}(X \times X)$ la notation $P(dx, dy) = P(x, y) dx dy = d\gamma(x, y)$ pour $\gamma \in \mathcal{P}(X \times X)$ correspondant à P :

$$(KP) := \inf_{P \in \mathcal{P}(X \times X)} \left\{ \int_{X \times Y} c(x, y) P(dx, dy), \forall j \in \{1, 2\}, \pi_{j\#}P = p_j \right\} \quad (3.23)$$

Soit l'espace produit $\mathcal{H} = X^N = \mathbb{R}^{Nd}$ qui est également un espace métrique complet et séparable. En adoptant la notation $P(dx_1, dx_2, \dots, dx_N) = d\gamma(x_1, x_2, \dots, x_N)$ pour $P, \gamma \in \mathcal{P}(\mathcal{H})$, on peut généraliser le problème pour N électrons en récrivant le problème de transport multi-marginales d'ordre N , où $X = X_1 = X_2 = \dots = X_N$ et les mesures marginales sont les $(p_i)_{i \in \{1, 2, \dots, N\}}$:

$$(MMP) := \inf_{P \in \mathcal{P}(\mathcal{H})} \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N), \forall i \in \{1, 2, \dots, N\}, \pi_{i\#}P = p_i \right\} \quad (3.24)$$

Le problème de Monge analogue (MP-N) est alors :

$$(MP-N) := \inf \left\{ \int_X c(f_1(x_1), \dots, f_N(x_N)) dp_1(x_1) | f_i : X \rightarrow X, f_1 = Id_X \text{ et } \forall i \in \{2, \dots, N\}, f_{i\#}p_1 = p_i \right\}$$

3.5 Reformulation

On s'attelle ici à reformuler le problème minimisation de $V_{ee}^{SCE}[\rho]$ défini en équation (8). On montre dans cette sous-section que ce problème correspond à un problème de Monge sous la contrainte définie en équation (6), où les fonctions de co-mouvement sont des fonctions de transport, ce qui ne pose pas de problème puisque la condition demandée en équation (21) est une forme faible de celle en équation (6) puisqu'elle ne nécessite pas que f soit différentiable.

Définition 25. On définit la mesure de probabilité de présence $P \in \mathcal{P}(\mathbb{R}^{Nd})$:

$$P(dx_1, dx_2, \dots, dx_N) = |\Psi(x_1, \dots, dx_N)|^2 dx_1 \dots dx_N$$

, où $|\Psi(x_1, \dots, dx_N)|^2$ est la densité de probabilité pour que les N électrons soient aux positions $x_i, i \in \{1, 2, \dots, N\}$.

Définition 26. On définit les mesures marginales $(p_i)_{i \in \{1,2,\dots,N\}}$ comme étant chacune égale à la densité ρ divisé par le nombre N de particules :

$$\forall i \in \{1, 2, \dots, N\}, p_i = \frac{1}{N} \rho dx$$

Définition 27. On choisit comme fonction de coût celle égale à la potentiel de répulsion électron-électron de Coulomb définie pour une structure électronique à N électrons comme :

$$c(x_1, x_2, \dots, x_N) = \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \quad (3.25)$$

Lorsque le coût est coulombien et que les marginales correspondent à la densité normalisée comme défini en définition 23 et 24, on remarque alors que la quantité à minimiser dans le problème (MP-N) est bien $V_{ee}^{SCE}[\rho]$ défini en équation (6) dans la section des résultats physiques. En effet :

$$\int_X c(f_1(x_1), \dots, f_N(x_N)) dp_1(x_1) = \int_X \sum_{1 \leq i < j \leq N} \frac{1}{|f_i(x_i) - f_j(x_j)|} \frac{1}{N} \rho dx_1 = \int_X ds \frac{\rho(s)}{N} \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|f_i(s) - f_j(s)|}$$

Le problème de minimisation de la quantité $V_{ee}^{SCE}[\rho]$ de gauche revient donc à un problème de transport optimal où les fonctions de transport sont les fonctions de co-mouvement $(f_i)_{i \in \{1,2,\dots,N\}}$.

Cependant, prouver l'existence de fonctions de transport est souvent un problème difficile à résoudre. De plus, la condition en équation (6), plus forte, a la forme d'un problème de transport multi-marginales comme celui figurant en équation (24).

3.6 Théorèmes énergétiques

Dans cette sous-section, on établit les résultats anticipés dans les sous-sections précédentes :

Théorème 1. *Si la fonction de coût c est positive et semi-continue inférieurement, alors il existe une mesure de probabilité optimale P_{opt} pour le problème de minimisation (MMP) (24)*

Démonstration. Soit une suite minimisante $(P_n)_{n \in \mathbb{N}}$ au problème (MMP) en équation (24). Par les conditions du problème (MMP), on a bien que $(P_n)_{n \in \mathbb{N}}$ est une suite de mesures de probabilités sur \mathcal{H} qui, par conséquent, est compacte selon la convergence faible- $*$ des mesures. Donc, quitte à passer à une sous-suite, elle converge faiblement- $*$ vers une mesure P sur \mathcal{H} positive par positivité des P_n . Ainsi, on a :

$$\int_{\mathcal{H}} cdP \leq \liminf_n \int_{\mathcal{H}} cdP_n$$

Donc on a :

$$\int_{\mathcal{H}} cdP \leq \inf(MMP)$$

Il suffit donc de montrer que P est bien une probabilité qui vérifie la condition sur les mesures marginales de (MMP) pour obtenir l'autre inégalité. Or pour tout $n \in \mathbb{N}$, P_n satisfait la condition sur les mesures marginales :

$$\forall i \in \{1, 2, \dots, N\}, \pi_{i\#} P_n = p_i$$

, laquelle ne dépend pas de n puisque les mesures marginales $(p_i)_{i \in \{1,2,\dots,N\}}$ sont les mêmes pour tout $n \in \mathbb{N}$. Donc, la suite de mesures de probabilités $(P_n)_{n \in \mathbb{N}}$ est bien tendue.

