Étude d'approximations de problèmes de transport optimal et application à la physique

Rafaël Coyaud

UNIVERSITÉ PARIS-EST

École doctorale : MATHÉMATIQUES ET SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

Thèse de doctorat

Spécialité : Mathématiques

présentée par

Rafaël COYAUD

Study of approximations of optimal transport problems and application to physics

Thèse dirigée par Aurélien Alfonsi et Virginie Ehrlacher
et préparée au CERMICS, École des Ponts ParisTech

Aurélien Alfonsi	*Directeur de thèse*
Virginie Ehrlacher	*Directrice de thèse*
Gero Friesecke	*Rapporteur*
Mathieu Lewin	*Rapporteur*
Jean-David Benamou	*Président du jury*
Paola Gori-Giorgi	*Examinatrice*
Didier Henrion	*Examinateur*
Luca Nenna	*Examinateur*

And, to crown the Herakleitean dialectic, indeterminism, by means of particular stochastic functions, took on color and structure, giving rise to generous possibilities of organization.

— Iannis Xenakis, Formalized Music

Remerciements

Les plus de trois années de recherche qu'a durée cette thèse furent rythmées d'idées, d'essais, d'échecs, d'erreurs, de nombreuses rencontres et échanges, de quelques instants "Eurêka !", souvent suivis de " Ah . . . peut-être que . . . non . . . " ou autrement qui menèrent aux résultats présentés dans ce manuscrit, qui doivent énormément à toutes les personnes citées dans ces remerciements qui m'ont tour à tout formé, aidé, soutenu, inspiré, accompagné, guidé et surtout contribué à rendre inoubliables les dites années.

Je voudrais en premier lieu remercier chaleureusement mes directeurs de thèse, Virginie Ehrlacher et Aurélien Alfonsi sans qui tout ce travail n'aurait été possible et qui ont su calmement m'accompagner et me guider dans mon apprentissage de la recherche et la découverte d'un petit bout du merveilleux monde mathématique. Leur soutien et leur assurance ont été précieux pour naviguer ces plus de trois années de recherche.

Je voudrais aussi tout particulièrement remercier Éric Cancès qui m'a introduit au monde passionnant de la recherche en mathématiques appliquées à la chimie quantique, m'y a accompagné et conseillé puis fait connaître cette thèse tout en continuant de m'encadrer sur d'autres travaux.

Je voudrais aussi remercier avec beaucoup de gratitude mes rapporteurs Mathieu Lewin — pour ses cours, ses nombreuses remarques, corrections et échanges avant et pendant la thèse ainsi que sur ce manuscrit, et Gero Friesecke pour son rapport et ses remarques et échanges tout au long de mes recherches. Je voudrais aussi, remercier tous les membres de mon jury, Paola Gori-Giorgi, Luca Nenna, Didier Henrion et Jean-David Benamou, avec qui j'ai adoré échanger suite à la présentation de mes travaux ainsi que lors de différentes rencontres, à distance ou entre l'Italie, l'Autriche et le Canada.

Ainsi que les autres persones qui m'ont formé, encadré, et avec qui j'ai pu échanger sur les sujets de cette thèse ou d'autres, et qui rendent le monde de la recherche si plaisant, L. Ridgway Scott, Augusto Gerolin, Daniela Vögler, Tony Lelièvre, Gabriel Stoltz, Pierre Monmarché, Frédéric Meunier, Bernard Lapeyre, Antoine Levitt, Louis Garrigue.

Je dois aussi un grand "Merci !" au CERMICS et Isabelle Simunic où, et avec qui j'ai passé tant de bon moments, et où j'ai été formé et aidé avec bienveillance, ainsi que ses doctorants et docteurs, Ezéchiel, William, Sami, Adrien, Adrien, Adrien, Adel, Adèle, Lingling, Grégoire, Laura, Julien, Alexandre, Oumaima, Thomas, Thomas, Sébastien, Victor, Maël, Guillaume, Pierre-Loïk, Athmane, Boris, Étienne, Marion, Dylan, Clément, Sofiane, Sophian, Mouad, Inass, Zineb, Robert, Benoît, Mohamed, Cyrille, Olga, Rémi, Gaspard.

Enfin, je voudrais remercier mes parents pour avoir conntribué à me donner goût aux maths et aux sciences et pour, évidemment, les 27 dernières années, mes frères et Loulia, sans oublier ni le reste de ma famille, ni mes amis, Romain pour ta passion, rigueur, volonté et le Cuarenta y Tres, Marie, Charlotte et Bianca, chez chacune de qui des parties de cette thèse ont été découvertes ou rédigées, et bien sûr Marion, Léa, Armand, Louise, Rémi, Julia, Max, Tiphaine, Maximilien, Grégoire, Florence, Alexandre, Martin, Joséphine, Clément, Sarah, Laurent, Khalid, Franck, Florian, Claire, Claire, Aarohi, Pierre-Louis, Nicolas, Louis, Louise, Daphné, Rami, Benjamin, Paul, Rémi, Vincent, Gwen-Jiro, Victoire, Alexandre, Aey, Yuan, Ali, Violette, Alexandre, Gauthier, Édouard, Tom, Louise, Rémi, Clara, Vincent, Kata, Terence, Raja, Matthieu, Lucie, Bastien, Thomas, Paul, pour de nombreux moments magiques passés et à venir.

Titre : Étude d'approximations de problèmes de transport optimal et application à la physique

Résumé : Le transport optimal (TO) a de nombreuses applications; mais son approximation numérique est complexe en pratique. Nous étudions une relaxation du TO pour laquelle les contraintes marginales sont remplacées par des contraintes de moments (TOCM), et montrons la convergence de ce dernier vers le problème OT. Le théorème de Tchakaloff nous permet de montrer qu'un minimiseur du problème TOCM est une mesure discrète chargeant un nombre fini de points, qui, pour les problèmes multimarginaux, est linéaire en le nombre de marginales, ce qui permet de contourner le fléau de la dimension. Cette méthode est aussi adaptée aux problèmes de TO martingale. Dans certains cas importants en pratique, nous obtenons des vitesses de convergence en $O(1/N)$ ou $O(1/N^2)$, où N est le nombre de moments, ce qui illustre leur rôle.

Nous présentons un algorithme, basé sur un processus de Langevin sur-amorti contraint, pour résoudre le problème TOCM. Nous prouvons que tout minimiseur local du problème TOCM en est un minimiseur global. Et illustrons l'algorithme sur des exemples de larges problèmes TOCM symétriques.

Dans la seconde partie de la thèse, nous étendons une méthode (E. Cancès et L.R. Scott, SIAM J. Math. Anal., 50, 2018, 381–410) pour calculer un nombre arbitraire de termes dans la série asymptotique de l'interaction de van der Waals entre deux atomes d'hydrogène. Ces termes sont obtenus en résolvant un ensemble d'EDP de Slater–Kirkwood modifiées. La précision de cette méthode est montrée par des exemples numériques et une comparaison avec d'autres méthodes issues de la littérature. Nous montrons aussi que les états de diffusion de l'atome d'hydrogène ont une contribution majeure au coefficient C_6 de la série de van der Waals.

Mots-clefs : Transport Optimal, Transport Optimal multimarginal, Transport Optimal martingale, Processus de Langevin sur-amorti contraint, coefficients de dispersion de van der Waals, schéma de Galerkin.

Title: Study of approximations of optimal transport problems and application to physics

Abstract: Optimal Transport (OT) problems arise in numerous applications. Numerical approximation of these problems is a practical challenging issue. We investigate a relaxation of OT problems when marginal constraints are replaced by some moment constraints (MCOT problem), and show the convergence of the latter towards the former. Using Tchakaloff's theorem, we show that the MCOT problem is achieved by a finite discrete measure. For multimarginal OT problems, the number of points weighted by this measure scales linearly with the number of marginal laws, which allows to bypass the curse of dimension. This method is also relevant for Martingale OT problems. In some fundamental cases, we get rates of convergence in $O(1/N)$ or $O(1/N^2)$ where N is the number of moments, which illustrates the role of the moment functions.

We design a numerical method, built upon constrained overdamped Langevin processes, to solve MCOT problems; and proved that any local minimizer to the MCOT problem is a global one. We provide numerical examples for large symmetrical multimarginal MCOT problems.

We extend a method (E. Cancès and L.R. Scott, SIAM J. Math. Anal., 50, 2018, 381–410) to compute more terms in the asymptotic expansion of the van der Waals attraction between two hydrogen atoms. These terms are obtained by solving a set of modified Slater–Kirkwood PDE's. The accuracy of the method is demonstrated by numerical simulations and comparison with other methods from the literature. We also show that the scattering states of the hydrogen atom (the ones associated with the continuous spectrum of the Hamiltonian) have a major contribution to the C_6 coefficient of the van der Waals expansion.

Keywords: Optimal Transport, Multimarginal Optimal Transport, Martingale Optimal Transport, Constrained Overdamped Langevin process, van der Waals dispersion coefficients, Galerkin scheme.

Résumé substantiel:

Le travail de cette thèse se concentre sur deux problèmes rencontrés en chimie quantique, et plus spécifiquement pour des applications concernant les calculs de structure électronique des molécules.

Une première partie de ce travail concerne des résultats théoriques sur une méthode pour calculer la fonctionnelle de Levy-Lieb dans la limite des *électrons strictement corrélés* (SCE) en Théorie de la Fonctionnelle de Densité (DFT). Pour une densité électronique donnée, la limite SCE de la fonctionnelle de Levy-Lieb est un problème de transport optimal multimarges symétrique avec un coût de Coulomb, où le nombre de marginales est égal au nombre d'électrons dans le système, qui peut être très large dans les applications considérées. Une des contributions de cette thèse est l'étude théorique et numérique d'une méthode numérique pour la résolution de ce problème de transport optimal, qui consiste en la relaxation des contraintes marginales en un nombre fini de contraintes de moments. En particulier, nous prouvons que les minimiseurs de ce problème approché existent et que certains d'entre eux peuvent être écrits comme chargeant un nombre fini de points, qui croît linéairement avec le nombre de marginales. Ceci est exploité pour la conception d'algorithmes efficaces pour la résolution de ce problème approché et des résultats numériques illustrent la performance de l'algorithme proposé, qui utilise un processus de Langevin sur-amorti contraint. La méthode numérique proposée peut être utilisée pour résoudre d'autres types de problèmes de transport optimal multimarges ainsi que des problèmes de transport optimal martingale venant d'applications financières.

Une seconde contribution de cette thèse s'intéresse à une méthode de perturbations et un développement en série asymptotique afin de calculer la fonction d'onde électronique dans l'approximation de Born-Oppenheimer de deux atomes d'hydrogène à grande distance. Ce travail étend un article de É. Cancès et L.R. Scott [78] et fournit une méthode itérative pour calculer les coefficients de dispersion de l'interaction de van der Waals à un ordre arbitraire pour deux atomes d'hydrogène.

Transport Optimal La théorie du transport optimal a été d'abord formulée par Monge en 1781 dans [260]. Son intérêt a été croissant dans le seconde moitié du XX$^{\text{e}}$ siècle après l'introduction de sa formulation relaxée par Kantorovich dans [189] et sa résolution numérique par la programmatin linéaire par Dantzig [115, 116]. Depuis la fin du XX$^{\text{e}}$ siècle, des progrès ont été fait dans l'étude de ses propriétés mathématiques par Brenier [57, 59], Gangbo [149] et McCann [150, 151], et de ses connections avec l'équation de Monge-Ampère (voir Caffarelli dans [72, 73, 74]). Les travaux suivants, incluant ceux de Otto [187, 264], Caffarelli [71], Villani [318, 319], Ambrosio et Savaré [12] et Figalli [135, 137, 138] ont encore développé cette théorie.

Transport Optimal multimarges Soit $M \in \mathbb{N}^*$ et pour tout $1 \leq i \leq M$, soit $\mathcal{X}_i = \mathbb{R}^{d_i}$ avec $d_i \in \mathbb{N}^*$. Nous considérons M mesures de probabilité $\mu_1 \in \mathcal{P}(\mathcal{X}_1), ..., \mu_M \in \mathcal{P}(\mathcal{X}_M)$ et une fonction de coût semi-continue inférieurement $c : \mathcal{X}_1 \times ... \times \mathcal{X}_M \to \mathbb{R}_+ \cup \{\infty\}$.

Le problème de transport optimal multimarges est défini par

$$I^* = \inf_{\pi \in \Pi(\mu_1, ..., \mu_M)} \left\{ \int_{\mathcal{X}_1 \times ... \times \mathcal{X}_M} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ..., x_M) \right\}, \tag{1}$$

où

$$\Pi(\mu_1, ..., \mu_M) = \{\pi \in \mathcal{P}(\mathcal{X}_1 \times ... \times \mathcal{X}_M) \\ \text{t.q.} \, \forall 1 \leq i \leq M, \int_{\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_M} \mathrm{d}\pi = \mathrm{d}\mu_i \Big\}.$$

Un tel problème apparaît en chimie quantique, objet de l'application de cette thèse ainsi qu'en mécanique des fluides [43] et en science des données [232].

D'un point de vue théorique, ces problèmes ont été grandement étudiés par les mathématiciens [49, 161, 274], avec la caractérisation de mesures optimales [152, 258, 259, 270], si elles peuvent être de type Monge [195, 269] ou non [141]. Il y a aussi des études de ce problème (1) pour des coûts particuliers [158], dans le cas symétrique [162] ou son utilisation comme métrique [254].

Des extensions du problème (1) comme le transport optimal multimarges partiel [198], avec un nombre infini de marginales [271] ou sur une variété Riemanienne [196] ont aussi été étudiées, ainsi que ses connections avec des systèmes d'équations [208, 163], les couplages multi-agents [273] and et les effets de quantification [53].

Transport Optimal martingale Nous introduisons dans ce paragraphe le transport optimal martingale dans le cas avec deux marginales. Nous supposons que $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ avec $d \in \mathbb{N}^*$, et considérons deux mesures de probabilités $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ telles que

$$\int_{\mathbb{R}^d} |y| \mathrm{d}\nu(y) < \infty$$

et μ est plus petite que ν dans l'ordre convexe, i.e.

$$\int_{\mathbb{R}^d} \varphi(x) \mathrm{d}\mu(x) \leq \int_{\mathbb{R}^d} \varphi(y) \mathrm{d}\nu(y), \tag{2}$$

pour toute fonction convexe $\varphi : \mathbb{R}^d \to \mathbb{R}$ positive et intégrable par rapport à μ et ν. Cette dernière condition est équivalente, d'après le théorème de Strassen [311], à l'existence d'un coupage martingale entre μ et ν, i.e.

$$\exists \pi \in \Pi(\mu, \nu), \ \forall x \in \mathbb{R}^d, \ \int_{\mathbb{R}^d} y \mathrm{d}\pi(x, y) = x.$$

Le problème de transport optmal martingale consiste alors en la résolution du problème de minimisation

$$\inf_{\substack{\pi \in \Pi(\mu,\nu) \\ \forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} y \mathrm{d}\pi(x,y) = x}} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \mathrm{d}\pi(x, y) \right\}, \tag{3}$$

où $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ est une fonction de coût semi-continue inférieurement.

Ces problèmes sont difficiles à résoudre d'un point de vue numérique car leur discrétisation par les méthodes classiquement utilisées pour les problèmes de transport optimal à deux marginales ont une complexité exponentielle en le nombre de marginales.

Partie I : Étude de l'approximation de problèmes de transport optimal par des problèmes de transport avec contraintes de moments Une première contribution de cette thèse a été d'introduire une relaxation du problème de transport optimal multimarges (1) ou martingale (3) pour laquelle les contraintes marginales et martingales sont relaxées en des contraintes de moments. Plus précisément, pour chaque loi marginale μ_i, nous choisissons N *fonctions tests* $\phi_n^{(i)} : \mathcal{X}_i \to \mathbb{R}$ $(1 \leq i \leq M,\ 1 \leq n \leq N)$, le problème de transport optimal approché par des contraintes de moments s'écrit alors

$$I^N = \inf_{\pi \in \Pi(\mu_1,...,\mu_M;(\phi_n^{(1)})_{1\leq n\leq N},...,(\phi_n^{(M)})_{1\leq n\leq N})} \left\{ \int_{\mathcal{X}_1\times...\times\mathcal{X}_M} c(x_1,...,x_M)\mathrm{d}\pi(x_1,...,x_M) \right\}, \tag{4}$$

où

$$\Pi\left(\mu_1,...,\mu_M;(\phi_n^{(1)})_{1\leq n\leq N},\ldots,(\phi_n^{(M)})_{1\leq n\leq N}\right) = \left\{ \pi \in \mathcal{P}(\mathcal{X}_1 \times ... \times \mathcal{X}_M) \text{ t.q.} \forall 1 \leq i \leq M,\ \forall 1 \leq n \leq N, \int_{\mathcal{X}_1\times\cdots\times\mathcal{X}_M} \phi_n^{(i)}(x_i)\mathrm{d}\pi(x_1,\ldots,x_M) = \int_{\mathcal{X}_i} \phi_n^{(i)}\mathrm{d}\mu_i \right\}.$$

Nous prouvons le théorème suivant

Théorème 0.1. *Sous des hypothèses appropriées sur les fonctions tests et des contraintes techniques additionnelles, nous avons que*

$$I^N \underset{N\to+\infty}{\longrightarrow} I(\nu).$$

De plus, il existe au moins un minimiseur $\pi^N \in \mathcal{P}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_M)$ à (4) qui s'écrit

$$\pi^N = \sum_{k=1}^{K} w_k \delta_{(x_1^k,\ldots,x_M^k)}$$

pour un certain $1 \leq K \leq NM+2$, et certains $w_k \geq 0$ et $(x_1^k,\ldots,x_M^k) \in \mathcal{X}_1\times\cdots\times\mathcal{X}_M$ pour tout $1 \leq k \leq K$.

L'existence de ce minimiseur discret chargeant un faible nombre de points est intéressant d'un point de vue numérique car cela permet de concevoir une méthode numérique qui peut calculer une approximation du transport optimal multimarges avec un nombre de scalaires qui croît linéairement avec le nombre de lois marginales, ce qui casse le fléau de la dimension posé par les problèmes multimarges.

Nous établissons de plus qu'une telle relaxation s'applique aussi à la contrainte martingale du transport optimal multimarges martingale et avons dans ce cas des résultats de convergence analogues. Enfin, ce problème approché a un intérêt particulier en finance dans la mesure où les fonctions de moments permettent de prendre seulement en compte l'information disponible sur les mesures de probabilité considérées.

De plus, nous étudions aussi pour des classes particulières de fonctions tests, dans certains cas fondamentaux, la vitesse de convergence du problème approché (4) vers (1), illustrant l'influence du choix des fonctions tests sur l'approximation.

Calculs de structure électronique de molécules Dans l'approximation de Born-Oppenheimer, une molécule est un système composé de

- $M \in \mathbb{N}^*$ noyaux qui sont connsidérés comme des particules ponctuelles classiques et dont les positions sont notées $R_1, \cdots, R_M \in \mathbb{R}^3$ et leur charge électrique $Z_1, \cdots, Z_M \in \mathbb{N}^*$;
- N électrons qui sont modélisés comme des particules quantiques et dont l'état est décrit par une fonction

$$\psi : \left\{ \begin{array}{ccc} \mathbb{R}^{3N} & \to & \mathbb{C} \\ (x_1, \ldots, x_N) & \mapsto & \psi(x_1, \ldots, x_N), \end{array} \right.$$

  appelée *fonction d'onde* du système d'électrons.

Afin d'alléger les notations, nous omettons les variables de spin. En effet, dans les deux parties de cette thèse, ou bien la dépendance en le spin peut être séparé de celle en les positions (seconde contribution), ou bien elle disparaît dans la limite semiclassique considérée (troisième contribution).

L'interprétation physique d'une fonction d'onde ψ est la suivante: étant donné un ensemble $A \subset \mathbb{R}^{3N}$, $\int_A |\psi|^2$ représnete la probabilité que les positions des N électrons appartiennent à l'ensemble A. En particulier, ceci implique que $\|\psi\|^2_{L^2(\mathbb{R}^{3N})} = 1$. De plus, la fonction d'onde ψ est *antisymétrique* par rapport à ses variables. C'est une conséquence du fait que les électrons sont des fermions. Plus précisément, en notant $\mathcal{S}_N$ l'ensemble des permutations de l'ensemble $\{1, \ldots, N\}$, nous avons que pour tout $p \in \mathcal{S}_N$ et tout $(x_1, \ldots, x_N) \in \mathbb{R}^{3N}$,

$$\psi(x_{p(1)}, \ldots, x_{p(N)}) = \epsilon(p)\psi(x_1, \ldots, x_N),$$

où $\epsilon(p)$ est la signature de p.

L'énergie $E[\psi]$ d'un système de N électrons dont l'état est décrit par une fonction d'onde ψ dans la molécule décrite ci-dessus est la somme de trois contributions :

- l'énergie cinétique:

$$T[\psi] := \frac{1}{2} \int_{\mathbb{R}^{3N}} |\nabla \psi|^2;$$

- l'énergie de Coulomb associée aux interactions entre les électrons et les noyaux :

$$C_{\text{nuc}}[\psi] := \int_{\mathbb{R}^{3N}} \left(\sum_{i=1}^{N} V_{\text{nuc}}(x_i) \right) |\psi(x_1, \ldots, x_N)|^2 \, dx_1 \cdots dx_N,$$

  where, for all $x \in \mathbb{R}^3$,

$$V_{\text{nuc}}(x) := -\sum_{k=1}^{M} \frac{Z_k}{|x - R_k|};$$

- l'énergie de Coulomb associée aux interactions entre les électrons :

$$C_{\text{elec}}[\psi] := \int_{\mathbb{R}^{3N}} c(x_1, \ldots, x_N) |\psi(x_1, \ldots, x_N)|^2 \, dx_1 \ldots \, dx_N,$$

  où pour presque tout $(x_1, \ldots, x_N) \in \mathbb{R}^{3N}$,

$$c(x_1, \ldots, x_N) = \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}.$$

Calculer l'état fondamental des électrons dans la molécule revient à calculer la fonction d'onde ψ_0 qui minimise l'énergie du système parmi toutes les fonctions d'onde admissibles. Plus précisément, notons

$$\mathcal{A} := \left\{\psi \in L^2(\mathbb{R}^{3N}),\ \nabla\psi \in L^2(\mathbb{R}^{3N})^{3N},\ \psi \text{ antisymmetric},\ \|\psi\|_{L^2(\mathbb{R}^{3N})} = 1\right\}$$

l'ensemble de fonctions d'ondes associées à un système de N electrons avec une énergie cinétique finie. Alors, nous avons que

$$U(R_1, \cdots, R_M) = \min_{\psi\in\mathcal{A}} T[\psi] + C_{\text{nuc}}[\psi] + C_{\text{elec}}[\psi], \tag{5}$$

où nous avons signalé la dépendance de la valeur de cet infimum en la position des noyaux de la molécule $R_1, \cdots, R_M$. Soit $H := -\frac{1}{2}\Delta + \sum_{i=1}^N V_{\text{nuc}}(x_i) + c(x_1, \cdots, x_N)$ que l'on appelle *opérateur de Schrödinger à plusieurs corps*. L'opérateur H est auto-adjoint, borné inférieurement, et opérateur sur

$$L^2_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in L^2(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}$$

avec pour domaine

$$H^2_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in H^2(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}.$$

Notons aussi

$$H^1_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in H^1(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}.$$

Dans le cas où $U(R_1, \cdots, R_M) := \inf\sigma(H)$ est une valeur propre discrète de H (ce qui arrive par exemple quand la molécule est neutre ou chargée positivement d'après le théorème de Zhislin [326]), il existe au moins un minimiseur ψ_0 à (5), et tout minimiseur est nécessairement un vecteur propre de H associé à la valeur propre $U(R_1, \cdots, R_M)$. Ainsi, résoudre le problème de Schrödinger électronique revient à résoudre un problème aux valeurs propres linéaire *de grande dimension* de la forme

$$H\psi_0 = U(R_1, \cdots, R_M)\psi_0. \tag{6}$$

Seconde contribution de cette thèse : interactions de van der Waals entre deux atomes d'hydrogène Bien que pour de grandes valeurs de N et M des approximations et des méthodes numériques sont nécessaires pour évaluer $U(R_1, \cdots, R_M)$, pour de petits systèmes, des techniques analytiques peuvent permettre de résoudre l'équation de Schrödinger.

C'est le cas lorsque l'on considère les intéractions électroniques en entre deux atomes hydrogène à grande distance. Ces interactions sont appelées interactions de van der Waals, sont attractives et jouent un rôle important dans les systèmes en phase condensée tels que les molécules biologiques [21, 288] ou les matériaux 2D [153]. Étudiées depuis 1873 [317], elles ont d'abord été comprises mathématiquement par London [238]. Dans le cas de deux atomes d'hydrogène, Slater et Kirkwood [309] ont amené une équation aux dérivées partielles qui permet de calculer les coefficients de dispersion de l'énergie dans la limite des grandes distances (qui décroît en $-C_6/R^6$ au premier ordre, où R est la distance entre les noyaux). Cancès et Scott dans [78] ont modifié leur technique et ont prouvé que le problème qu'ils ont proposé est bien posé et à l'aide d'une approximation de Galerkin ont calculé le coefficient C_6.

Une extension de la technique de Cancès et Scott a été étudiée pendant cette thèse afin de calculer les coefficients de dispersion de van der Waals a n'importe

quel ordre. Cette technique repose sur une méthode de perturbation afin d'analyser le développement en série asymptotique de l'attraction de van der Waals ainsi que sur une séparation des interactions entre une partie radiale et une partie angulaire ce qui ramène le problème original en six dimensions à des équations aux dérivées partielles en deux dimensions. Les coefficients de dispersion peuvent enfin être calculés récursivement par des approximations de Galerkin; les valeurs calculées avec cette méthode sont en accord avec celles de [204, 265] pour lesquelles les auteurs ont utilisé d'autres techniques.

Theorie de la fonctionnelle de la densité La grande dimensionalité de l'équation (6) la rend difficile à résoudre d'un point de vue numérique par des méthodes standard dans le cas où N est grand, en particulier pour des systèmes d'électrons fortement corrélés où les interactions coulombiennes entre les noyaux jouent un rôle important.

Le principe de la Théorie de la Fonctionnelle de la Densité (DFT), et de tous les modèles qui en sont dérivés est une reformulation du problème (5) où la densité (et non plus la fonction d'onde) est la variable principale. Le principal avantage de cette méthode est que le problème est maintenant formulé sur le domaine $\mathbb{R}^3$ plutôt que $\mathbb{R}^{3N}$.

La justification théorique des modèles de DFT a été introduite par Hohenberg et Kohn [182], puis par Levy [225] et complétée par Lieb [229]. En effet, le théorème de Hohenberg-Kohn [182] dit que l'énergie de la densité électronique de l'état fondamental du problème électronique (5) peut être trouvée en résolvant un problème de la forme

$$U(R_1, \cdots, R_M) = \inf \left\{ F(\rho) + \int_{\mathbb{R}^3} \rho V,\ \rho \in L^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \rho = N \right\},$$

où F est une fonctionnelle de la densité électronique ρ. Plus précisément, la DFT repose sur le calcul suivant [182, 229]:

$$\begin{aligned} U(R_1, \cdots, R_M) &= \inf \{ \langle \psi_e, H_V \psi_e \rangle,\ \psi_e \in \mathcal{A} \} \\ &= \inf \left\{ \inf \{ \langle \psi_e, H_1 \psi_e \rangle,\ \psi_e \in \mathcal{A},\ \rho_{\psi_e} = \rho \} + \int_{\mathbb{R}^3} \rho V,\ \rho \in \mathcal{I}_N \right\} \\ &= \inf \left\{ F_{LL}(\rho) + \int_{\mathbb{R}^3} \rho V,\ \rho \in \mathcal{I}_N \right\}, \end{aligned}$$

où $\mathcal{I}_N$ est l'ensemble des densités électroniques associées à des fonctions d'onde admissible et qui peut s'écrire [229]

$$\mathcal{I}_N = \left\{ \rho \geq 0,\ \sqrt{\rho} \in H^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \rho = N \right\}.$$

et où

$$F_{LL}(\rho) := \inf \{ \langle \psi_e, H_1 \psi_e \rangle,\ \psi_e \in \mathcal{A},\ \rho_{\psi_e} = \rho \}$$

est appelée la *fonctionnelle de Levy-Lieb*. Cette fonctionnelle est universelle au sens où elle ne dépend pas du système moléculaire étudié (qui n'intervient qu'à travers le potentiel V et le nombre d'électrons N). De façon équivalente, la fonctionnelle de Levy-Lieb peut être réécrite

$$F_{LL}(\rho) := \{ T[\psi] + C_{\text{elec}}[\psi],\ \psi \in \mathcal{A},\ \rho_\psi = \rho \}.$$

Cette théorie est attrayante, toutefois en pratique, le calcul exact de $F_{LL}(\rho)$ est hors de portée dans la mesure où il nécessite la résolution d'un problème aussi complexe que le problème de Schrödinger électronique original.