En effet, pour $\epsilon > 0$ fixé, pour tout $i \in \{1, 2, \dots, N\}$, il existe un compact K_i inclus dans X tel que

$$p_i(X \setminus K_i) < \epsilon^{\frac{1}{N}}$$

Posons $K = K_1 \times K_2 \times \dots \times K_N \in \mathcal{H}$, on a pour tout $n \in \mathbb{N}$:

$$\begin{aligned} P_n(\mathcal{H} \setminus K) &= \int_{\mathcal{H} \setminus K} P_n(dx_1, \dots, dx_N) \\ &= \int_{\mathcal{H} \setminus K} \pi_{1\#} P_n(dx_1) \dots \pi_{N\#} P_n(dx_N) \\ &= \int_{\mathcal{H} \setminus K} p_1(dx_1) \dots p_N(dx_N) \\ &= \int_{\mathcal{H} \setminus K} dp_1(x_1) \dots dp_N(x_N) \\ &= \left(\int_{X \setminus K_1} dp_1(x_1) \right) \dots \left(\int_{X \setminus K_N} dp_N(x_N) \right) < \epsilon^{\frac{1}{N}} \dots \epsilon^{\frac{1}{N}} = (\epsilon^{\frac{1}{N}})^N = \epsilon \end{aligned}$$

Par le théorème de Prokhorov, quitte à passer par une sous-suite, $P_n \rightharpoonup P$ dans la dualité avec $\mathcal{C}_b(X)$, ce qui suffit à montrer, par passage à la limite, que P est bien une probabilité et vérifie bien la condition sur les mesures marginales. On a donc bien l'autre inégalité :

$$\int_{\mathcal{H}} c dP \geq \inf(MMP)$$

qui donne l'égalité voulue pour P . Ainsi $P_{opt} = P$ qui convient. \square

Remarque. Si les mesures marginales p_1, \dots, p_N sont toutes égales et si la fonction de coût satisfait la condition de symétrie :

$$\forall \phi \in \mathfrak{S}_n, c(x_1, \dots, x_N) = c(x_{\phi(1)}, \dots, x_{\phi(N)}) \quad (3.26)$$

où \mathfrak{S}_n est l'ensemble des permutations de $\{1, \dots, N\}$, alors le théorème d'existence précédent est valable avec P_{opt} qui satisfait la même condition de symétrie.

En fait, il suffit de remarquer que si l'on se munit d'une probabilité P , alors la nouvelle probabilité \tilde{P} définie ci-dessous a un coût inférieur ou égal à celui de toute probabilité P et les mêmes mesures marginales :

$$\tilde{P}(x_1, \dots, x_N) = \frac{1}{N!} \sum_{\phi \in \mathfrak{S}_n} P(x_{\phi(1)}, \dots, x_{\phi(N)})$$

Le problème peut alors être reformulé en un problème dual.

Théorème 2. *Si c est semi-continue inférieurement et à valeur finie, alors :*

$$\begin{aligned} &\inf_P \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) \mid \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ &= \sup_{u_j} \left\{ \sum_{j=1}^N \int_{X_j} u_j dp_j \mid u_j : X_j \rightarrow \mathbb{R}, u_j \in \mathcal{C}_b(X_j), \sum_{j=1}^N u_j(x_j) \leq c(x_1, \dots, x_N) \right\} \end{aligned}$$

Corollaire 2. *Le problème dual de maximisation (DMMP) admet aussi une solution :*

$$(DMMP) := \sup_{u_j} \left\{ \sum_{j=1}^N \int_{X_j} u_j dp_j \mid u_j : X_j \rightarrow \mathbb{R}, u_j \in \mathcal{C}_b(X_j), \sum_{j=1}^N u_j(x_j) \leq c(x_1, \dots, x_N) \right\}$$

Lemme 1. *Si c est semi-continue inférieurement et bornée inférieurement et que $(c_k)_{k \in \mathbb{N}}$ est une suite croissante de fonctions semi-continues inférieurement, bornées inférieurement et convergant vers c , alors :*

$$\begin{aligned} & \lim_{k \rightarrow +\infty} \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ &= \inf_{\bar{P}} \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \end{aligned}$$

preuve du lemme. Par croissance de la suite $(c_k)_{k \in \mathbb{N}}$, on a pour tout $k, n \in \mathbb{N}$ tels que $k < n$, $c_k \leq c_n$. En faisant tendre $n \rightarrow +\infty$, on a $c_k \leq c$. Donc :

$$\begin{aligned} & \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ & \leq \inf_P \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \end{aligned}$$

En passant à la limite lorsque $k \rightarrow +\infty$, on a par un théorème de convergence monotone :

$$\begin{aligned} & \lim_{k \rightarrow +\infty} \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ & \leq \inf_P \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \end{aligned}$$

Montrons l'autre inégalité. Pour tout $k \in \mathbb{N}$, on a une probabilité optimale P_k au problème (MMP) relatif au coût c_k :

$$\inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\}$$

Puisque la suite $(P_k)_{k \in \mathbb{N}}$ est tendue, quitte à passer par une sous-suite, on peut supposer sans perte de généralités que $P_k \rightharpoonup \bar{P}$. Fixons j un entier tel que $k \geq j$. On a alors $c_k \geq c_j$ et par conséquent :

$$\begin{aligned} & \liminf_k \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ &= \lim_k \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P_k(dx_1, dx_2, \dots, dx_N) \geq \liminf_k \int_{\mathcal{H}} c_j(x_1, x_2, \dots, x_N) P_k(dx_1, dx_2, \dots, dx_N) \end{aligned}$$

Puisque c_j est semi-continue :

$$\liminf_k \int_{\mathcal{H}} c_j(x_1, x_2, \dots, x_N) P_k(dx_1, dx_2, \dots, dx_N) \geq \int_{\mathcal{H}} c_j(x_1, x_2, \dots, x_N) \bar{P}(dx_1, dx_2, \dots, dx_N)$$

Par conséquent :

$$\begin{aligned} & \liminf_k \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ & \geq \int_{\mathcal{H}} c_j(x_1, x_2, \dots, x_N) \bar{P}(dx_1, dx_2, \dots, dx_N) \end{aligned}$$

Puisque j est arbitraire et que $\lim_j \int c_j d\bar{P} = \int c d\bar{P}$ par convergence monotone. On a alors :

$$\begin{aligned} \liminf_k \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ \geq \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) \bar{P}(dx_1, dx_2, \dots, dx_N) \\ \geq \inf_P \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \end{aligned}$$

D'où la seconde inégalité et la preuve. \square

preuve du théorème. On a déjà une inégalité puisque le deuxième terme est plus petit que le premier terme en passant à la borne inférieure sur P à droite puis à la borne supérieure sur les u_j à gauche dans cette inégalité :

$$\sum_{j=1}^N \int_{X_j} u_j dp_j = \int_{\mathcal{H}} \sum_{j=1}^N u_j P(dx_1, dx_2, \dots, dx_N) \leq \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N)$$

Il suffit maintenant de montrer l'autre inégalité. Soit une suite croissante $(c_k)_{k \in \mathbb{N}}$ de fonctions k -lipshitziennes convergeant vers c . Pour tout $k \in \mathbb{N}$, c_k est uniformément continue car lipshitzienne. Donc on a la formule de dualité :

$$\begin{aligned} \inf_P \left\{ \int_{\mathcal{H}} c_k(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ = \sup_{u_j} \left\{ \sum_{j=1}^N \int_{X_j} u_j dp_j | u_j : X_j \rightarrow \mathbb{R}, u_j \in \mathcal{C}_b(X_j), \sum_{j=1}^N u_j(x_j) \leq c_k(x_1, \dots, x_N) \right\} \\ \leq \sup_{u_j} \left\{ \sum_{j=1}^N \int_{X_j} u_j dp_j | u_j : X_j \rightarrow \mathbb{R}, u_j \in \mathcal{C}_b(X_j), \sum_{j=1}^N u_j(x_j) \leq c(x_1, \dots, x_N) \right\} \end{aligned}$$