Limite semi-classique de la fonctionnelle de Levy-Lieb L'une des approximations, suggérée par des chimistes théoriques dans [302, 304], consiste à regarder la limite *semi-classique* où celle des *électrons strictement corrélés* (SCE) de la fonctionnelle de Levy-Lieb, afin de l'utiliser pour concevoir des modèles approchant la DFT pour les systèmes fortement corrélés. Cette limite semi-classique est la limite lorsque α tend vers 0 de la fonctionnelle F_{LL}^{α} définie comme suit pour tout $\rho \in \mathcal{I}_N$ et $0 < \alpha \leq 1$:

$$F_{LL}^{\alpha}(\rho) := \{\alpha T[\psi] + C_{\text{elec}}[\psi],\ \psi \in \mathcal{A},\ \rho_\psi = \rho\}\,.$$

Dans cette limite semi-classique, l'influence du terme d'énergie cinétique $T[\psi]$ est alors négligé par rapport aux contributions dues au terme d'interactions coulombiennes électron-électron $C_{\text{elec}}[\psi]$. Il a été rigoureusement prouvé par une série de travaux [106, 107, 226] que la limite lorsque α tend vers 0 de la fonctionnelle $F_{LL}^{\alpha}(\rho)$ s'écrit comme un *problème de transport optimal multimarges avec coût de Coulomb*. Plus précisément, pour tout $\rho \in \mathcal{I}_N$, notons ν_ρ la mesure de probabilité sur $\mathbb{R}^3$ définie par $d\nu_\rho(x) := \frac{\rho(x)}{N}\,dx$ et $\mathcal{P}_{\text{sym}}(\mathbb{R}^{3N})$ l'ensemble des mesures de probabilité symétriques sur $\mathbb{R}^{3N}$. Pour tout $\gamma \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{3N})$, notons μ_γ la mesure de probabilité sur $\mathbb{R}^3$ définie comme la marginale de γ, i.e.

$$d\mu_\gamma(x) := \int_{(x_2,\ldots,x_N)\in\mathbb{R}^{3(N-1)}} d\gamma(x, x_2, \ldots, x_N).$$

Des travaux de Buttazzo, De Pascale et Gori Giorgi [67], Cotar, Friesecke et Klüppelberg [106] pour des preuves pour $N = 2$ et avec Mendl et Pass [142], Bindini et De Pascale [50] étendus par Lewin [226] pour $N \geq 2$ dans le cas des états mixes fermioniques et Cotar, Friesecke et Klüppelberg [107] pour $N \geq 2$, ont prouvé, en utilisant un lissage approprié des plans de transport que dans la limite semi-classique

$$\lim_{\alpha\to 0} F_{LL}^{\alpha}(\rho) = I(\nu_\rho),$$

où pour toute mesure de probabilité ν sur $\mathbb{R}^3$,

$$I(\nu) := \inf_{\substack{\gamma \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{3N}),\\ \mu_\gamma = \nu}} \int_{\mathbb{R}^{3N}} c\,d\gamma. \tag{7}$$

Troisième contribution de cette thèse : Développement d'un nouvel algorithme numérique pour des problèmes de transport optimal multimarges symétriques Une troisième contribution de cette thèse est de proposer et analyser d'un point de vue mathématique une nouvelle méthode pour approcher le problème de transport optimal multimarges symétrique (7). Dans cette approche, nous considérons toujours l'espace d'états continus $\mathbb{R}^3$, mais les contraintes marginales apparaissant dans (7) sont relaxées en un nombre fini de contraintes de moments. C'est un cadre dans lequel les résultats introduits en première contribution s'appliquent et peuvent être symétrisés, permettant des gains de complexités additionnels.

Par simplicité, nous présentons nos résultats ci-après dans le cas où le support de la mesure ν est inclus dans un ensemble compact $Y \subset \mathbb{R}^3$. Soit $(f_m)_{m\in\mathbb{N}^*} \subset \mathcal{C}(Y)$, vérifiant l'hypothèse de densité naturelle suivante

$$\forall f \in \mathcal{C}(Y), \qquad \inf_{g_M \in \text{Span}\{f_1,\ldots,f_M\}} \|f - g_M\|_{L^\infty} \underset{M\to+\infty}{\longrightarrow} 0,$$

et considérons le problème de transport optimal approché avec contraintes de moments

$$I^M(\nu) := \inf_{\substack{\gamma \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{3N}),\\ \forall 1 \le m \le M,\\ \int_{\mathbb{R}^{3N}} \left(\frac{1}{N} \sum_{i=1}^N f_m(x_i) \right) d\gamma(x_1, \dots, x_N) = \int_{\mathbb{R}^3} f_m \, d\nu}} \int_{\mathbb{R}^{3N}} c \, d\gamma. \quad (8)$$

Nous avons prouvé le théorème suivant, où $\mathcal{P}(\mathbb{R}^{3N})$ est l'ensemble des mesures de probabilités sur $\mathbb{R}^{3N}$ (non nécessairement symétriques).

Théorème 0.2. *Sous les hypothèses précédentes, nous avons que*

$$I^M(\nu) \underset{M \to +\infty}{\longrightarrow} I(\nu).$$

De plus,

$$I^M(\nu) = \inf_{\substack{\gamma \in \mathcal{P}(\mathbb{R}^{3N}),\\ \forall 1 \le m \le M,\\ \int_{\mathbb{R}^{3N}} \left(\frac{1}{N} \sum_{i=1}^N f_m(x_i) \right) d\gamma(x_1, \dots, x_N) = \int_{\mathbb{R}^3} f_m \, d\nu}} \int_{\mathbb{R}^{3N}} c \, d\gamma, \quad (9)$$

et il existe au moins un minimiseur $\gamma^M \in \mathcal{P}(\mathbb{R}^{3N})$ à (9) qui s'écrit

$$\gamma^M = \sum_{k=1}^K w_k \delta_{(x_1^k, \dots, x_N^k)}$$

pour un certain $1 \le K \le M+2$, et certains $w_k \ge 0$ et $(x_1^k, \dots, x_N^k) \in Y^N$ pour tout $1 \le k \le K$. De plus,

$$\gamma^M_{\mathrm{sym}} = \frac{1}{N!} \sum_{p \in \mathcal{S}_N} \sum_{k=1}^K w_k \delta_{(x^k_{p(1)}, \dots, x^k_{p(N)})},$$

la version symétrisée de γ^M, est un minimiseur de (8).

Le théorème 0.2 établit deux choses : (i) il est possible de retirer la contrainte de symétricité de la mesure γ dans le problème (8) pour calculer $I^M(\nu)$; (ii) il existe un minimiseur de (9) qui s'écrit comme une mesure discrète chargeant un faible nombre de points (moins de $M+2$), et un minimiseur à (8) peut être obtenu comme le symétrisé de cette mesure discrète. En particulier, ceci signifie qu'il est suffisant d'identifier au plus $\mathcal{O}(NM)$ scalaires pour calculer γ^M. Ceci suggère que considérer le problème d'optimisation suivant pour le calcul de $I^M(\nu)$, puisque

$$I^M(\nu) = \min_{\substack{(w_k)_{1 \le k \le M+2} \in \mathbb{R}_+^{M+2},\\ \sum_{k=1}^{M+2} w_k = 1,\\ \left(x_1^k, \dots, x_N^k\right) \in Y^N, \ \forall 1 \le k \le M+2,\\ \sum_{k=1}^{M+2} w_k \left(\frac{1}{N} \sum_{i=1}^N f_m(x_i^k) \right) = \int_{\mathbb{R}^3} f_m \, d\nu}} \sum_{k=1}^{M+2} w_k c(x_1^k, \dots, x_N^k). \quad (10)$$

L'utilisation de cette structure parcimonieuse pour la conception de méthode numériques pour la résolution de (8) a été l'objet de cette thèse. Nous prouvons en particulier que tout minmiseur local de (10) en est un minimseur *global* . De plus, la méthode numérique proposée pour la résolution de ce problème utilise un processus de Langevin sur-amorti contraint, et cela permet de résoudre un problème de transport optimal multimarges symétrique ayant 100 marginales, ce qui est plus large que l'état de l'art pour ce type de problèmes.

Contents

Chapter 1

Introduction

1.1 Introduction

The work of this thesis focuses on two mathematical problems arising from quantum chemistry, and more specifically from applications concerning electronic structure calculations of molecules.

A first part of this work concerns some theoretical results about a numerical method for computing the so-called *strictly correlated electrons* (SCE) limit of the so-called Levy-Lieb functional in Density Functional Theory (DFT). For a given electronic density, the SCE limit of the Levy-Lieb functional gives rise to a symmetric multi-marginal optimal transport problem with Coulomb cost, where the number of marginal laws is equal to the number of electrons in the system which can be very large in relevant applications. One contribution of this thesis is the theoretical and numerical study of a numerical method for the resolution of this optimal transport problem, which consists in relaxing the marginal constraints into a finite number of moment constraints. In particular, it is proved that minimizers to the resulting approximate optimization problem exist and that some of them can be written as discrete measures charging a low number of points, which scales linearly with the number of electrons. This can be exploited for the design of efficient algorithms for the resolution of such approximate problems. Numerical results illustrate the performance of the proposed numerical method, which makes use of constrained overdamped Langevin dynamics. The proposed numerical method can be used for the resolution of other types of multi-marginal optimal transport problems, including problems with martingale constraints arising from finance applications.

A second contribution of the thesis focuses on a perturbation method and asymptotic expansion to compute the electronic wavefunction in the Born-Oppenheimer approximation of two hydrogen atoms at large distance. This work extends an article by E. Cancès and R. L. Scott [78], and provides an iterative method to compute van der Waals dispersion coefficients up to an arbitrary order for two hydrogen atoms.

This introductory chapter is structured as follows: an introduction on optimal transport theory, its applications and existing numerical methods in the general case is given in Section 1.2. Section 1.3 presents an introduction on electronic structure calculation problems for molecules in quantum chemistry, in particular for the electronic Schrödinger many-problem together with its link to van der Waals interactions, and the Density Functional Theory. In this section are also presented the links between optimal transport problems and the SCE limit of the so-called Levy-Lieb functional, together with the main contributions of this thesis.

1.2 Introduction to optimal transportation

Optimal transport theory has been first formulated by Monge in 1781 in [260]. Its interest has been increasing in the second half of XX^{th} century after the introduction of its relaxed formulation by Kantorovich in [189] and its numerical solution by linear programming by Dantzig [115, 116]. From the end of the XX^{th} century, progress was made in the study of its mathematical properties by Brenier [57, 59], Gangbo [149] and McCann [150, 151], and of its connection with Monge-Ampère equation (see Caffarelli in [72, 73, 74]). Latter works including the one of Otto [187, 264], Caffarelli [71], Villani [318, 319], Ambrosio and Savaré [12] and Figalli [12, 137, 138] developped this theory further.

Optimal transport has a wide range of applications, and we refer the reader to

the review articles [283, 284]. Different points of views can be cast on this family of problems, depending on the types of applications one considers.

First, a "transport" point of view allows computing given a starting distribution and an end distribution, a way of transporting the first one on the last one which is optimal relative to some physical cost or constraints. This can be leveraged in the planning of cities [69, 80] and urban network [68], the study of crowd motion [65, 216, 293] as well as in fluid mechanics [39, 43, 56, 58, 60, 61] or propagation through porous media [87] and even for the reconstruction of initial conditions of the universe [62, 144].

A second point of view on optimal transport is to consider this theory as a mean to compute metrics between two probability distributions (for instance the Wasserstein distance). This approach is used in computer vision and image analysis [23, 131, 176, 203, 241, 242, 282, 289, 294, 298], signal analysis [202, 315], shape matching or reconstruction [117, 132, 239, 312] and data science and machine learning (for instance for linguistics or Wasserstein Generative Adversarial Networks) [20, 155, 183, 277, 295].

A third field of applications is economics, which often makes use of the dual formulation of optimal transport as an equilibrium state that maximizes the interest of two (or more) actors [81, 82, 94], and econometrics [146, 147].

Section 1.2.1 contains an introduction to the mathematical properties of two-marginal optimal transport problems. In this thesis, we will focus more specifically on applications stemming from quantum chemistry and finance, where multimarginal optimal transport and martingale optimal transport naturally arise. Hence, Section 1.2.2 will be devoted to the presentation of such problems, together with the existing numerical methods used to compute a numerical approximation of their solutions.

1.2.1 Two-marginal optimal transport problem

We begin by presenting definitions and mathematical properties related to the study of two-marginal optimal transport problems, following results from [11, 318, 319, 291].

1.2.1.1 Monge and Kantorovich formulation

For any Polish space $\mathcal{Z}$, (i.e. complete and separable metric space), let us denote by $\mathcal{P}(\mathcal{Z})$ the set of probability measures on $\mathcal{Z}$, and by $C_b(\mathcal{Z})$ the set of continuous bounded functions on $\mathcal{Z}$.

Let us consider two probability measures $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$, where $\mathcal{X}$ and $\mathcal{Y}$ are Polish spaces.

Definition 1.1 (Push-forward measure)**.** *Let $T : \mathcal{X} \to \mathcal{Y}$ be a measurable map. Then, the push-forward measure of μ by T (denoted $T\#\mu$) in $\mathcal{P}(\mathcal{Y})$ is defined by*

$$T\#\mu(A) = \mu(T^{-1}(A)) \qquad \text{for every measurable set } A \subset \mathcal{Y}. \tag{1.1}$$

In addition, let $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ be a lower semi-continuous function. The function c will be called hereafter a *cost function*. For a given cost function c, the Monge formulation of optimal transport reads as

$$\text{find } T^* : \mathcal{X} \to \mathcal{Y} \text{ a measurable map s.t. } \quad T^* \in \underset{\substack{T:\mathcal{X}\to\mathcal{Y} \text{ measurable} \\ T\#\mu=\nu}}{\arg\min} \int_{\mathcal{X}} c(x, T(x)) d\mu(x). \tag{1.2}$$

This problem can be interpreted as follows: the map T^* encodes a displacement of some mass distributed according to the probability measure μ to match the distribution of mass given by ν in such a way that it generates the lowest displacement cost relative to c.

In general, Problem 1.2 is not well-posed as there may not exist any minimizer T^*, depending on the choice of c, μ and ν.

Let us denote the set of couplings measures on $\mathcal{X} \times \mathcal{Y}$ between μ and ν by

$$\Pi(\mu, \nu) := \left\{ \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \mid \int_{\mathcal{X}} d\pi = d\nu, \int_{\mathcal{Y}} d\pi = d\mu \right\}. \tag{1.3}$$

The Kantorovich formulation of optimal transport then reads as

$$\text{find } \pi^* \in\in \Pi(\mu, \nu) \text{ s.t.} \quad \pi^* \in \underset{\pi \in \Pi(\mu,\nu)}{\arg\min} \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) d\pi(x, y). \tag{1.4}$$

The measure π^* can be seen as a coupling measure between μ and ν which encodes correlations between two random variables whose marginal laws are given respectively by μ and ν which, integrated against c, have the lowest cost.

Theorem 1.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Polish spaces, $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$, $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ be a lower semi-continuous cost function, then* (1.4) *admits at least one solution.*

Note that in some cases, solutions to (1.4) can be obtained from solutions to (1.2). If there exists $T : \mathcal{X} \to \mathcal{Y}$ a measurable map such that $d\pi^*(x, y) = d\mu(x) \otimes \delta_{T(x)}(y)$ where π^* is a solution to (1.4), we say that π^* is a *Monge minimizer* to (1.4).

In the case when $c(x, y) = h(x - y)$ with h a strictly convex function, μ and ν are probability measures on a compact subset $\Omega \subset \mathbb{R}^d$, μ is absolutely continuous and $\partial\Omega$ is negligible, then the minimizer of (1.4) is unique and of Monge form.

1.2.1.2 Wasserstein distance

Optimal transport is a natural way to define metrics between probability measures, such as the Wasserstein distance introduced below.

Definition 1.2 (Wasserstein distance)**.** *Let $1 < p < +\infty$. The* Wasserstein *distance to the power p between μ and ν is defined as*

$$W_p(\mu, \nu) = \left(\min_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} |x - y|^p d\pi(x, y) \right)^{1/p}.$$

For any $\phi \in C_b(\mathcal{X})$ and $\psi \in C_b(\mathcal{Y})$, let us denote by

$$\phi \oplus \psi : \left\{ \begin{array}{ccc} \mathcal{X} \times \mathcal{Y} & \to & \mathbb{R}, \\ (x, y) & \mapsto & \phi(x) + \psi(y). \end{array} \right.$$

1.2.1.3 Dual formulation

The dual formulation of the Kantorovich formulation (1.4) reads as

$$\text{find } (\phi^*, \psi^*) \in C_b(\mathcal{X}) \times C_b(\mathcal{Y}) \text{ s.t.} \quad (\phi^*, \psi^*) \in \underset{\substack{\phi \in C_b(\mathcal{X}), \psi \in C_b(\mathcal{Y}) \\ \phi \oplus \psi \leq c}}{\arg\max} \int_{\mathcal{X}} \phi \, d\mu + \int_{\mathcal{Y}} \psi d\nu. \tag{1.5}$$

Any solution $(\phi^*, \psi^*) \in C_b(\mathcal{X}) \times C_b(\mathcal{Y})$ to (1.4) is called a set of *Kantorovich potentials.*

The following theorem states some theoretical properties of the dual problem (1.4).

Theorem 1.2 (From [291, Theorem 1.39 and 1.42]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be Polish spaces.*

- *If $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded from below, then there exists a supremum value to (1.5); besides, it is equal to the minimal value of (1.4).*

- *If $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is uniformly continuous and bounded, then the optimal value of (1.5) is equal to the optimal value of (1.4); besides, there exists maximizers to (1.5). Moreover, for any π^* solution of (1.4) and any (ϕ^*, ψ^*) solution to (1.5), it necessarily holds that*

$$\phi^*(x) + \psi^*(y) = c(x, y) \qquad \pi^*\text{-a.e. on } \mathcal{X} \times \mathcal{Y} \tag{1.6}$$

Let us mention here that numerous works are related to regularity properties of transport maps [26, 194, 235], Kantorovich potentials [240, 281], c-cyclical monotonicity [89], optimality [280] and duality [30, 31, 32, 191, 281]

1.2.1.4 Extensions of the two-marginal optimal transport problem

Let us briefly mention here some works on extensions or particular aspects of the two-marginal optimal transport problems introduced in the preceding sections: optimal transport theory on Riemanian or non-compact manifolds [129, 139, 250]; study of particular cost functions (e.g. the determinant [85] or repulsive costs [104, 123]); problems with unequal dimensions [263]; unbalanced optimal transport [54, 96, 97]; partial transport problem [134]; dynamical formulation [186, 212, 213]; transport with obstacle problem [75].

Second, studies involving extensions of the Wasserstein distance [79, 90] or in the Wasserstein space [249, 292] such as the problem of finding barycenters in the Wasserstein space [2].

Let us also mention studies involving relaxations of optimal transport such as the entropic relaxation and some statistical properties [95, 124, 133, 136, 199, 201, 231] or its use to approximate Wasserstein gradient flows [276] or Schrödinger problem [248].

Last, let us mention the link between optimal transport and PDE's [127] and analytical problems such as the Schrödinger problem [221, 222], Cournot-Nash equilibria [51].

1.2.2 Multimarginal and martingale optimal transport

In this section are introduced multimarginal and martingale optimal transport and optimal transport problems, which are the main focus of this thesis.

Multimarginal optimal transport Let $M \in \mathbb{N}^*$ (where $\mathbb{N}^*$ denotes the set of positive integers $\{1, 2, 3, \ldots\}$) and for all $1 \leq i \leq M$, let $\mathcal{X}_i = \mathbb{R}^{d_i}$ with $d_i \in \mathbb{N}^*$. We consider M probability measures $\mu_1 \in \mathcal{P}(\mathcal{X}_1), ..., \mu_M \in \mathcal{P}(\mathcal{X}_M)$ and a lower semi continuous cost function $c : \mathcal{X}_1 \times ... \times \mathcal{X}_M \to \mathbb{R}_+ \cup \{\infty\}$.

The multimarginal optimal transport problem is defined as follows

$$\inf_{\pi \in \Pi(\mu_1, ..., \mu_M)} \left\{ \int_{\mathcal{X}_1 \times ... \times \mathcal{X}_M} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ..., x_M) \right\}, \tag{1.7}$$

where

$$\Pi(\mu_1,...,\mu_M) = \{\pi \in \mathcal{P}(\mathcal{X}_1 \times ... \times \mathcal{X}_M) \\ \text{s.t.} \, \forall 1 \leq i \leq M, \, \int_{\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_M} \mathrm{d}\pi = \mathrm{d}\mu_i \Big\}.$$

Such multi-marginal optimal transport problems arise in quantum chemistry applications which will be detailed in the next section. Let us mention here that it also appears in fluid mechanics [43] and data science [232].

From a theoretical point of view, such problems have received a lot of attention from mathematicians [49, 161, 274], with some characterization of its optimal measures [152, 258, 259, 270], whether they can be of Monge form [195, 269] or not [141]. There are also studies of problem (1.7) for some particular costs [158], the symmetric case [162] or its use as a metric [254].

Extensions of problem (1.7) such as partial multimarginal optimal transport [198], with an infinite number of marginal laws [271] or on a Riemanian manifold [196] have also been studied, as well as connections with systems of equations [208, 163] or multi-agent matching [273] and quantization effects [53].

Martingale optimal transport In this paragraph, we introduce martingale optimal transport in the two marginal case. Let us assume that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ with $d \in \mathbb{N}^*$, and consider two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} |y| \mathrm{d}\nu(y) < \infty$$

and μ is lower than ν for the convex order, i.e.

$$\int_{\mathbb{R}^d} \varphi(x) \mathrm{d}\mu(x) \leq \int_{\mathbb{R}^d} \varphi(y) \mathrm{d}\nu(y), \tag{1.8}$$

for any convex function $\varphi : \mathbb{R}^d \to \mathbb{R}$ non-negative or integrable with respect to μ and ν. This latter condition is equivalent, by Strassen's theorem [311], to the existence of a martingale coupling between μ and ν, i.e.

$$\exists \pi \in \Pi(\mu,\nu), \; \forall x \in \mathbb{R}^d, \; \int_{\mathbb{R}^d} y \mathrm{d}\pi(x,y) = x.$$

The original martingale optimal transport then consists in the resolution of the minimization problem

$$\inf_{\substack{\pi \in \Pi(\mu,\nu) \\ \forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} y \mathrm{d}\pi(x,y) = x}} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) \mathrm{d}\pi(x,y) \right\}, \tag{1.9}$$

with $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ being a l.s.c. cost function.

Financial application of multimarginal and martingale optimal transport Considering an asset with a price S_t at time t. An option on this asset is a tradable product which gives a payoff $\lambda(S_T)$ at time T (called maturity) to its owner. Typical options are calls ($\lambda : s \mapsto (s-K)^+$) or puts ($\lambda : s \mapsto (K-s)^+$) traded for various values of K (called strike) at fixed maturities ($T_1 < T_2 < \dots$). Knowing the price of

such options for any value of K is equivalent to knowing the probability distribution of the price of the asset at time T.

Thus, the modeling of the price of any financial product based on the asset must be in accordance with its option prices, and, under a no-arbitrage assumption, must follow a martingale. Recent works [27, 125, 128, 145] then used martingale optimal transport to provide model-free bounds on the considered financial product, which was used recently for VIX options [120, 173, 174]. This financial application led to several theoretical studies of martingale optimal transport [29, 181], its duality [33, 35, 93], martingale transport plans [119, 160] and its link with Skorokhod embeddings [25, 28, 172].

Numerical methods using sampling techniques [5, 6], entropic regularization and the Sinkhorn algorithm [118, 171, 247] have been used. However, those techniques might struggle in the case of multimarginal martingale optimal transport, i.e. when taking into account two or more maturities.

1.2.3 Numerical methods for optimal transportation

Several numerical methods have been introduced to solve optimal transport problems [224, 277], often in view of some particular applications. We summarize in this section the most widely used methods, and highlight the advantages and drawbacks of them, in particular with respect to the resolution of multimarginal or martingale optimal transport problems. Methods dedicated to the resolution of optimal transport problems stemming from quantum chemistry applications will be detailed in a forthcoming section.

In order to solve optimal transport problems involving discrete distributions, linear programming methods naturally arise such as the simplex [4], the Hungarian method [207] or the Auction algorithm [47, 70] or related improved linear programming methods [164]. The sparsity and the ordered structure of the minimas for particular cost function and particular space dimension, can be used in order to improve the complexity of the algorithms through the resolution of local subproblems [296, 299, 300] or via the use of proximal splitting [267], or solving local versions of the dual problem [253].

Semi-discrete optimal transport, for which only one marginal law is discrete, is another numerical approach and can be used for applications in dimension 2 or 3 [45, 177, 197, 215, 223, 252].

Differential methods using PDE's [176], gradient flows [19, 36, 175], Lagrangian method [184], its connection with fluid mechanics [37] or Newton method, and leveraging the connection between optimal transport and the Monge-Ampère equation [44, 234, 236, 294]) have also been studied to solve L^2 optimal transport with two marginal laws; the extension of these methods to unbalanced optimal transport has been considered in [38, 237].

Last but not least, let us mention the Sinkhorn algorithm (and some variants, such as Greenkhorn) [41, 83, 111, 154, 233, 297] which relies on the use of an entropic regularization of optimal transport which yields a strictly convex problem and which is very efficient to solve discrete two marginal relaxed optimal transport. This method can be adapted to unbalanced optimal transport problems [305]. Although data science and image processing [310] behaves nicely regarding the relaxation, it has to be noted that the speed of convergence of this class of algorithms decreases as the relaxation parameter goes to zero [200, 206, 306]. Let us mention the computation of Wasserstein barycenters as a multimarginal optimal transport application

for which the Sinkhorn algorithm is also efficient [112] and for which ad hoc methods also exist [101]. The use of the Sinkhorn algorithm for multimarginal optimal transport can also be analysed thanks to the multimarginal Schrödinger problem [84]. A study on some other regularization methods for optimal transport can also be found in [130].

Let us comment on the limitations of the aforementioned methods with respect to multimarginal optimal transport applications. In this context, the size of the resulting discrete optimal transport problems typically scales exponentially with the number of marginal laws. The semi-discrete methods make a crucial use of the two marginal structure and henceforth are not easily applicable to multimarginal problems. The differential methods mentioned above relies on results specific to the two marginal case. Lastly, the Sinkhorn algorithm relies on the computation of a cost matrix on a tensorized discretization grid. Although multimarginal problems [232] can be efficiently solved by means of these approaches, their complexity still scales exponentially with the number of marginal laws, making it not transferable to large systems.

1.3 Electronic structure calculations for molecules and main contributions of the thesis

The aim of this section is to give an introduction to *ab initio* modeling in quantum chemistry, more precisely electronic structure calculations of molecules, in particular to the many-body Schrödinger problem and Density Functional Theory. We refer to [77, 185, 268] for a more complete introduction.

1.3.1 The many-body Schrödinger electronic problem

In this section, we will make the use of *atomic units* so that

$$m_e = 1, \quad e = 1, \quad \hbar = 1, \quad \frac{1}{4\pi\epsilon_0} = 1, \tag{1.10}$$

where m_e is the mass of an electron, e is the elementary charge, $\hbar$ is the reduced Planck's constant, and ϵ_0 is the dielectric permittivity of the vacuum.

In the Born-Oppenheimer approximation, a molecule is a system composed of

- $M \in \mathbb{N}^*$ nuclei, which are considered as classical point-like particles, whose positions are denoted by $R_1, \dots, R_M \in \mathbb{R}^3$ and electrical charges by $Z_1, \dots, Z_M \in \mathbb{N}^*$;
- N electrons, which are modeled as quantum particles, and whose state is described by a function

$$\psi : \left\{ \begin{array}{ccc} \mathbb{R}^{3N} & \to & \mathbb{C} \\ (x_1, \dots, x_N) & \mapsto & \psi(x_1, \dots, x_N), \end{array} \right.$$

  called the *wavefunction* of the system of electrons.

Note that, in order to lighten notations, spin variables are omitted, since in the two part of this thesis, either the spin dependency can be separated from the treatment of the positions one (Part II), or it disapears in the semiclassical limit considered (Part I).

The physical interpretation of a wavefunction ψ is the following: given $A \subset \mathbb{R}^{3N}$, $\int_A |\psi|^2$ represents the probability that the positions of the N electrons belong to the set A. In particular, this implies that $\|\psi\|^2_{L^2(\mathbb{R}^{3N})} = 1$. In addition, the wavefunction ψ is *antisymmetric* with respect to its variables. This is a consequence of the fact that the electrons are fermionic particles. More precisely, denoting by $\mathcal{S}_N$ the set of permutations of the set $\{1, \dots, N\}$, it holds that for all $p \in \mathcal{S}_N$ and all $(x_1, \dots, x_N) \in \mathbb{R}^{3N}$,

$$\psi(x_{p(1)}, \dots, x_{p(N)}) = \epsilon(p)\psi(x_1, \dots, x_N),$$

where $\epsilon(p)$ denotes the signature of p.

The energy $E[\psi]$ of a system of N electrons whose state is described by a wavefunction ψ in the molecule described above is the sum of three contributions:

- the kinetic energy:
$$T[\psi] := \frac{1}{2} \int_{\mathbb{R}^{3N}} |\nabla \psi|^2;$$

- the Coulomb energy associated to the interactions between the electrons and the nuclei:
$$C_{\text{nuc}}[\psi] := \int_{\mathbb{R}^{3N}} \left(\sum_{i=1}^{N} V_{\text{nuc}}(x_i) \right) |\psi(x_1, \dots, x_N)|^2 \, dx_1 \cdots dx_N,$$
where, for all $x \in \mathbb{R}^3$,
$$V_{\text{nuc}}(x) := -\sum_{k=1}^{M} \frac{Z_k}{|x - R_k|};$$

- the Coulomb energy associated to the interactions between the electrons:
$$C_{\text{elec}}[\psi] := \int_{\mathbb{R}^{3N}} c(x_1, \dots, x_N) |\psi(x_1, \dots, x_N)|^2 \, dx_1 \dots dx_N,$$
where for almost all $(x_1, \dots, x_N) \in \mathbb{R}^{3N}$,
$$c(x_1, \dots, x_N) = \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}.$$

Computing a ground state of the electrons in the molecule amounts to computing a wavefunction ψ_0 among all admissible wavefunctions which minimize the energy of the system. More precisely, let us denote by

$$\mathcal{A} := \left\{ \psi \in L^2(\mathbb{R}^{3N}),\ \nabla\psi \in L^2(\mathbb{R}^{3N})^{3N},\ \psi \text{ antisymmetric},\ \|\psi\|_{L^2(\mathbb{R}^{3N})} = 1 \right\}$$

the set of wavefunctions associated to a system of N electrons with finite kinetic energy. Then, it holds that

$$U(R_1, \dots, R_M) = \min_{\psi \in \mathcal{A}} T[\psi] + C_{\text{nuc}}[\psi] + C_{\text{elec}}[\psi], \tag{1.11}$$

where we have highlighted the dependence of this infimum value with respect to the positions of the nuclei of the molecule $R_1, \dots, R_M$. Let $H := -\frac{1}{2}\Delta + \sum_{i=1}^{N} V_{\text{nuc}}(x_i) +$

$c(x_1, \ldots, x_N)$ be the so-called *many-body Schrödinger operator*. The operator H is a self-adjoint, bounded from below, operator on

$$L^2_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in L^2(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}$$

with domain

$$H^2_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in H^2(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}.$$

We also denote by

$$H^1_{\text{antisym}}(\mathbb{R}^{3N}) := \{\psi \in H^1(\mathbb{R}^{3N}),\ \psi \text{ antisymmetric}\}.$$

In the case when $U(R_1, \ldots, R_M) := \inf \sigma(H)$ is a discrete eigenvalue of H (which occurs for instance when the molecule is neutral or positively charged from Zhislin's theorem [326]), there exists at least one minimizer ψ_0 to (1.11), and any minimizer is necessarily an eigenvector of H associated to the eigenvalue $U(R_1, \ldots, R_M)$. Thus, solving the electronic Schrödinger problem amounts to solving a linear *high-dimensional* eigenvalue problem of the form

$$H\psi_0 = U(R_1, \ldots, R_M)\psi_0. \tag{1.12}$$

1.3.2 First contribution of the thesis: Van der Waals interaction between two hydrogen atoms

Although for large values of N and M approximations and numerical techniques must be used in order to evaluate $U(R_1, \ldots, R_M)$, for small systems analytical techniques can provide a way of solving Schrödinger equation.