Par croissance de $(c_k)_{k \in \mathbb{N}}$ convergeant vers c , on a $c_k \leq c$ qui assure que tout N -uplet (u_1, \dots, u_N) qui satisfait $\sum_{j=1}^N u_j(x_j) \leq c_k(x_1, \dots, x_N)$ satisfait aussi $\sum_{j=1}^N u_j(x_j) \leq c(x_1, \dots, x_N)$. D'où l'inégalité ci-dessus. On conclut par le lemme 1 en faisant tendre $k \rightarrow +\infty$ dans cette inégalité pour obtenir la seconde inégalité. D'où la preuve. \square

Proposition 13. *Si les mesures marginales sont toutes égales et si la fonction de coût satisfait la condition de symétrie (26), alors le problème dual devient :*

$$(DMMP2) := \sup_u \left\{ N \int_X u dp | u : X \rightarrow \mathbb{R}, u \in \mathcal{C}_b(X), \sum_{i=1}^N u(x_i) \leq c(x_1, \dots, x_N) \right\} \quad (3.27)$$

où $X = X_1 = \dots = X_N$. Une fonction u solution au problème dual (DMMP2) (27) est appelé potentiel de Kantorovitch.

Démonstration. Si les mesures marginales $p_1 = \dots = p_N = p$ et si la fonction de coût satisfait la condition de symétrie (26), alors le problème dual admet une solution $u_1 = \dots = u_N = u$. En fait, si u_1, \dots, u_N est une solution au problème dual (DMMP), alors la fonction :

$$u(x) = \frac{1}{N} [u_1(x) + \dots + u_N(x)]$$

a le même coût maximal au problème dual car :

$$\sum_{j=1}^N \int_{X_j} u_j dp_j = \sum_{j=1}^N \int_X u_j dp = \int_X \sum_{j=1}^N u_j dp = \int_X N u dp = N \int_X u dp$$

De plus, elle satisfait la contrainte :

$$u(x_1) + \dots + u(x_N) \leq c(x_1, \dots, x_N)$$

En effet, on a :

$$u(x_1) + \dots + u(x_N) = \frac{1}{N} [u_1(x_1) + \dots + u_N(x_1)] + \dots + \frac{1}{N} [u_1(x_N) + \dots + u_N(x_N)]$$

En réarrangeant les termes selon les u_i , on a :

$$\begin{aligned} u(x_1) + \dots + u(x_N) &= \frac{1}{N} [u_1(x_1) + \dots + u_1(x_N)] + \dots + \frac{1}{N} [u_N(x_1) + \dots + u_N(x_N)] \\ &= \frac{1}{N} [u_1(x_1) + \dots + u_N(x_N)] + \dots + \frac{1}{N} [u_1(x_1) + \dots + u_N(x_N)] \\ &\leq \frac{1}{N} c(x_1, \dots, x_N) + \dots + \frac{1}{N} c(x_1, \dots, x_N) = N \frac{1}{N} c(x_1, \dots, x_N) \end{aligned}$$

D'où la contrainte satisfaite

□

Cependant, ce théorème ne s'applique pas directement au problème de transport optimal étudié ici car la fonction de coût donné en (25) prend la valeur $+\infty$ sur l'ensemble $\{(x_1, \dots, x_N) \in \mathbb{R}^{Nd} | \exists i, j \in \{1, \dots, N\}, i \neq j \text{ et } x_i = x_j\}$. On reformule alors le problème :

Théorème 3. *Soit c une fonction borélienne à valeur dans $\mathbb{R}_+ \cup \{+\infty\}$. Si c est finie presque partout relativement à la mesure $p_1 \otimes \dots \otimes p_N$ et si il existe un plan de transport \hat{P} de coût fini. Alors il existe u_i des maximiseurs borel-mesurables solution au problème dual (DMMP2) à valeurs dans $\mathbb{R} \cup \{-\infty\}$ tels que*

$$\begin{aligned} &\inf_P \left\{ \int_{\mathcal{H}} c(x_1, x_2, \dots, x_N) P(dx_1, dx_2, \dots, dx_N) | \forall j \in \{1, \dots, N\}, \pi_{j\#} P = p_j \right\} \\ &= \sup_{u_j} \left\{ \int_{\mathcal{H}} \sum_{j=1}^N u_j(x_j) d\hat{P}(x_1, \dots, x_N) | u_j : X_j \rightarrow \mathbb{R}, \sum_{j=1}^N u_j(x_j) \leq c(x_1, \dots, x_N) \right\} \end{aligned}$$

L'hypothèse caractérisant c comme une fonction borélienne est assez large pour inclure le cas des fonctions continues et semi-continues inférieurement prenant éventuellement la valeur $+\infty$, en particulier, le potentiel de Coulomb en équation (25)

La forme duale du théorème 3 ne permet pas de calcul explicite puisqu'il nécessite un plan qui n'est pas toujours explicitement connu. Pour contourner cette difficulté, on peut donner, sous certaines hypoythèses, une forme duale plus utile :

Théorème 4. *Soit c la fonction de coût (25). Si toutes les mesures marginales p_j coïncident. Alors il existe un maximiseur u solutions du problème dual (DMMP2) qui satisfait la formule suivante :*

$$u(x) = \inf_{y_i} \left\{ c(x, y_1, \dots, y_{N-1}) - \sum_{i=1}^{N-1} u(y_i), y_i \in \mathbb{R}^d \right\}$$

De plus, un tel potentiel de Kantorovitch u est borné et vérifie :

$$\int_{\mathcal{H}} \sum_{j=1}^N u_j(x_j) d\hat{P}(x_1, \dots, x_N) = N \int u(x) dp(x)$$

Une preuve de ce théorème ainsi que des exemples de maximiseurs sont déjà donnés dans [7] pour le cas de à deux électrons $N = 2$ et reviennent donc à l'étude d'un problème de Monge-Kantorovitch classique mais avec un coût non fini sur un ensemble fini de points.

Corollaire 3. *Si $p = \frac{1}{N} \rho dx$, alors u est différentiable presque partout et ∇u est localement borné.*

4 L'équilibre de Cournot-Nash

4.1 Le modèle d'équilibre

On introduit, dans cette sous-section, les notions nécessaires à la définition d'un équilibre de Cournot-Nash [9].

Soit X, Y des espaces métriques compacts, où X est un espace de types d'individus et Y un espace d'actions. Soit $\mu \in \mathcal{P}(X)$, donnant la distribution des types. On considère également une mesure borélienne positive de référence m_0 et $\nu \in \mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$ la distribution des actions.

Le phénomène où ν a une masse infinie en un point est appelé congestion. Le rôle de m_0 est d'interdire les effets de congestion et de modéliser ν puisque ν doit être absolument continue par rapport à m_0 .

De plus, on considère le coût $C : X \times Y \times \mathcal{P}(Y) \rightarrow \mathbb{R}$ séparable additivement, c'est à dire :

$$C(x, y, \nu) = c(x, y) + \mathcal{V}[\nu](y),$$

où $c \in \mathcal{C}(X \times Y)$ et $\mathcal{V}[\nu]$ est défini m_0 -p.p. La fonction C représente le coût que les individus de type x paient lorsqu'ils prennent l'action y , étant donné ν . La fonction $\mathcal{V}[\nu](y)$ est appelée fonction d'interaction et représente les différentes interactions entre les individus, telles que la congestion ou d'autres coûts spécifiques à l'action y .

On cherche une mesure de probabilité $\gamma \in \mathcal{P}(X \times Y)$ de première et seconde mesure marginale respectives μ et ν . Alors $\gamma(A \times B)$ représente la probabilité que le type d'un individu soit dans A et qu'il prenne une action dans B . On impose également que l'intégrale du coût par rapport à γ soit finie, c'est-à-dire que l'intégrale suivante soit finie.

$$\iint_{X \times Y} (c(x, y) + \mathcal{V}[\nu](y)) d\gamma(x, y) = \iint_{X \times Y} c(x, y) d\gamma(x, y) + \int_Y \mathcal{V}[\nu](y) d\nu(y).$$

Puisque $c \in \mathcal{C}(X \times Y)$, le premier terme est fini et si

$$\nu \in \mathcal{D} := \{\nu \in \mathcal{L}^1(m_0) : \mathcal{V}[\nu] \in \mathcal{L}^1(\nu)\},$$

le second terme est fini également.

On peut illustrer le modèle d'équilibre par l'exemple suivant.

Exemple. Soit X l'espace des résidences d'une ville et Y l'espace des restaurants de cette même ville. Nous pouvons supposer que X et Y sont des sous-ensembles compacts de \mathbb{R}^2 . Soit $\mu \in \mathcal{P}(X)$ soit la distribution des résidences. Les habitants choisissent le restaurant dans lequel ils vont manger. Le coût $C(x, y, \nu) = c(x, y) + \mathcal{V}[\nu](y)$ sera la somme du coût de déplacement

$c(x, y)$ et du coût donné par la fonction d'interaction $\mathcal{V}[\nu](y)$. La carte d'interaction peut être écrite comme une somme d'autres coûts liés à différents effets.

Le premier effet est la congestion. Dans cet exemple, cela signifie que les restaurants ne peuvent pas servir plus de personnes que leur nombre de sièges. De plus, certaines personnes peuvent choisir des restaurants moins encombrés. Ce coût peut s'écrire sous la forme $f(\frac{d\nu}{dm_0}(y))$, où f est croissant. Ainsi, plus le restaurant y a de masse, plus ce coût sera élevé.

Le deuxième effet est géographique et qualitatif : ainsi, les restaurants situés en centre-ville ou offrant une belle vue peuvent être plus chers. La qualité de la nourriture ou une décoration agréable peuvent aussi influencer sur le prix. De plus, ces types de restaurants peuvent procurer plus de plaisir, et cet effet est donc également pris en compte dans le coût. Ainsi, le coût peut augmenter ou diminuer en raison des facteurs géographiques et qualitatifs et il peut être modélisé par une fonction $g_1(y) + g_2(y)$, où le premier terme est le coût géographique et le second terme est le coût qualitatif.

Enfin, on doit tenir compte du fait que les restaurants isolés peuvent être moins attrayants, c'est-à-dire que la fonction $y \mapsto \int_Y h(y, z) d\nu(z)$ modélise ce coût et qu'il augmente lorsque le restaurant y est éloigné du reste des restaurants et que $z = y$ minimise $h(y, z)$.

Donc, on peut écrire la fonction d'interaction comme

$$\mathcal{V}[\nu](y) = f\left(\frac{d\nu}{dm_0}(y)\right) + g_1(y) + g_2(y) + \int_Y h(y, z) d\nu(z).$$

Définition 28. Soit $\gamma \in \mathcal{P}(X \times Y)$. On dit que γ est un *équilibre de Cournot-Nash* si sa première marginale est μ , sa seconde marginale est ν , appartient à \mathcal{D} et qu'il existe $\varphi \in \mathcal{C}(X)$ tel que :

$$\begin{aligned} c(x, y) + \mathcal{V}[\nu](y) &\geq \varphi(x) \quad \forall x \in X \text{ et } m_0\text{-p.p. } y \\ c(x, y) + \mathcal{V}[\nu](y) &= \varphi(x) \quad \gamma\text{-p.p.} \end{aligned} \tag{4.1}$$

De plus, un équilibre de Cournot-Nash γ est dit *pure* si il existe une fonction borélienne $T : X \rightarrow Y$ telle que $\gamma = (\text{id}, T)_\# \mu$.

Remarque. De manière équivalente, $\gamma \in \mathcal{P}(X \times Y)$ est un équilibre de Cournot-Nash de première et seconde marginale μ et ν si :

$$\gamma(\{(x, y) \in X \times Y : C(x, y, \nu) = \min_{z \in Y} C(x, z, \nu)\}) = 1,$$

qui illustre que γ est un équilibre si les individus prennent des actions minimisant le coût.

Définition 29. On définit $\Pi_o(\mu, \nu)$ l'ensemble des plans de transports γ solutions au problème de Kantorovitch. Notons que cet ensemble est non vide par le théorème d'existence de Kantorovitch.

Lemme 2. Si γ est un équilibre de Cournot-Nash, alors $\gamma \in \Pi_o(\mu, \nu)$.

Démonstration. Soit $\lambda \in \Pi(\mu, \nu)$ et $\varphi \in \mathcal{C}(X)$ satisfaisant (4.1). Alors

$$\begin{aligned} \iint_{X \times Y} c(x, y) d\lambda(x, y) &\geq \iint_{X \times Y} (\varphi(x) - \mathcal{V}[\nu](y)) d\lambda(x, y) \\ &= \int_X \varphi(x) d\mu(x) - \int_Y \mathcal{V}[\nu](y) d\nu(y) = \iint_{X \times Y} c(x, y) d\gamma(x, y), \end{aligned}$$

où la dernière égalité vient de (4.1). Donc $\gamma \in \Pi_o(\mu, \nu)$. □

En outre, il existe un lien entre les équilibres de Cournot-Nash et le problème dual.

Lemme 3. Soit $\varphi \in \mathcal{C}(X)$ satisfaisant (4.1). Alors (φ, φ^c) est solution du problème dual (1.1).