This is the case for instance when considering the electronic interaction between two hydrogen atoms in the dissociation limit. In this asymptotic regime, the interaction between the two hydrogen atoms is called the van der Waals interaction. It is attractive and plays a crucial role in systems in the condensed phase such as biological molecules [21, 288] or 2D materials [153]. Studied from 1873 [317], van der Waals interaction has first been mathematically understood by London [238]. Its rigorous mathematical foundations have been investigated in the pioneering work by Morgan and Simon [261], inspired by the one of Ahlrichs in [3], and later by Lieb and Thiring [230], followed by many authors (see in particular [13, 205] and references therein). For H_2^+, the expansion of the interaction energy as a function of the distance R between the nuclei is a diverging series — yet Borel summable, as predicted in [63] and later proved by [100, 114, 168]. Recent articles have studied this expansion for collection of atoms [14, 18], with terms up to $1/R^9$ [22], molecules [15, 16] and its differentiability [17]. In the case of two hydrogen atoms, Slater and Kirkwood [309] provided a PDE, which allows to compute the first dispersion coefficient of the energy in the dissociation limit (which scales like $-C_6/R^6$, with R the distance between the nuclei). Cancès and Scott in [78], modified their technique, proved the well-posedness of the problem they proposed and used a Galerkin approximation to compute the C_6 coefficient.

An extension of the technique by Cancès and Scott has been studied during this thesis, and is presented in Chapter 4 in order to compute van der Waals dispersion coefficients up to any order. The technique relies on a perturbation method in order to analyse the asymptotic expansion of the van der Waals attraction and on a separation between radial and angular interactions which brings the original six-dimensional problem to the study of two-dimensional PDE's. Dispersion coefficients

can then be computed recursively by Galerkin approximations; values were found following this approach in accordance with the ones in [204, 265], where the authors used other techniques.

1.3.3 Density Functional Theory

The high-dimensional character of equation (1.12) makes it very difficult to solve from a numerical point of view with standard numerical methods in the case when N is large, especially for strongly correlated systems where the Coulombic interactions between the nuclei play a significant role.

The principle of Density Functional Theory (DFT), and of all the models which are derived from it, is the reformulation of problem (1.11) with the density (and not anymore the wavefunction) as the main variable. The key advantage of this method is that problems are then formulated over the domain $\mathbb{R}^3$ instead of $\mathbb{R}^{3N}$.

Theoretical justification of DFT models has been introduced by Hohenberg and Kohn [182], followed by Levy [225] and completed by Lieb [229]. We refer the reader to the review chapter [228]. Indeed, the Hohenberg-Kohn theorem [182] states that the energy and the electronic density of the ground state of the electronic problem (1.11) can be found by solving a problem of the form

$$U(R_1, \ldots, R_M) = \inf \left\{ F(\rho) + \int_{\mathbb{R}^3} \rho V,\ \rho \in L^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \rho = N \right\},$$

where F is a functional of the electronic density ρ.

Let us rewrite H under the following form

$$H = H_V = H_0 + \sum_{i=1}^{N} V_{\text{nuc}}(x_i),$$

where

$$H_0 := -\sum_{i=1}^{N} \frac{1}{2}\Delta_{x_i} + \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}, \tag{1.13}$$

in order to highlight the dependence of the Hamiltonian on the potential V_{nuc}. The minimization problem (1.11) can then be rewritten as

$$U(R_1, \ldots, R_M) = \inf \left\{ \langle \psi_e, H_V \psi_e \rangle,\ \psi_e \in \mathcal{A} \right\}. \tag{1.14}$$

We also denote by

$$\mathcal{I}_N := \{ \rho,\ \exists \psi_e \in \mathcal{A},\ \rho_{\psi_e} = \rho \}$$

the set of all electronic densities associated with some admissible wavefunction, where the density ρ_{ψ_e} associated to the wavefunction ψ_e is defined for $x \in \mathbb{R}^3$

$$\rho_{\psi_e}(x) = N \int_{\mathbb{R}^{3(N-1)}} |\psi_e(x, x_2, \ldots, x_N)|^2 dx_2 \ldots dx_N.$$

It is proved in [229] that $\mathcal{I}_N$ can be characterized equivalently as

$$\mathcal{I}_N = \left\{ \rho \ge 0,\ \sqrt{\rho} \in H^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \rho = N \right\}.$$

DFT relies on the following elementary calculus [182, 229]:

$$\begin{aligned} U(R_1,\dots,R_M) &= \inf\left\{\langle\psi_e, H_V\psi_e\rangle,\ \psi_e\in\mathcal{A}\right\} \\ &= \inf\left\{\inf\left\{\langle\psi_e, H_0\psi_e\rangle,\ \psi_e\in\mathcal{A},\ \rho_{\psi_e}=\rho\right\} + \int_{\mathbb{R}^3}\rho V,\ \rho\in\mathcal{I}_N\right\} \\ &= \inf\left\{F_{LL}(\rho) + \int_{\mathbb{R}^3}\rho V,\ \rho\in\mathcal{I}_N\right\}, \end{aligned}$$

where

$$F_{LL}(\rho) := \inf\left\{\langle\psi_e, H_0\psi_e\rangle,\ \psi_e\in\mathcal{A},\ \rho_{\psi_e}=\rho\right\}$$

is called the *Levy-Lieb functional*. It is universal in the sense that it does not depend on the molecular system under consideration (which only comes into play through the potential V and the number of electrons N). Equivalently, the Levy-Lieb functional can be rewritten as

$$F_{LL}(\rho) := \left\{T[\psi] + C_{\text{elec}}[\psi],\ \psi\in\mathcal{A},\ \rho_\psi=\rho\right\}.$$

This is a very appealing theory, but unfortunately, the exact computation of $F_{LL}(\rho)$ is out-of-reach since it requires the resolution of a problem almost as complex as the original electronic Schrödinger problem.

In practice then, approximations of the functional F_{LL} are used, which gives rise to a wide zoology of DFT models.

1.3.4 Semi-classical limit of the Levy-Lieb functional

One of these approximations, which was suggested by theoretical chemists in [302, 304], consists in looking to the *semi-classical* or *strongly correlated electrons* (SCE) limit of the Levy-Lieb functional, with a view to use it in order to design approximate DFT models for strongly correlated systems. This semi-classical limit is the limit as α goes to 0 to the functional F_{LL}^α defined as follows for $\rho\in\mathcal{I}_N$ and $0<\alpha\leq 1$:

$$F_{LL}^\alpha(\rho) := \left\{\alpha T[\psi] + C_{\text{elec}}[\psi],\ \psi\in\mathcal{A},\ \rho_\psi=\rho\right\}.$$

In this semi-classical limit, the influence of the kinetic term $T[\psi]$ is then neglected in front of the contributions due to the electron-electron Coulombic interaction term $C_{\text{elec}}[\psi]$. It has been rigorously proven in the series of works [106, 107, 226] that the limit as α goes to 0 of the functional $F_{LL}^\alpha(\rho)$ reads as a *symmetric multi-marginal optimal transport problem with Coulomb cost*. More precisely, for all $\rho\in\mathcal{I}_N$, let us denote by ν_ρ the probability measure on $\mathbb{R}^3$ defined by $d\nu_\rho(x) := \frac{\rho(x)}{N}\,dx$ and by $\mathcal{P}_{\text{sym}}(\mathbb{R}^{3N})$ the set of symmetric probability measures on $\mathbb{R}^{3N}$. For all $\gamma\in\mathcal{P}_{\text{sym}}(\mathbb{R}^{3N})$, we denote by μ_γ the probability measure on $\mathbb{R}^3$ defined as the marginal of γ, i.e.

$$d\mu_\gamma(x) := \int_{(x_2,\dots,x_N)\in\mathbb{R}^{3(N-1)}} d\gamma(x,x_2,\dots,x_N).$$

With work of Buttazzo, De Pascale and Gori Giorgi [67], Cotar, Friesecke and Klüppelberg [106] for proofs for $N=2$ and with Mendl and Pass [142] , Bindini and De Pascale [50] extended by Lewin [226] for $N\geq 2$ in the fermonic mixed-states case and Cotar, Friesecke and Klüppelberg [107] for $N\geq 2$, it is proved, using appropriate smoothing of transport plans, that in the semi-classical limit

$$\lim_{\alpha\to 0} F_{LL}^\alpha(\rho) = I(\nu_\rho),$$

where for all probability measure ν on $\mathbb{R}^3$,

$$I(\nu) := \inf_{\substack{\gamma \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{3N}),\\ \mu_\gamma = \nu}} \int_{\mathbb{R}^{3N}} c \, d\gamma. \tag{1.15}$$

A different (but closely related) approach to this limit also exists through density scaling [91]. The asymptotic expansion to the next orders of $F^{\alpha}_{LL}(\rho)$ with respect to α has been studied in [167, 170].

This problem has been well studied in the recent years. Although some Monge maps can be exhibited in spherically symmetric cases (which are close to optimality [105, 303]), there exists when $N \geq 3$ in general some non-Monge minimizers [272] concentrated on higher dimensional submanifolds (non-necessary unique), and the minimizer is unique and non-Monge when $N = \infty$ [108]. Let us mention as studied subjects the continuity of multimarginal optimal transport and of its maps (with studies with repulsive costs other than the Coulombic one) [66, 66, 102, 104], duality theory [103, 121, 157], relaxation [156, 159]. Let us also note [52, 109, 169, 209, 255] as other works on the subject.

The SCE formulation of DFT has also already been applied to model quantum systems [110, 165, 244, 246, 256], and appears in the study of uniform electron gas [227]; comparisons with other DFT methods can be found in [166, 243, 245].

A classical way to approximate the problem (1.15) is to use a (fixed) discrete state space $\{y_1, \dots, y_M\} \subset \mathbb{R}^3$ for some $M \in \mathbb{N}^*$ and compute an approximation of a solution γ to (1.15) under the form

$$\gamma \approx \sum_{1 \leq m_1, \dots, m_N \leq M} \lambda_{m_1, \dots, m_N} \delta_{(y_{m_1}, \dots, y_{m_N})}$$

where the M^N real coefficients $(\lambda_{m_1, \dots, m_N})_{1 \leq m_1, \dots, m_N \leq M}$ have to be determined. This leads to a very high-dimensional linear optimization problem.

1.3.5 Numerical methods for the resolution of (1.15)

The challenges raised by multimarginal optimal transport with a Coulombic cost led to the development of dedicated numerical methods which are exposed in this section.

First, let us mention the work of Mendl and Lin [251], which, using the Kantorovich dual formulation computes the SCE limit for atoms and small molecules. Let us also note the work of Chen, Friesecke and Mendl [92] in the 2 electrons case which uses a smart meshing method to compute precisely a minimizer for the SCE formulation of DFT for the H_2 molecule. Nenna in [42, 262] uses the Sinkhorn algorithm to solve a relaxed multimarginal optimal transport problem for atoms, up to $N = 3$ electrons in the radially symmetric case. However, in the duality case, checking the inequality constraint is not easy and in the two later methods, scaling to more electrons is not easy either.

In more recent works by Friesecke and Vögler [143, 320] and Khoo, Ying, Lin and Lindsey [192, 193], numerical methods on finite state space break the curse of dimensionality, with complexity growing linearly with the number of electrons. Note that, for some particular multimarginal problems (Wasserstein barycenters), this linear complexity had been showed [86].

1.3.6 Second contribution of the thesis: moment constrained approximation of multi-marginal optimal transportation problems

A second contribution of the thesis is to propose and analyze from a mathematical point of view an alternative approach in order to approximate the symmetric optimal transport problem (1.15). In this approach, we still consider a continuous state space $\mathbb{R}^3$, but the marginal contraint appearing in (1.15) is relaxed into a finite number of moment constraints. For the sake of simplicity, let us present our results here in the case when the support of the measure ν is included in a compact set $Y \subset \mathbb{R}^3$. Let $(f_m)_{m\in\mathbb{N}^*} \subset \mathcal{C}(Y)$, satisfying the following natural density assumption

$$\forall f \in \mathcal{C}(Y), \qquad \inf_{g_M \in \mathrm{Span}\{f_1,\dots,f_M\}} \|f - g_M\|_{L^\infty} \underset{M\to+\infty}{\longrightarrow} 0,$$

and consider the approximate moment constrained optimal transport problem

$$I^M(\nu) := \inf_{\substack{\gamma \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{3N}),\\ \forall 1 \le m \le M,\\ \int_{\mathbb{R}^{3N}} \left(\frac{1}{N}\sum_{i=1}^N f_m(x_i)\right) d\gamma(x_1,\dots,x_N) = \int_{\mathbb{R}^3} f_m \, d\nu}} \int_{\mathbb{R}^{3N}} c \, d\gamma. \tag{1.16}$$

It is proved in Chapter 2 where $\mathcal{P}(\mathbb{R}^{3N})$ denotes the set of (not necessarily symmetric) probability measures on $\mathbb{R}^{3N}$.

Theorem 1.3. *Under the preceding assumptions, it holds that*

$$I^M(\nu) \underset{M\to+\infty}{\longrightarrow} I(\nu).$$

Besides, it holds that

$$I^M(\nu) = \inf_{\substack{\gamma \in \mathcal{P}(\mathbb{R}^{3N}),\\ \forall 1 \le m \le M,\\ \int_{\mathbb{R}^{3N}} \left(\frac{1}{N}\sum_{i=1}^N f_m(x_i)\right) d\gamma(x_1,\dots,x_N) = \int_{\mathbb{R}^3} f_m \, d\nu}} \int_{\mathbb{R}^{3N}} c \, d\gamma, \tag{1.17}$$

and there exists at least one minimizer $\gamma^M \in \mathcal{P}(\mathbb{R}^{3N})$ to (1.17) which reads as

$$\gamma^M = \sum_{k=1}^K w_k \delta_{(x_1^k,\dots,x_N^k)}$$

for some $1 \le K \le M+2$, and for some $w_k \ge 0$ and $(x_1^k,\dots,x_N^k) \in Y^N$ for all $1 \le k \le K$. Besides,

$$\gamma^M_{\mathrm{sym}} = \frac{1}{N!}\sum_{p\in\mathcal{S}_N}\sum_{k=1}^K w_k \delta_{(x^k_{p(1)},\dots,x^k_{p(N)})},$$

the symmetrized version of γ^M, is a minimizer to (1.16).

Theorem 1.3 states two things: (i) it is possible to drop the symmetry constraint of the measure γ in problem (1.16) to compute $I^M(\nu)$; (ii) there exists a minimizer of (1.17) which reads as a discrete measure which charges a low number of points (less

than $M+2$), and a minimizer to (1.16) can be obtained as the symmetrized version of this discrete measure. In particular, this means that it is sufficient to identify at most $\mathcal{O}(NM)$ scalars to compute γ^M. This suggests considering the following optimization problem for the computation of $I^M(\nu)$, since

$$I^M(\nu) = \min_{\substack{(w_k)_{1\leq k\leq M+2} \in \mathbb{R}_+^{M+2}, \\ \sum_{k=1}^{M+2} w_k = 1, \\ \left(x_1^k, \dots, x_N^k\right) \in Y^N,\ \forall 1 \leq k \leq M+2, \\ \sum_{k=1}^{M+2} w_k \left(\frac{1}{N}\sum_{i=1}^N f_m(x_i^k)\right) = \int_{\mathbb{R}^3} f_m \, d\nu}} \sum_{k=1}^{M+2} w_k c(x_1^k, \dots, x_N^k). \quad (1.18)$$

The use of this sparse structure for the design of efficient numerical methods for the resolution of (1.16) is the object of Chapter 3. It is proved in particular that any local minimizer to (1.18) is actually a *global* minimizer. In addition, the numerical method proposed for the resolution of this problem builds on the use of constrained overdamped Langevin processes.

Let us stress on the fact that the theorems and results presented in Chapter 2 and Chapter 3 can be extended to general multi-marginal optimal transport problems, as well as martingale optimal transport problems, and thus can be used, in addition to the quantum chemistry applications highlighted in this section, for the financial applications mentioned in Section 1.2.2.

Part I

Moment Constrained Optimal Transport

Chapter 2

Moment Constrained Optimal Transport

This chapter is an article written with Aurélien Alonsi, Virginie Ehrlacher and Damiano Lombardi and published in *Mathematics of Computations* [8].

Abstract

Optimal Transport (OT) problems arise in a wide range of applications, from physics to economics. Getting numerical approximate solutions of these problems is a challenging issue of practical importance. In this work, we investigate the relaxation of the OT problem when the marginal constraints are replaced by some moment constraints. Using Tchakaloff's theorem, we show that the Moment Constrained Optimal Transport problem (MCOT) is achieved by a finite discrete measure. Interestingly, for multimarginal OT problems, the number of points weighted by this measure scales linearly with the number of marginal laws, which is encouraging to bypass the curse of dimension. This approximation method is also relevant for Martingale OT problems. We show the convergence of the MCOT problem toward the corresponding OT problem. In some fundamental cases, we obtain rates of convergence in $O(1/N)$ or $O(1/N^2)$ where N is the number of moments, which illustrates the role of the moment functions. Last, we present algorithms exploiting the fact that the MCOT is reached by a finite discrete measure and provide numerical examples of approximations.

2.1 Introduction

The aim of this paper is to investigate a new direction to approximate optimal transport problems [291, 319]. Such problems arise in a variety of application fields ranging from economics [82, 146] to quantum chemistry [108] or machine learning [277] for instance. The simplest prototypical example of optimal transport problem is the two-marginal Kantorovich problem, which reads as follows: for some $d \in \mathbb{N}^*$, let μ and ν be two probability measures on $\mathbb{R}^d$ and consider the optimization problem

$$\inf \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\, \mathrm{d}\pi(x,y) \tag{2.1}$$

where c is a non-negative lower semi-continuous cost function defined on $\mathbb{R}^d \times \mathbb{R}^d$ and where the infimum is taken over the set of probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginal laws μ and ν.

The most straightforward approach for the resolution of problems of the form (2.1) consists in introducing discretizations of the state spaces, which are fixed a priori. More precisely, N points $x^1, \cdots, x^N \in \mathbb{R}^d$ are chosen a priori and fixed, marginal laws μ and ν are approximated by discrete measures of the form $\mu \approx \sum_{i=1}^N \mu_i \delta_{x^i}$ and $\nu \approx \sum_{i=1}^N \nu_i \delta_{x^i}$ with some non-negative coefficients μ_i and ν_i for $1 \leq i \leq N$. An optimal measure π minimizing (2.1) is then approximated by a discrete measure $\pi \approx \sum_{1 \leq i,j \leq N} \pi_{ij} \delta_{x^i, x^j}$ where the non-negative coefficients $(\pi_{ij})_{1 \leq i,j \leq N} \in \mathbb{R}_+^{N^2}$ are solution to the optimization problem

$$\inf \sum_{1 \leq i,j \leq N} \pi_{ij} c(x^i, x^j) \tag{2.2}$$

and satisfy the following discrete marginal constraints:

$$\forall 1 \leq i, j \leq N, \quad \sum_{j=1}^N \pi_{ij} = \mu_i \quad \text{and} \quad \sum_{i=1}^N \pi_{ij} = \nu_j.$$

This boils down to a classical linear programming problem, which becomes computationally prohibitive when N is large.

Several numerical methods have already been proposed in the literature for the resolution of optimal transport problems at a lower computational cost. Most of them rely on an a priori discretization of the state spaces as presented above. One of the most successful approach consists in minimizing a regularized cost involving the Kullback-Leibler divergence (or relative entropy) via iterative Bregman projections: the so-called Sinkhorn algorithm [41, 262, 306]. Let us also mention other approaches such as the auction algorithm [48], numerical methods based on Laguerre cells [148], multiscale algorithms [252, 296] and augmented Lagrangian methods using the Benamou-Brenier dynamic formulation [39, 40].

In this work, we are also interested in studying multi-marginal and martingale-constrained optimal transport problems.

Multimarginal optimal transport problems arise in a wide variety of contexts [291, 319], like for instance the computation of Wasserstein barycenters [2] or the approximation of the correlation energy for strongly correlated systems in quantum chemistry [102, 108, 304]. Such problems read as follows: let $M \in \mathbb{N}^*$ and $\mu_1, \cdots, \mu_M$ be M probability measures on $\mathbb{R}^d$ and consider the optimization problem

$$\inf \int_{(\mathbb{R}^d)^M} c(x_1, \cdots, x_M) \, \mathrm{d}\pi(x_1, \cdots, x_M) \tag{2.3}$$

where c is a lower semi-continuous cost function defined on $(\mathbb{R}^d)^M$ and where the infimum runs on the set of probability measures π on $(\mathbb{R}^d)^M$ with marginal laws given by $\mu_1, \cdots, \mu_M$. Approximations of such multi-marginal problems on discrete state spaces can be introduced similar to (2.2), leading to a linear programming problem of size N^M. For large values of M, such discretized problems become intractable numerically. The most successful method up to now for solving such problems, which avoids this curse of dimensionality, is based on an entropic regularization together with the Sinkhorn algorithm [41, 262].

Constrained martingale transport arise in problems related to finance [28]. Few numerical methods have been proposed so far for the resolution of such problems. In [5, 6], algorithms using sampling techniques preserving the convex order is proposed, which enables then to use linear programming solvers. Algorithms making

use of an entropy regularization and the Sinkhorn algorithm have been studied in [118, 171].

In this paper, we consider an alternative direction to approximate optimal transport problems, with a view to the design of numerical schemes. In this approach, the state space is *not* discretized, but the approximation consists in relaxing the marginal laws constraints (or the martingale constraint) of the original problem into a finite number of moment constraints against some well-chosen test functions. More precisely, in the case of Problem (2.1), let us introduce some real-valued bounded functions $\phi_1, \cdots, \phi_N$ defined on $\mathbb{R}^d$, which are called hereafter *test functions*, and consider the approximate optimal transport problem, called hereafter the Moment Constrained Optimal Transport (MCOT) problem:

$$\inf \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\, d\pi(x,y)$$

where the infimum is taken over the set of probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ satisfying for all $1 \leq i, j \leq N$,

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \phi_i(x)\, \mathrm{d}\pi(x,y) = \int_{\mathbb{R}^d} \phi_i(x)\, \mathrm{d}\mu(x) \text{ and } \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi_j(y)\, \mathrm{d}\pi(x,y) = \int_{\mathbb{R}^d} \phi_j(y)\, \mathrm{d}\nu(y).$$

The aim of this paper is to study the properties of this alternative approximation of optimal transport problems, and its generalization for multi-marginal and martingale-constrained optimal transport problems. A remarkable feature of this approximation is that it circumvents the curse of dimensionality with respect to the number of marginal laws in the case of multimarginal optimal transport problems. Note that in the martingale constrained case, our method enables to consider the original formulation of the financial problem that has moment constraints (see for instance Example 2.6 of [180]), which is not the case of the previous methods.

Our first contribution in this paper is to characterize some minimizers of the MCOT problem. Using Tchakaloff's theorem, we prove that, under suitable assumptions, there exists at least one minimizer which is a discrete measure charging a finite number of points. Interestingly, in the multi-marginal case, the number of charged points scales at most linearly in the number of marginals. In the particular case of problems issued from quantum chemistry applications, due to the inherent symmetries of the system, the number of charged points is independent of the number of marginals, and only scales linearly with the number of imposed moments. This formulation of the multimarginal optimal transport problem thus does not suffer from the curse of dimensionality. The result obtained in the quantum chemistry case is close in spirit to the one of [143] where the authors studied a multimarginal Kantorovich problem with Coulomb cost on finite state spaces.

These considerations motivate us to consider a new family of algorithms for the resolution of multi-marginal and martingale constrained optimal transport problems, in which an optimal measure is approximated by a discrete measure charging a relatively low number of points. The points and weights of this discrete measure are then optimized in order to satisfy a finite number of moment constraints and to minimize the cost of the original optimal transport problem.

Of course, another interesting issue consists in determining how the choice of the particular test functions influences the quality of the approximation with respect to the exact optimal transport problem. In this paper, we prove interesting one-dimensional results in this direction. More precisely, for piecewise affine test

functions defined on a compact interval, and for the two-marginal optimal transport problems involved in the computation of the W_2 or the W_1 distance between two sufficiently regular measures, the convergence of the approximate optimal cost with respect to the optimal cost scales like $\mathcal{O}\left(\frac{1}{N^2}\right)$ where N is the number of test functions. These results indicate that the choice of appropriate test functions has an influence on the rate of convergence of the approximate problem to the exact problem. Besides, there is very few results, up to our knowledge, concerning the speed of convergence of approximations of optimal transport problems. The study of these rates of convergence for more general set of test functions and of optimal transport problems is an interesting issue which is left for future research.

The article is organized as follows. Some preliminaries, including the Tchakaloff theorem, are recalled in Section 2.2. In Section 2.3, we introduce the approximate MCOT problem and prove under suitable assumptions that one of its minimizers reads as a discrete measure whose number of charged points is estimated depending on the number of moment constraints and on the nature of the optimal transport problem considered. Under additional assumptions, we prove that the MCOT problem converges to the exact optimal transport problem as the number of test functions grows in Section 2.4. Rates of convergence of the approximate problem to the exact problem depending on the choice of test functions are proved in Section 2.5. Finally, algorithms which exploits the particular structure of the MCOT problem are proposed in Section 2.6 and tested on some numerical examples.

2.2 Preliminaries

2.2.1 Presentation of the problem and notation

We begin this section by recalling the classical form of the 2-marginal optimal transport (OT) problem, which will be the prototypical problem considered in this paper, and introduce the notation used in the sequel.

Let $d_x, d_y \in \mathbb{N}^*$. We assume that $\mathcal{X} \subset \mathbb{R}^{d_x}$ (resp. $\mathcal{Y} \subset \mathbb{R}^{d_y}$) is a G_δ-set, i.e. a countable intersection of open sets. This property ensures by Alexandroff's lemma (see e.g. [9], p. 88) that $\mathcal{X}$ (resp. $\mathcal{Y}$) is a Polish space for a metric which is equivalent to the original one on $\mathbb{R}^{d_x}$ (resp. $\mathbb{R}^{d_y}$). In particular, $\mathcal{X}$ can either be a closed or an open set of $\mathbb{R}^{d_x}$.

Let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$ be probability measures on $\mathcal{X}$ and $\mathcal{Y}$ and let us define

$$\Pi(\mu,\nu) := \left\{ \pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}) : \int_{\mathcal{X}} \mathrm{d}\pi(x,y) = \mathrm{d}\nu(y),\ \int_{\mathcal{Y}} \mathrm{d}\pi(x,y) = \mathrm{d}\mu(x) \right\},$$

the set of probability couplings between μ and ν. We consider a non-negative cost function $c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$, which we assume to be lower semi-continuous (l.s.c.). The Kantorovich optimal transport (OT) problem with the two marginal laws μ and ν associated to the cost function c is the following minimization problem:

$$I = \inf \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y) : \pi \in \Pi(\mu,\nu) \right\}. \tag{2.4}$$

Under our assumptions, it is known (see e.g. Theorem 1.7 in [291]) that there exists $\pi^* \in \Pi(\mu,\nu)$ such that $I = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi^*(x,y)$. Problem (2.4) will be referred hereafter as the *exact* OT problem, with respect to the *approximate* problem which we define hereafter.

The aim of this paper is to study a relaxation of Problem (2.4) with a view to the design of numerical schemes to approximate the exact OT problem. More precisely, the *approximate* problem considered in this paper consists in relaxing the marginal constraints into a *finite number of* moments constraints. Let $M, N \in \mathbb{N}^*$ and $(\phi_m)_{1\leq m\leq M} \subset L^1(\mathcal{X}, \mu; \mathbb{R})$ (respectively $(\psi_n)_{1\leq n\leq N} \subset L^1(\mathcal{Y}, \nu; \mathbb{R})$) measurable real-valued functions that are integrable with respect to μ (resp. ν). The functions $(\phi_m)_{1\leq m\leq M}$ and $(\psi_n)_{1\leq n\leq N}$ will be called *test functions* hereafter. We define for such families of functions

$$\Pi(\mu, \nu; (\phi_m)_{1\leq m\leq M}, (\psi_n)_{1\leq n\leq N}) := \Bigg\{ \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \tag{2.5}$$

$$\forall 1 \leq m \leq M, 1 \leq n \leq N, \int_{\mathcal{X}\times\mathcal{Y}} |\phi_m(x)| + |\psi_n(y)| \mathrm{d}\pi(x, y) < \infty,$$

$$\int_{\mathcal{X}\times\mathcal{Y}} \phi_m(x)\mathrm{d}\pi(x, y) = \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\mu(x),\ \int_{\mathcal{X}\times\mathcal{Y}} \psi_n(y)\mathrm{d}\pi(x, y) = \int_{\mathcal{X}} \psi_n(y)\mathrm{d}\mu(x) \Bigg\},$$

which is the set of probability measures on $\mathcal{X} \times \mathcal{Y}$ that have the same moments as μ and ν for the test functions. We are then interested in the moment constrained optimal transport (MCOT) problem, which we define as the following minimization problem :

$$I^{M,N} = \inf \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x, y)\mathrm{d}\pi(x, y) : \pi \in \Pi(\mu, \nu; (\phi_m)_{1\leq m\leq M}, (\psi_n)_{1\leq n\leq N}) \right\}. \tag{2.6}$$

Since $\Pi(\mu, \nu) \subset \Pi(\mu, \nu; (\phi_m)_{1\leq m\leq M}, (\psi_n)_{1\leq n\leq N})$, we clearly have $I^{M,N} \leq I$. In this paper, we are interested in the following question.