Ce lemme vient du théorème de dualité en section 1.4

Si (φ, φ^c) est solution du problème dual (1.1), on a :

$$\begin{aligned}\frac{\delta}{\delta\mu}\mathcal{KP}(\mu, \nu) &= \varphi(x) \\ \frac{\delta}{\delta\nu}\mathcal{KP}(\mu, \nu) &= \psi(y).\end{aligned}$$

Si ν^* minimise $\mathcal{J}_\mu[\nu]$, alors :

$$\frac{\delta}{\delta\nu}\mathcal{J}_\mu[\nu^*] = 0 \quad \nu^*\text{-p.p.}$$

D'où,

$$\frac{\delta}{\delta\nu}\mathcal{J}_\mu[\nu^*] = \frac{\delta}{\delta\nu}\mathcal{KP}(\mu, \nu^*) + \frac{\delta}{\delta\nu}\mathcal{E}[\nu^*] = \psi(y) + \mathcal{V}[\nu^*](y) = 0 \quad \nu^*\text{-p.p.}$$

Or $\psi(y) + \varphi(x) \leq c(x, y)$,

$$c(x, y) + \mathcal{V}[\nu^*](y) \geq \varphi(x)$$

4.2 Recherche d'équilibre

Dans cette sous-section, on souhaite trouver l'équilibre $\gamma \in \mathcal{P}(X \times Y)$. On fixe $p \in [1, +\infty[$ et on suppose que $\mathcal{V}[\nu]$ est la différentielle d'une fonctionnelle \mathcal{E} :

$$\mathcal{V}[\nu] = \frac{d}{d\nu}\mathcal{E}[\nu] \quad \text{pour tout } \nu \in \mathcal{D}_p := \mathcal{P}(Y) \cap \mathcal{L}^p(m_0). \quad (4.2)$$

Alors l'objectif est de résoudre le problème variationnel

$$\inf_{\nu \in \mathcal{D}_p} \mathcal{J}_\mu[\nu] \quad \text{où} \quad \mathcal{J}_\mu[\nu] := \mathcal{KP}(\mu, \nu) + \mathcal{E}[\nu]. \quad (4.3)$$

Le théorème suivant montre que si ν minimise $\mathcal{KP}(\mu, \nu) + \mathcal{E}[\nu]$ et γ minimise $\mathcal{KP}(\mu, \nu)$, alors γ est un équilibre de Cournot-Nash. Donc, au lieu de rechercher une solution au problème de Kantorovitch, on va rechercher une solution au problème variationnel pour trouver l'équilibre.

Théorème 5. On suppose que $\Omega \subset \mathbb{R}^d$ est borné de frontière négligeable, que $X = \overline{\Omega}$ et que $c(\cdot, y)$ est différentiable pour tout $y \in Y$ tel que $\frac{\partial}{\partial x}c(x, y)$ est borné sur $X \times Y$. De plus, on suppose que μ est absolument continue par rapport à la mesure de Lebesgue sur X et que $\mathcal{V}[\nu]$ est la différentielle d'une fonctionnelle au sens de (4.2). Si ν est solution au problème variationnel (4.3) et $\gamma \in \Pi_o(\mu, \nu)$, alors γ est un équilibre de Cournot-Nash.

4.3 Calculs numérique

Dans cette sous-section, nous allons implémenter un algorithme pour résoudre un cas unidimensionnel du problème variationnel :

$$\inf_{\nu \in \mathcal{D}_p} \mathcal{J}_\mu[\nu] \quad \mathcal{J}_\mu[\nu] = \mathcal{KP}(\mu, \nu) + \mathcal{E}[\nu].$$

Afin de résoudre ce problème, on aura besoin des théorèmes suivants.

Définition 30. Soit $\mu \in \mathcal{P}(\mathbb{R})$. La fonction de répartition de μ $F_\mu : \mathbb{R} \rightarrow [0, 1]$ est définie comme :

$$F_\mu(x) = \mu((-\infty, x])$$

et son *pseudo-inverse* $F_\mu^{[-1]} : [0, 1] \rightarrow \overline{\mathbb{R}}$ est donné par :

$$F_\mu^{[-1]}(x) = \inf\{t \in \mathbb{R} : F_\mu(t) \geq x\}.$$

Remarque. Si l'ensemble précédent est non vide et borné inférieurement, l'infimum est un minimum. S'il est vide, alors $x = 1$ et $F_\mu^{[-1]}(1) = +\infty$ et si elle n'est pas bornée inférieurement, alors $x = 0$ et $F_\mu^{[-1]}(0) = -\infty$.

Théorème 6. On suppose que $\mu, \nu \in \mathcal{P}(\mathbb{R})$ et que ν est sans atomes. Alors il existe une unique fonction croissante $S : \mathbb{R} \rightarrow \mathbb{R}$ telle que $S_\# \nu = \mu$ et cette fonction est donnée par :

$$S(x) = F_\mu^{[-1]}(F_\nu(x)).$$

Remarque. De même, si μ est sans atomes, la fonction $T : \mathbb{R} \rightarrow \mathbb{R}$ avec $T(x) = F_\nu^{[-1]}(F_\mu(x))$ donne l'unique fonction croissante satisfaisant $T_\# \mu = \nu$.

On introduit maintenant les notions utilisées dans l'algorithme. On suppose que $X = Y = [0, 1]$, que la fonction de coût c est convexe et que μ et ν sont toutes les deux absolument continues par rapport à la mesure de Lebesgue.

Par le théorème 6, il existe une fonction $S : Y \rightarrow X$ avec

$$S(y) = F_\mu^{[-1]}(F_\nu(y)) \quad \text{and} \quad S_\# \nu = \mu.$$

Si (φ, ψ) est solution au problème dual, on a que :

$$\varphi(x) + \psi(y) = c(x, y) \quad \gamma\text{-p.p.}$$

En différentiant par rapport à y ,

$$\psi'(y) = \frac{\partial}{\partial y} c(x, y)$$

Puis en intégrant et en écrivant $x = S(y)$,

$$\psi(y) = \int_{-\infty}^y \frac{\partial}{\partial y} c(S(t), t) dt.$$

Si ν^* est un point critique de $\mathcal{J}_\mu[\nu] = \mathcal{KP}(\mu, \nu) + \mathcal{E}[\nu]$, alors

$$\frac{d}{d\nu} \mathcal{J}_\mu[\nu^*] = \tilde{c} \quad \nu^*\text{-p.p.},$$

où \tilde{c} est une constante et

$$\begin{aligned} \frac{d}{d\nu} \mathcal{J}_\mu[\nu] &= \frac{\partial}{\partial \nu} \mathcal{KP}(\mu, \nu) + \frac{d}{d\nu} \mathcal{E}[\nu] \\ &= \psi(y) + f(\nu(y)) + v(y) + \int_{-\infty}^{+\infty} \varphi(y, z) d\nu(z) \\ &= \int_{-\infty}^y \frac{\partial}{\partial y} c(S(t), t) dt + f(\nu(y)) + v(y) + \int_{-\infty}^{+\infty} \varphi(y, z) d\nu(z) \end{aligned}$$

Si on prend $f(x) := \log(x)$, on obtient

$$\begin{aligned}\nu(y) &= \exp \left\{ \tilde{c} - \int_{-\infty}^y \frac{\partial}{\partial y} c(S(t), t) dt - v(y) - \int_{-\infty}^{+\infty} \varphi(y, z) d\nu(z) \right\} \\ &= c' \exp \left\{ - \int_0^y \frac{\partial}{\partial y} c(S(t), t) dt - v(y) - \int_0^1 \varphi(y, z) d\nu(z) \right\},\end{aligned}$$

où c' est telle que $\int \nu(y) = 1$.

4.4 Un algorithme de calculs numériques

On fixe μ sur une distribution uniforme $[0, 1]$. L'algorithme consiste à itérer ν avec $S = F_\mu^{-1}(F_{\nu^k}(y)) = F_{\nu^k}(y)$:

$$\nu^{k+1}(y) = c' \exp \left\{ - \int_0^y \frac{\partial}{\partial y} c(S(t), t) dt - v(y) - \int_0^1 \varphi(y, z) d\nu^k(z) \right\}.$$

Dans le cas discret, toutes les intégrales sont remplacées par des sommes de la manière suivante

$$\nu_i^{k+1} = c' \exp \left\{ - \frac{1}{N} \sum_{j=0}^i \frac{\partial}{\partial y} c(S_j, y_j) - v_i - \frac{1}{N} \sum_{j=1}^N \varphi(y_j, z_j) \nu_j^k \right\}.$$

Ceci nous permet d'effectuer des simulations pour un coût euclidien en faisant varier les fonctions paramètres sur les 3 configurations respectives :

Première simulation : $f(x) = \log(x)$, $\phi(y, z) = 10^{-1}(y - z)^2$ et $V(y) = 10|y - 0.7|^2$

Seconde simulation : $f(x) = \log(x)$, $\phi(y, z) = 10^3|y - z|^2$ et $V(y) = 10^2|y - 0.4|^2$

Troisième simulation : $f(x) = x^p$, $\phi(y, z) = 10|y - z|^2$ et $V(y) = 10|y - 0.6|^2$

*

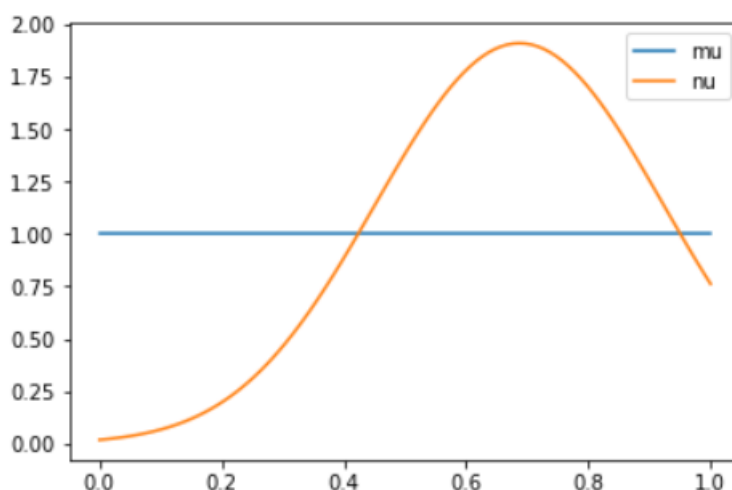


FIGURE 8 – Première simulation de μ et ν

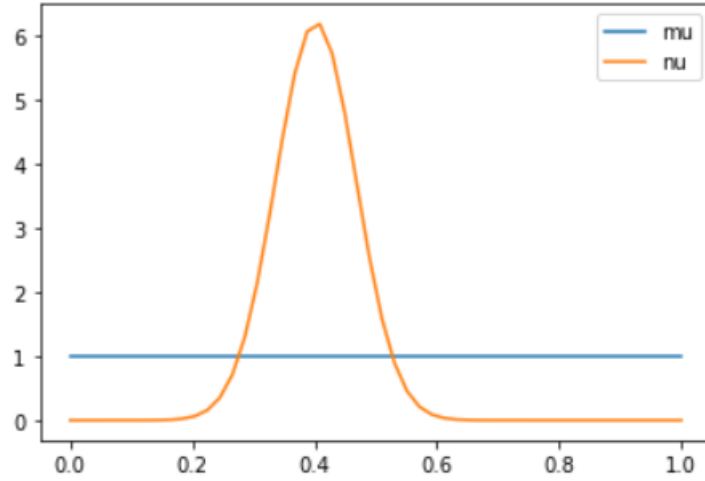


FIGURE 9 – Seconde simulation de μ et ν

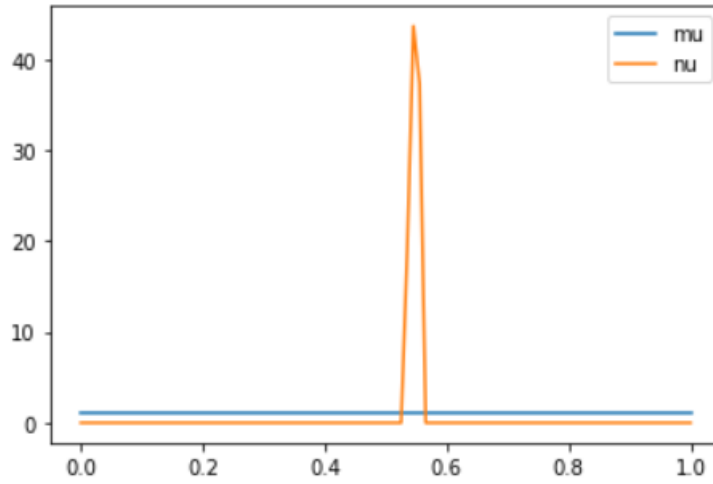


FIGURE 10 – Troisième simulation de μ et ν

4.5 Équilibre de Cournot-Nash dans le cas fini

Dans cette section on se restreindra au cas fini [10] d'équilibre de Cournot-Nash. On réintroduit alors les notions déjà abordées mais cette fois-ci dans le cas fini.

Dans une population donnée, on considère un nombre fini de types d'agent $X = \{x_1, \dots, x_n\}$. Chaque individu est associé à un unique type d'agent et on désigne par $\mu_i = \mu(x_i)$ la fréquence d'un certain type d'agent au sein de la population. On remarque que μ est une mesure de probabilités.

On suppose également disposer d'un nombre fini de stratégies $Y = \{y_1, \dots, y_m\}$. Chaque individu doit choisir une unique stratégie dans Y qui lui sera d'un certain coût.

On s'intéresse ici aux matrices $(\gamma_{i,j})_{i,j \in \{1, \dots, n\} \times \{1, \dots, m\}} \in \mathcal{M}_{n,m}$ telles que $\gamma_{i,j}$ représente la probabilité qu'un individu soit de type x_i et choisisse la stratégie y_j .

On a en sommant les colonnes :

$$\sum_{j=1}^m \gamma_{ij} = \mu_i \quad \forall i \in \llbracket 1; n \rrbracket \quad (*1)$$

Et posons, en sommant en lignes :

$$\sum_{i=1}^n \gamma_{ij} =: \nu_j \quad \forall j \in \llbracket 1; m \rrbracket \quad (*2)$$

ν est une mesure de probabilité et ν_j représente la probabilité qu'un joueur quelconque choisisse la stratégie ν_j .