- Is the infimum of the MCOT problem attained by some probability measure $\pi^* \in \Pi(\mu, \nu; (\phi_m)_{1\leq m\leq M}, (\psi_n)_{1\leq n\leq N})$?
- Under which assumptions does it hold: $I^{M,N} \underset{M,N\to+\infty}{\to} I$? Can the speed of convergence be estimated?

For simplicity, we will assume that $M = N$ in the whole paper, and we will denote for $1 \leq m, n \leq N$:

$$\bar{\mu}_m := \int_{\mathcal{X}} \phi_m \mathrm{d}\mu \quad \text{and} \quad \bar{\nu}_n := \int_{\mathcal{Y}} \psi_n \mathrm{d}\nu. \tag{2.7}$$

For all $x \in \mathcal{X}$ (respectively for all $y \in \mathcal{Y}$), we define $\phi(x) := (\phi_1(x), ..., \phi_N(x)) \in \mathbb{R}^N$ (respectively $\psi(y) := (\psi_1(y), ..., \psi_N(y)) \in \mathbb{R}^N$) and $\Phi(x) := (1, \phi(x)) \in \mathbb{R}^{N+1}$ (respectively $\Psi(y) := (1, \psi(y)) \in \mathbb{R}^{N+1}$).

2.2.2 Tchakaloff's theorem

In this section, we present a corollary of the Tchakaloff theorem which is the backbone of our results concerning the existence of a minimizer to the MCOT problem. A general version of the Tchakaloff theorem has been proved by Bayer and Teichmann [24] and Bisgaard [46]. The next proposition is a rather immediate consequence of Tchakaloff's theorem, see Corollary 2 in [24]. We recall first that

Theorem 2.1. *Let π be a measure on $\mathbb{R}^d$ concentrated on a Borel set $A \in \mathcal{F}$, i.e. $\pi(\mathbb{R}^d \setminus A) = 0$. Let $N_0 \in \mathbb{N}^*$ and $\Lambda : \mathbb{R}^d \to \mathbb{R}^{N_0}$ a Borel measurable map. Assume that the first moments of $\Lambda\#\pi$ exist, i.e.*

$$\int_{\mathbb{R}^{N_0}} \|u\| \mathrm{d}\Lambda\#\pi(u) = \int_{\mathbb{R}^d} \|\Lambda(z)\| \mathrm{d}\pi(z) < \infty,$$

where $\|\cdot\|$ denotes the Euclidean norm of $\mathbb{R}^{N_0}$. Then, there exist an integer $1 \leq K \leq N_0$, points $z_1, ..., z_K \in A$ and weights $p_1, ..., p_K > 0$ such that

$$\forall 1 \leq i \leq N_0, \quad \int_{\mathbb{R}^d} \Lambda_i(z) \mathrm{d}\pi(z) = \sum_{k=1}^{K} p_k \Lambda_i(z_k),$$

where Λ_i denotes the i-th component of Λ.

We recall here that $\Lambda\#\pi$ is the push-forward of π through Λ, and is defined as $\Lambda\#\pi(A) = \pi(\Lambda^{-1}(A))$ for any Borel set $A \subset \mathbb{R}^{N_0}$. Let us note here that even if π is a probability measure, we may have $\sum_{k=1}^{K} p_k \neq 1$. In the sequel, we will apply this proposition to functions Λ such that the constant 1 is a coordinate of Λ, which will ensure $\sum_{k=1}^{K} p_k = 1$.

Let us remark that the number of points K needed may be significantly smaller than N_0. Lemma A.1 gives, for any $N \in \mathbb{N}^*$, an example with $N_0 = 2N + 1$ and $K = N + 1$.

Last, let us mention that Theorem 2.1 is a consequence of Caratheodory's theorem [287, Corollary 17.1.2] applied to $\int_{\mathbb{R}^{N_0}} u \mathrm{d}\Lambda\#\pi(u)$ which lies in the (convex) cone induced by $\mathrm{spt}(\Lambda\#\pi)$, the support of the measure $\Lambda\#\pi$.

2.2.3 An admissibility property

A natural requirement for the MCOT Problem (2.6) to make sense is to assume that it has finite value. This is precisely our definition of admissibility.

Definition 2.1 (Admissibility)**.** *Let $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$ and a l.s.c. cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+ \cup \{\infty\}$. Then, a set of test functions $((\phi_m)_{1\leq m\leq N}, (\psi_n)_{1\leq n\leq N}) \in L^1(\mathcal{X}, \mu; \mathbb{R})^N \times L^1(\mathcal{Y}, \nu; \mathbb{R})^N$ is said to be admissible for (μ, ν, c) if*

$$\exists \gamma \in \Pi(\mu, \nu; (\phi_m)_{1\leq m\leq M}, (\psi_n)_{1\leq n\leq N}), \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\gamma(x, y) < \infty. \tag{2.8}$$

Thanks to Tchakaloff's theorem, the admissibility can be checked on finite probability measure as stated in the next Lemma.

Lemma 2.2. *Let $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$ and $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ a l.s.c. function. A set $((\phi_m)_{1\leq m\leq N}, (\psi_n)_{1\leq n\leq N}) \in L^1(\mathcal{X}, \mu; \mathbb{R})^N \times L^1(\mathcal{Y}, \nu; \mathbb{R})^N$ is admissible for (μ, ν, c) if, and only if, there exist weights $w_1, \ldots, w_{2N+1} \geq 0$ and points $(x_1, y_1), \ldots, (x_{2N+1}, y_{2N+1}) \in \mathcal{X} \times \mathcal{Y}$ such that*

$$\sum_{k=1}^{2N+1} w_k \delta_{(x_k, y_k)} \in \Pi(\mu, \nu; (\phi_m)_{1\leq m\leq N}, (\psi_n)_{1\leq n\leq N}) \text{ and } \sum_{k=1}^{2N+1} w_k c(x_k, y_k) < \infty.$$

In particular, if c is finite valued (i.e. $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$), any set of test functions $((\phi_m)_{1\leq m\leq N}, (\psi_n)_{1\leq n\leq N}) \in L^1(\mathcal{X}, \mu; \mathbb{R})^N \times L^1(\mathcal{Y}, \nu; \mathbb{R})^N$ is admissible for (μ, ν, c) in the sense of Definition 2.1.

Proof. Let $\Lambda : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^{2N+1}$ be defined by $\Lambda_m(x,y) = \phi_m(x)$ and $\Lambda_{m+N}(x,y) = \psi_m(y)$ for $m \in \{1,\dots,N\}$, $\Lambda_{2N+1}(x,y) = 1$. Let $A = \{(x,y) \in \mathcal{X} \times \mathcal{Y} : c(x,y) = +\infty\}$. Since the set of test function is admissible, there exists a probability measure $\gamma \in \Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})$ such that $\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\gamma(x,y) < \infty$. In particular, $\gamma(A) = 0$. We can thus apply Theorem 2.1, which gives the implication. The reciprocal result is obvious.

Last, when c is finite valued ($A = \emptyset$), any $\gamma \in \Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})$ satisfies $\gamma(A) = 0$ and the claim follows by using again Theorem 2.1. □

2.3 Existence of discrete minimizers for MCOT problems

2.3.1 Two-marginal case

When Definition 2.1 is satisfied, in order to have the existence of a minimizer for the MCOT problem, we make two further assumptions.

- We assume that the test function are continuous.
- We add to the MCOT problem (2.6) a moment inequality constraint.

The additional moment constraint will ensure the tightness of a minimizing sequence satisfying the moment equality and inequality constraints, while the continuity of the test functions will ensure that any limit satisfies the moment constraints. Our main result is stated in Theorem 2.3 thereafter.

Theorem 2.3. *Let $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$ and $c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ a l.s.c. function. Let $\Sigma_\mu \subset \mathcal{X}, \Sigma_\nu \subset \mathcal{Y}$ be Borel sets such that $\mu(\Sigma_\mu) = \nu(\Sigma_\nu) = 1$. Let $N \in \mathbb{N}^*$ and let $((\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N}) \in L^1(\mathcal{X},\mu;\mathbb{R})^N \times L^1(\mathcal{Y},\nu;\mathbb{R})^N$ be an admissible set of test functions for (μ,ν,c) in the sense of Definition 2.1. We assume that*

1. *For all $n \in \{1,\dots,N\}$, the functions ϕ_n and ψ_n are continuous.*
2. *There exist $\theta_\mu : \mathbb{R}_+ \to \mathbb{R}_+$ and $\theta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ two non-negative non-decreasing continuous functions such that $\theta_\mu(r) \underset{r\to+\infty}{\longrightarrow} +\infty$ and $\theta_\nu(r) \underset{r\to+\infty}{\longrightarrow} +\infty$, and such that there exist $C > 0$ and $0 < s < 1$ such that for all $1 \le n \le N$, and all $(x,y) \in \mathcal{X}\times\mathcal{Y}$,*

$$|\phi_n(x)| \le C(1+\theta_\mu(|x|))^s \quad \text{and} \quad |\psi_n(y)| \le C(1+\theta_\nu(|y|))^s. \tag{2.9}$$

For all $A > 0$, let us introduce

$$I_A^N = \inf_{\substack{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N}) \\ \int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|)+\theta_\nu(|y|))\mathrm{d}\pi(x,y)\le A}} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y). \tag{2.10}$$

Then, there exists $A_0 > 0$ such that for all $A \ge A_0$, I_A^N is finite and is a minimum. Moreover, for all $A \ge A_0$, there exists a minimizer π_A^N for Problem (2.10) such that $\pi_A^N = \sum_{k=1}^K p_k \delta_{x_k,y_k}$, for some $0 < K \le 2N+2$, with $p_k \ge 0$, $x_k \in \Sigma_\mu$ and $y_k \in \Sigma_\nu$ for all $1 \le k \le K$.

Remark 2.1. *Let us make here a few remarks:*

(i) When I defined by (2.4) is finite and

$$A_0' = \int_{\mathcal{X}} \theta_\mu(|x|)\mathrm{d}\mu(x) + \int_{\mathcal{Y}} \theta_\nu(|y|)\mathrm{d}\nu(y) < \infty,$$

we have for all $A \geq A_0'$, $I_A^N \leq I < \infty$.

(ii) When the functions ϕ_m and ψ_n are bounded continuous (which holds automatically when $\mathcal{X}$ and $\mathcal{Y}$ are compact), Assumption (2.9) is obviously satisfied.

(iii) When $\mathcal{X}$ and $\mathcal{Y}$ are compact sets, we can then take the positive constant $C = \max_{1\leq n\leq N}(\max(\|\phi_n\|_\infty, \|\psi_n\|_\infty))$ and $\theta_\mu = \theta_\nu = 0$, and therefore we get for all $A > 0$, $I_A^N = I^N$ with

$$I^N = \inf_{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y).$$

(iv) An alternative statement of Theorem 2.3 that avoids imposing the constraint $\int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|) + \theta_\nu(|y|))\mathrm{d}\pi(x,y) \leq A$ can be obtained under stronger assumptions on the test functions and on the cost. More precisely, such a result can be obtained if the test functions are assumed to be compactly supported. The precise statement of this result is given in Section A.2.1 of Appendix A.2.

Proof of Theorem 2.3. Let us introduce the function

$$\Lambda : \begin{cases} \mathcal{X}\times\mathcal{Y} & \to & \mathbb{R}^{2N+2} \\ (x,y) & \mapsto & \begin{pmatrix} \phi(x) \\ \psi(y) \\ 1 \\ c(x,y) \end{pmatrix} \end{cases} \tag{2.11}$$

and let us denote by Λ_i the i^{th} component of Λ for all $1 \leq i \leq 2N+2$. By assumption there exists $\gamma \in \Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})$ such that $\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\gamma(x,y) < \infty$. We apply Theorem 2.1 with $N_0 = 2N+2$ and get that there exist $K \in \{1,\ldots,2N+2\}$, $x_1,...,x_K \in \mathcal{X}$, $y_1,...,y_K \in \mathcal{Y}$ and weights $w_1,...,w_K \in \mathbb{R}_+^*$ such that

$$\int_{\mathcal{X}\times\mathcal{Y}} \Lambda(x,y)\mathrm{d}\gamma(x,y) = \sum_{k=1}^K w_k\Lambda(x_k,y_k). \tag{2.12}$$

Denoting by $\widetilde{\gamma} := \sum_{k=1}^K w_k\delta_{x_k,y_k}$, we have that

$$\int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|) + \theta_\nu(|y|))\,\mathrm{d}\tilde{\gamma}(x,y) < \infty.$$

We thus get that, for all $A \geq A_0 := \int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|) + \theta_\nu(|y|))\mathrm{d}\tilde{\gamma}(x,y)$, I_A^N is finite, since we have $\tilde{\gamma} \in \Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})$.

Let us now assume that $A \geq A_0$ and let us prove that this infimum is a minimum. Let $(\pi_l)_{l\in\mathbb{N}}$ be a minimizing sequence for the minimization problem (2.10). We first prove the tightness of this sequence. For $M_1, M_2 > 0$, let us introduce the compact sets

$$\mathcal{K}_1 = \{x \in \mathcal{X}, \text{ s.t. } |x| \leq M_1\}, \ \mathcal{K}_2 = \{y \in \mathcal{Y}, \text{ s.t. } |y| \leq M_2\}.$$

Then, we have

$$\pi_l((\mathcal{K}_1 \times \mathcal{K}_2)^c) = \int_{\mathcal{X}\times\mathcal{Y}} \mathbf{1}_{(x,y)\notin\mathcal{K}_1\times\mathcal{K}_2} \mathrm{d}\pi_l(x,y) \leq \int_{\mathcal{X}\times\mathcal{Y}} \mathbf{1}_{x\notin\mathcal{K}_1} + \mathbf{1}_{y\notin\mathcal{K}_2} \mathrm{d}\pi_l(x,y)$$
$$\leq \int_{\mathcal{X}\times\mathcal{Y}} \frac{\theta_\mu(|x|)}{\theta_\mu(M_1)} + \frac{\theta_\nu(|y|)}{\theta_\nu(M_2)} \mathrm{d}\pi_l(x,y) \leq \frac{A}{\min(\theta_\mu(M_1), \theta_\nu(M_2))},$$

which implies the tightness of the sequence $(\pi_l)_{l\in\mathbb{N}}$. We can thus extract a subsequence that weakly converges. For notational simplicity, we still denote $(\pi_l)_{l\in\mathbb{N}}$ this subsequence, and there exists $\pi_\infty \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ such that $\pi_l \xrightarrow[l\to\infty]{} \pi_\infty$.

Since c and $\Theta : \mathcal{X}\times\mathcal{Y} \ni (x,y) \mapsto \theta_\mu(|x|) + \theta_\nu(|y|)$ are non-negative lower semi-continuous functions, using [291][Lemma 1.6], we have

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi^\infty(x,y) \leq \liminf_{l\to+\infty} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi_l(x,y) = I_A^N.$$
$$\int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|) + \theta_\nu(|y|))\, d\pi_\infty(x,y) \leq \liminf_{l\to+\infty} \int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|) + \theta_\nu(|y|))\, d\pi_l(x,y) \leq A.$$

Besides, using (2.9), we obtain that for all $1 \leq m, n \leq N$,

$$\max\left(\int_{\mathcal{X}\times\mathcal{Y}} |\phi_m(x)|\, d\pi_l(x,y), \int_{\mathcal{X}\times\mathcal{Y}} |\psi_n(y)|\, d\pi_l(x,y)\right) \leq C(1+A). \tag{2.13}$$

Therefore, we get from (2.13) and the continuity of ϕ_m and ψ_n that

$$\int_{\mathcal{X}\times\mathcal{Y}} \phi_m(x)\, d\pi^\infty(x,y) = \lim_{l\to+\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi_m(x)\, d\pi_l(x,y) = \bar{\mu}_m,$$
$$\int_{\mathcal{X}\times\mathcal{Y}} \psi_n(x)\, d\pi^\infty(x,y) = \lim_{l\to+\infty} \int_{\mathcal{X}\times\mathcal{Y}} \psi_n(x)\, d\pi_l(x,y) = \bar{\nu}_n.$$

This shows that π_∞ satisfies the constraints of Problem (2.10) and thus that

$$I_A^N \leq \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi_\infty(x,y).$$

Thus, $I_A^N = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi_\infty(x,y)$ and π_∞ is a minimizer of Problem (2.10).