Chaque agent du type x_i en choisissant la stratégie y_j subit un coût qui dépend non seulement de x_i et y_j mais aussi de $\nu = \{\nu_j ; 1 \leq j \leq m\}$. Les choix des autres joueurs ont donc un impact sur la stratégie choisie par l'un d'entre eux.

Exemple : On peut imaginer la situation suivante : Des médecins souhaitent ouvrir un cabinet médical. Les différents types d'agent X représentent l'endroit où vivent les médecins et les stratégies Y les lieux où ils vont exercer. Pour avoir un maximum de clients, les médecins ont intérêt à ne pas ouvrir leur cabinets là où beaucoup de leurs collègues exercent.

Notation : Notons $\Psi_{ij}[\nu]$ ce coût subi par l'agent du type x_i en choisissant y_j .

Définition 31. Un équilibre de Cournot-Nash est une matrice $\gamma = (\gamma_{ij})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}} \in \mathcal{M}_{n \times m}(\mathbb{R})$ qui vérifie (*1) et (*2) et telle que :

$$\forall i, j \in \{1, \dots, n\} \times \{1, \dots, m\} \quad \text{tq} \quad \gamma_{ij} > 0, \quad \Psi_{ij}[\nu] = \min_{k \in \{1, \dots, m\}} \Psi_{ik}[\nu]$$

Remarque. Soient μ et ν deux mesures de probabilité sur \mathbb{R} avec F et G leur fonctions de répartition respective. Posons $c(x, y) = |x - y|^2$ la norme euclidienne.

Alors :

$$\mathcal{T}_c(\mu, \nu) = \int_0^1 |F^{-1}(t) - G^{-1}(t)|^2 dt$$

Dans le cas plus général pour la fonction $c : X \times Y \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ telle que $c(x, y) = h(x - y)$ avec h une fonction convexe, on aura

$$\mathcal{T}_c(\mu, \nu) = \int_0^1 h(F^{-1}(t) - G^{-1}(t)) dt$$

Plus particulièrement pour $c(x, y) = |x - y|$ on aura :

$$\mathcal{T}_c(\mu, \nu) = \int_0^1 |F^{-1}(t) - G^{-1}(t)| dt$$

4.6 Une approche variationnelle des équilibres de Cournot-Nash

Dans cette partie on montre que pour une certaine forme de $\Psi_{i,j}$ la recherche d'un équilibre de Cournot-Nash peut se ramener au calcul du plan de transport optimal entre μ et une mesure

ν obtenue comme solution d'un problème d'optimisation.

Dans la suite on suppose que Ψ_{ij} a la forme variationnelle suivante :

$$\Psi_{ij}[\nu] := c_{ij} + f_j(\nu_j) + \sum_{k=1}^m \Phi_{kj} \nu_k$$

où :

- $c = (c_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{R}^+)$.
- Pour $j \in \llbracket 1; m \rrbracket$, f_j une fonction non décroissante et continue.
- $\Phi = (\Phi_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{R})$ une matrice symétrique.

Exemple : Si on reprend l'exemple précédent, un médecin habitant dans la zone x_i choisissant d'ouvrir un cabinet dans la zone y_j subira la somme de ces différents coûts :

- le coût de transport entre sa maison et son clinique $c_{ij} = c(x_i, y_j)$.
- le coût de congestion $f_j(\nu_j)$: Un quartier y_j avec beaucoup de médecins (ν_j élevé) va diminuer le bénéfice du médecin (d'où la propriété de non décroissance de f).
- le coût d'interaction avec les autres docteurs. Φ_{kj} est par exemple la distance entre y_k et y_j et $\sum_{k \in \{1, \dots, m\}} \Phi_{kj} \nu_k$ est la distance moyenne de y_j au reste de la population.

L'approche variationnelle consiste à chercher d'abord la mesure une mesure ν solution du problème d'optimisation suivant :

$$\inf_{\nu \in \mathcal{P}(X)} \mathcal{T}_c(\mu, \nu) + E(\nu)$$

Où :

$$E(\nu) = \sum_{j \in J} F_j(\nu_j) + \sum_{k, j \in J} \Phi_{j,k} \nu_j \nu_k \text{ et } F_j = \int f_j$$

Le résultat suivant montre que si ν est solution de ce problème d'optimisation, alors le plan de transport optimal entre μ et ν donne un équilibre de Cournot-Nash.

Théorème 7. *Si ν vérifie le problème d'optimisation précédent et γ est un transport optimal pour les mesures μ et ν alors γ est un équilibre de Cournot-Nash.*

Une référence pour ce résultat est dans [9], page 12.

Remarque : Ce résultat est également vrai dans le cas continu où $\Psi(x, y)(\nu) = c(x, y) + f(\nu) + \int_Y \phi(y, z) d\nu(z)$

4.7 Approche numérique

Dans cette section, on cherche à approcher la mesure ν solution du problème d'optimisation précédent.

On se place d'abord dans un cas continu pour expliquer la méthode puis on reviendra au cas discret.

4.7.1 Cas continu

On prend $X = Y = [0, 1]$, $c(x, y) = \frac{1}{2}(x - y)^2$.

On suppose disposer de ν minimisant la quantité $\mathcal{T}_c(\mu, \nu) + \int_{[0,1]} \nu \log \nu + \int_{[0,1]} V d\nu + \int_{[0,1]^2} \phi(y, z) d\nu(y) d\nu(z)$ où $f(\nu) = \int_{[0,1]} \nu \log \nu + \int_{[0,1]} V d\nu$

On note F_μ et F_ν les fonctions de répartition des mesures μ et ν par rapport la mesure de Lebesgue.

Proposition 14. *Si ν vérifie le problème d'optimisation précédent et γ est un transport optimal pour les mesures μ et ν alors γ est un équilibre de Cournot-Nash.*

La preuve de ce résultat se trouve dans [6].

Soient ψ, φ les fonctions duales de T définies dans la partie 4.

On a : $\forall x \in [0, 1], \psi(x) + \varphi(T(x)) = c(x, T(x))$.

En dérivant par rapport à y on obtient : $\varphi'(y) = D_y c(S(y), y)$
D'où $\varphi(y) = \varphi(0) + \int_0^y D_y c(S(t), t) dt$. On suppose pour simplifier que $\varphi(0) = 0$.
On peut calculer $D_y c(S(y), y) = S(y) - y = F_\mu^{-1}(F_\nu(y)) - y$.

ν minimise la quantité $\mathcal{T}_c(\mu, \nu) + \int_{[0,1]} \nu \log \nu + \int_{[0,1]} V d\nu + \int_{[0,1]^2} \phi(y, z) d\nu(y) d\nu(z)$.
Donc si on dérive cette quantité par rapport à ν , on doit avoir :
 $\varphi(y) + \log \nu + V + \int \phi(y, z) d\nu(z) = cste$, où $cste$ est une constante qui doit normaliser ν .

On a donc $\nu(y) = \exp(cste) \exp(-\varphi(y) - V(y) - \int \phi(y, z) d\nu(z))$.
Pour que ν soit une mesure de probabilité, on aura : $cste = \frac{1}{\exp(-\varphi(y) - V - \int \phi(x, y) d\nu(y))}$.
On pose $\tilde{C} = \exp(cste)$

ν est un point fixe de l'application $\nu \rightarrow C \exp(-\varphi(y) - V(y) - \int \phi(y, z) d\nu(z))$. On est donc amené à utiliser la méthode du point fixe, que l'on détaille dans le point suivant, pour déterminer ν .

4.7.2 Discrétisation du cas continu

Pour déterminer ν , on discrétise le problème en prenant $X = Y = \{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\}$ et on procède par la méthode itérative suivante :

- On initialise $\nu_0 = \mu$ et $\varphi_0 = 0$
- Pour tout $k \in \mathbb{N}^*$, on pose $\forall j \in \llbracket 0; N \rrbracket$

$$\nu_{k+1}(y_j) = \tilde{C} \exp(-\varphi_k(y_j) - V_j - \sum_{n=1}^N \phi_{n,j} \nu_k(y_n))$$

et

$$\varphi_{k+1}(y_j) = \sum_{n=0}^j D_y c(S(y_n), y_n) \nu_{k+1}(y_n) = \sum_{n=0}^j (y_n - S(y_n)) \nu_k(y_n)$$

On choisit $\epsilon > 0$ et on arrête les itérations lorsque $\|\nu_{k+1} - \nu_k\| \leq \epsilon$.

On admet alors que l'on a une version approchée d'une mesure ν satisfaisant le problème d'optimisation.

Pour simplifier le calcul de $S = F_\mu^{-1}(F_\nu)$ on prendra μ la mesure uniforme sur X : pour $j \in \llbracket 0; N \rrbracket$, $\mu(x_j) = \frac{1}{N+1}$. On a alors $F_\mu^{-1} = Id$.

Voici quelques simulations pour différents $\phi_{i,j}$ et V pour $N = 100$ et $\epsilon = 10^{-8}$:

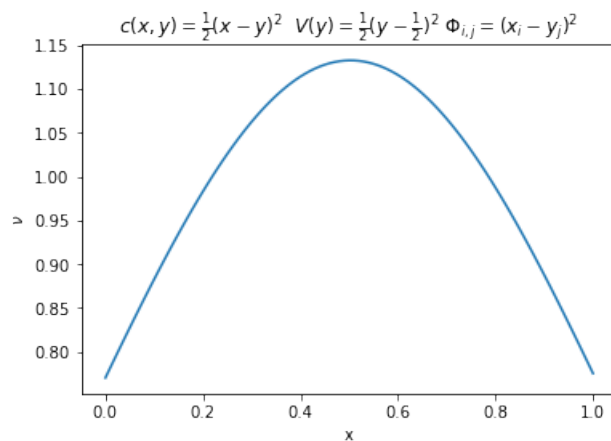


FIGURE 11 – Première simulation

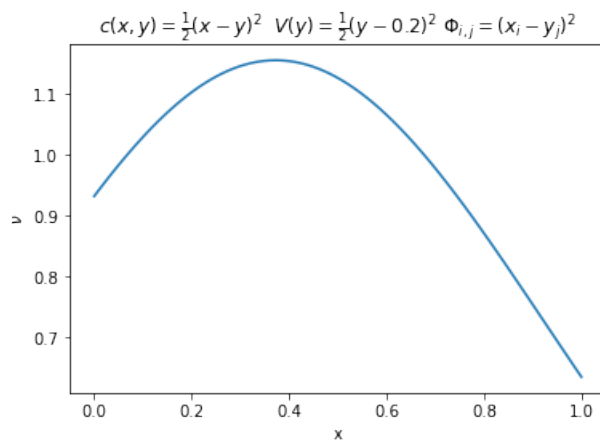


FIGURE 12 – Seconde simulation

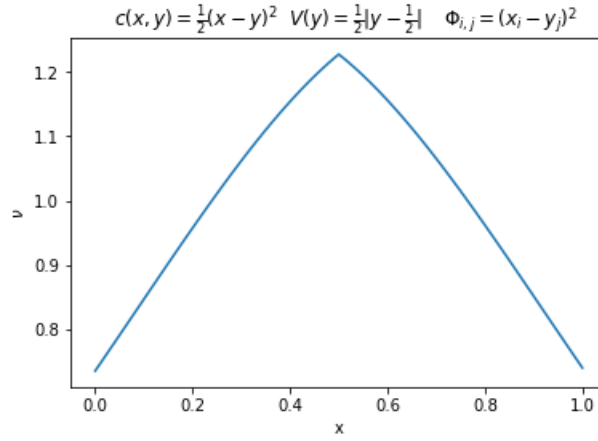


FIGURE 13 – Troisième simulation

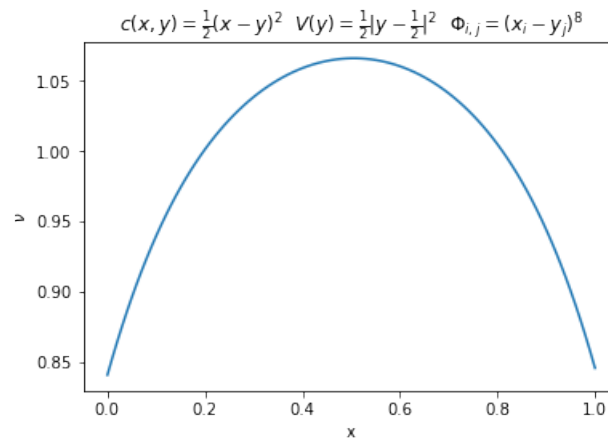


FIGURE 14 – Quatrième simulation

Références

- [1] Filippo Santambrogio, Optimal transport for applied mathematicians, Springer, 2015
- [2] G. Monge, Memoire sur la Théorie des Déblais et des Remblais, Histoire de l'Académie Royale des Sciences de Paris, 1781
- [3] Aldo Pratelli, On the equality between monge's infimum and kantorovich's minimum in optimal mass transportation, Annales de l'Institut Henri Poincaré (B) Probability and Statistics 43, no. 1, 1–13, 2007
- [4] L. V. kantorovitch, Dokl. Akad. Nauk SSSR 37, No. 7-8, 227-229 , 1942
- [5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, Gradient flows : in metric spaces and in the space of probability measures, Springer Science & Business Media, 2008
- [6] Cédric Villani, Topics in optimal transportation, no. 58, American Mathematical Soc., 2003
- [7] Giuseppe Buttazzo, Luigi De Pascale, and Paola Gori-Giorgi, Optimal-transport formulation of electronic density-functional theory, 2012
- [8] Codina Cotar, Gero Friesecke, and Claudia Klüppelberg, Density functional theory and optimal transportation with Coulomb cost, 2011

- [9] Adrien Blanchet, Guillaume Carlier, OPTIMAL TRANSPORT AND COURNOT-NASHEQUILIBRIA, 2012
- [10] Adrien Blanchet, Guillaume Carlier, Luca Nenna, Computation of Cournot-Nash equilibrium by entropic regularization, 2016