Last, we apply Theorem 2.1 to the measure π_∞ and the application Λ defined in (2.11) and get the existence of π_A^N. □

Example 2.1 below shows that the MCOT problem may not be a minimum if we remove the constraint $\int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|) + \theta_\nu(|y|)) \mathrm{d}\pi(x,y) \leq A$.

Example 2.1. *Let*

$$c : \left\{ \begin{array}{ccc} \mathbb{R}\times\mathbb{R} & \to & \mathbb{R}_+ \\ (x,y) & \mapsto & (x-y)^2 + \varphi(|x|) + \varphi(|y|), \end{array} \right.$$

where for $r \in \mathbb{R}_+$, $\varphi(r) = \mathbf{1}_{0\leq r\leq 1} r + \mathbf{1}_{1<r} e^{1-r}$. Let us consider the MCOT problem with the test functions $\phi_1 = \psi_1 = x \mapsto x$,

$$I = \inf_{\substack{\pi\in\mathcal{P}(\mathbb{R}\times\mathbb{R}) \\ \int_{\mathbb{R}} x \mathrm{d}\pi(x,y) = 1 \\ \int_{\mathbb{R}} y \mathrm{d}\pi(x,y) = 1}} \left\{ \int_{\mathbb{R}\times\mathbb{R}} c(x,y) \mathrm{d}\pi(x,y) \right\}.$$

The sequence defined for $l \in \mathbb{N}^*$ *by* $\pi_l = \left(1 - \frac{1}{l}\right)\delta_{(0,0)} + \frac{1}{l}\delta_{(l,l)}$ *is a minimizing sequence since* $\int_{\mathbb{R}\times\mathbb{R}} x \mathrm{d}\pi_l(x,y) = \int_{\mathbb{R}\times\mathbb{R}} y \mathrm{d}\pi_l(x,y) = 1$, $c \geq 0$ *and*

$$\int_{\mathbb{R}\times\mathbb{R}} c(x,y)\mathrm{d}\pi_l(x,y) = \frac{2}{l}e^{1-l} \xrightarrow[l\to\infty]{} 0.$$

Hence, $I = 0$. *Now, since* $\varphi(r) > 0$ *for* $r > 0$, *the only probability measure* $\pi \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ *such that* $\int c\mathrm{d}\pi = 0$ *is* $\delta_{(0,0)}$. *Since this probability measure does not satisfy the constraints* ($\int_{\mathbb{R}\times\mathbb{R}} x\mathrm{d}\delta_{(0,0)}(x,y) = \int_{\mathbb{R}\times\mathbb{R}} y\mathrm{d}\delta_{(0,0)}(x,y) = 0$), *this shows that* I *is not a minimum.*

Let us also note here that the test functions $(\phi_m)_{1\leq m\leq N}$ and $(\psi_n)_{1\leq n\leq N}$ are needed to be continuous to guarantee the existence of a minimum in Theorem 2.3 as Example 2.2 shows.

Example 2.2. *Let* $\mathcal{X} = \mathcal{Y} = [0,1]$, $\mathrm{d}\nu(x) = \left(\frac{1}{2}\mathbf{1}_{(0,\frac{1}{2})}(x) + \frac{3}{2}\mathbf{1}_{(\frac{1}{2},1)}(x)\right)\mathrm{d}x$, $\mathrm{d}\mu(x) = \mathrm{d}x$ *and* $c(x,y) = (y-x)^2$. *Let* $N = 4$, $\phi_1 = \mathbf{1}_{[0,\frac{1}{4}]}$, $\phi_m = \mathbf{1}_{(\frac{m-1}{4},\frac{m}{4}]}$ *for* $2 \leq m \leq 4$ *and* $\psi_m = \phi_m$ *for* $1 \leq m \leq 4$, *so that*

$$\bar{\mu}_1 = \bar{\mu}_2 = \bar{\mu}_3 = \bar{\mu}_4 = \frac{1}{4},\ \bar{\nu}_1 = \bar{\nu}_2 = \frac{1}{8} \text{ and } \bar{\nu}_3 = \bar{\nu}_4 = \frac{3}{8}.$$

For $l \in \mathbb{N}$, $l > 4$, *let*

$$\gamma_l = \frac{1}{8}\delta_{\frac{1}{8},\frac{1}{8}} + \frac{1}{8}\delta_{\frac{1}{4}-\frac{1}{l},\frac{1}{4}+\frac{1}{l}} + \frac{1}{4}\delta_{\frac{1}{2}-\frac{1}{l},\frac{1}{2}+\frac{1}{l}} + \frac{1}{8}\delta_{\frac{5}{8},\frac{5}{8}} + \frac{1}{8}\delta_{\frac{3}{4}-\frac{1}{l},\frac{3}{4}+\frac{1}{l}} + \frac{1}{4}\delta_{\frac{7}{8},\frac{7}{8}}. \tag{2.14}$$

For all $l > 4$, γ_l *satisfies the constraints of the MCOT problem, and*

$$\int_0^1\int_0^1 |x-y|^2\mathrm{d}\gamma_l(x,y) = \left(\frac{1}{8}+\frac{1}{4}+\frac{1}{8}\right)\frac{4}{l^2} = \frac{2}{l^2} \xrightarrow[l\to+\infty]{} 0.$$

Thus, the infimum value of the associated MCOT problem is 0. *Now, let* $\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ *be such that* $\int c\mathrm{d}\pi = 0$. *We have* $\pi(\{(x,y) \in \mathcal{X}\times\mathcal{Y} : y = x\}) = 1$ *and thus*

$$\forall m,\ \int_{\mathcal{X}\times\mathcal{Y}} \phi_m(x)\,\mathrm{d}\pi(x,y) = \int_{\mathcal{X}\times\mathcal{Y}} \phi_m(y)\,\mathrm{d}\pi(x,y).$$

Therefore, we cannot have the left hand side equal to $\bar{\mu}_m$ *and the right hand side equal to* $\bar{\nu}_m$, *which shows that there does not exist any minimizer to the MCOT problem.*

2.3.2 Multimarginal and martingale OT problem

In this section, two important extensions of the previous problem are introduced, the multimarginal problem and the martingale problem. As for Problem (2.10), several formulations and refinements can be established. We only keep here the more general ones for conciseness.

2.3.2.1 Multimarginal problem

The propositions introduced until now for two marginal laws can be extended to an arbitrary (finite) number of marginal laws. The proof can be straightforwardly adapted from the one of Theorem 2.3. For all $1 \leq i \leq M$, we consider $\mathcal{X}_i = \mathbb{R}^{d_i}$ with $d_i \in \mathbb{N}^*$ or more generally a G_δ-set $\mathcal{X}_i \subset \mathbb{R}^{d_i}$. We consider M probability measures

$\mu_1 \in \mathcal{P}(\mathcal{X}_1), ..., \mu_M \in \mathcal{P}(\mathcal{X}_M)$ and a l.s.c. cost function $c : \mathcal{X}_1 \times ... \times \mathcal{X}_M \to \mathbb{R}_+ \cup \{\infty\}$. We consider the following multimarginal optimal transport problem

$$I = \inf_{\pi \in \Pi(\mu_1,...,\mu_M)} \left\{ \int_{\mathcal{X}_1 \times ... \times \mathcal{X}_M} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ..., x_M) \right\}, \tag{2.15}$$

where $\Pi(\mu_1, ..., \mu_M) = \{\pi \in \mathcal{P}(\mathcal{X}_1 \times ... \times \mathcal{X}_M) \text{ s.t.} \forall 1 \leq i \leq M, \int_{\mathcal{X}_i} \mathrm{d}\pi = \mathrm{d}\mu_i\}$.

In order to build the moments constrained optimal transport problem, we introduce, for each i, $N_i \in \mathbb{N}^*$ test functions $(\phi_n^i)_{1 \leq n \leq N_i} \in L^1(\mathcal{X}_i, \mu_i; \mathbb{R})^{N_i}$. We say that this set of test functions is admissible for $(\mu_1, \ldots, \mu_M, c)$ if there exists $\gamma \in \mathcal{P}(\mathcal{X}_1 \times ... \times \mathcal{X}_M)$ such that

$$\forall i \in \{1, \ldots, M\}, \forall n \in \{1, \ldots, N_i\}, \int_{\mathcal{X}_1 \times ... \times \mathcal{X}_M} \phi_n^i(x_i) \mathrm{d}\gamma(x_1, \ldots, x_M) = \int_{\mathcal{X}_i} \phi_n^i(x) \mathrm{d}\mu_i(x)$$

and $\int_{\mathcal{X}_1 \times ... \times \mathcal{X}_M} c(x_1, \ldots, x_M) \mathrm{d}\gamma(x_1, \ldots, x_M) < \infty$. We can now state the analog of Theorem 2.3 for the multimarginal case.

Theorem 2.4. *For $i \in \{1, \ldots, M\}$, let $\mu_i \in \mathcal{P}(\mathcal{X}_i)$ and $\Sigma_{\mu_i} \subset \mathcal{X}_i$ a Borel set such that $\mu_i(\Sigma_{\mu_i}) = 1$. We assume that $c : \mathcal{X}_1 \times ... \times \mathcal{X}_M \to \mathbb{R}^+ \cup \{\infty\}$ is a l.s.c. cost function, and that the set of test functions $\phi_n^i \in L^1(\mathcal{X}_i, \mu_i; \mathbb{R})$ for $i \in \{1, \ldots, M\}$ and $n \in \{1, \ldots, N_i\}$ is admissible for $(\mu_1, \ldots, \mu_M, c)$. We make the following assumptions.*

1. *For all i and n, the function ϕ_n^i is continuous.*

2. *For all i, there exists $\theta_i : \mathbb{R}_+ \to \mathbb{R}_+$ a non-decreasing continuous function such that $\theta_i(r) \underset{r \to +\infty}{\to} +\infty$ and such that there exist $C > 0$ and $0 < s < 1$ such that for all $1 \leq n \leq N_i$, we have*

$$\forall x \in \mathcal{X}_i, \ |\phi_n^i(x)| \leq C(1 + \theta_i(|x|))^s. \tag{2.16}$$

We note $\mathbf{N} = (N_1, \ldots, N_M)$, $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_M$ and consider the following problem

$$I_A^{\mathbf{N}} = \inf_{\substack{\pi \in \mathcal{P}(\mathcal{X}) \\ \forall i,n, \int_{\mathcal{X}} \phi_n^i(x_i) \mathrm{d}\pi(x_1,...,x_M) \\ = \int_{\mathcal{X}_i} \phi_n^i(x) \mathrm{d}\mu_i(x) \\ \int_{\mathcal{X}} \sum_{i=1}^M \theta_i(|x_i|) \mathrm{d}\pi(x_1,...,x_M) \leq A}} \left\{ \int_{\mathcal{X}} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ...x_M) \right\}. \tag{2.17}$$

Then, there exists $A_0 > 0$ such that for all $A \geq A_0$, $I_A^{\mathbf{N}}$ is finite and is a minimum. Moreover, for all $A \geq A_0$, there exists a minimizer $\pi_A^{\mathbf{N}}$ for the problem (2.10) such that $\pi_A^{\mathbf{N}} = \sum_{k=1}^K p_k \delta_{x_1^k, ..., x_M^k}$, for some $0 < K \leq \sum_{i=1}^M N_i + 2$, with $p_k \geq 0$ and $x_i^k \in \Sigma_{\mu_i}$ for all $1 \leq i \leq M$ and $1 \leq k \leq K$.

Remark 2.2. *An interesting point to remark in Theorem 2.4 is that the number of weighted points of the discrete measure $\pi_A^{\mathbf{N}}$ is linear with respect to the number of moment constraints. In particular, if we take the same number of moments $N_i = N$ for each marginal, the number of weighted points is equal to $2 + MN$ and thus grows linearly with respect to M. Since each point has dM coordinates, the dimension of the discrete measure is in $O(M^2)$. For this reason, the development of algorithms for minimizing $\pi_A^{\mathbf{N}}$ by using finite discrete measures may be a way to avoid the curse of dimensionality when M is getting large.*

We make here a specific focus on the multimarginal optimal transport problem which arises in quantum chemistry applications [304, 108]. In this particular case, the multi-marginal optimal transport of interest reads as (2.15), with $\mathcal{X}_1 = \cdots \mathcal{X}_M = \mathbb{R}^3$, $N_1 = \cdots = N_M = N$ for some $N \in \mathbb{N}^*$, $\mu_1 = \cdots = \mu_M = \mu$ for some $\mu \in \mathcal{P}(\mathbb{R}^3)$ and c is given by the Coulomb cost

$$c(x_1, \cdots, x_M) := \sum_{1 \leq i < j \leq M} \frac{1}{|x_i - x_j|}.$$

The integer M represents here the number of electrons in the system of interest. The inherent symmetries of the system yield interesting consequences for the MCOT problem (2.17), which are summarized in the following proposition.

Proposition 2.5. *Let $M \in \mathbb{N}^*$, $N \in \mathbb{N}^*$, $\mu \in \mathcal{P}(\mathcal{X})$ and $\Sigma_\mu \subset \mathcal{X}$ a Borel set such that $\mu(\Sigma_\mu) = 1$. We assume that $c : \mathcal{X}^M \to \mathbb{R}^+ \cup \{\infty\}$ is a symmetric l.s.c. cost function. More precisely, we denote by $\mathcal{S}_M$ the set of permutations of the set $\{1, \cdots, M\}$ and assume that*

$$\forall \sigma \in \mathcal{S}_M, \quad c(x_{\sigma(1)}, \cdots, x_{\sigma(M)}) = c(x_1, \cdots, x_M), \quad \text{for almost all } x_1, \cdots, x_M \in \mathcal{X}.$$

For all $1 \leq n \leq N$, let $\phi_n \in L^1(\mathcal{X}, \mu; \mathbb{R})$. We define $\phi_n^i := \phi_n$ for all $1 \leq i \leq M$ and assume the set of test functions ϕ_n^i for $n \in \{1, \ldots, N\}$ and $i \in \{1, \cdots, M\}$ is admissible for $(\mu, \ldots, \mu, c)$. We make the following assumptions.

1. *For all n, the function ϕ_n is continuous.*
2. *There exists $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ a non-decreasing continuous function such that $\theta(r) \underset{r \to +\infty}{\longrightarrow} +\infty$ and such that there exist $C > 0$ and $0 < s < 1$ such that for all $1 \leq n \leq N$, we have*

$$\forall x \in \mathcal{X}, \ |\phi_n(x)| \leq C(1 + \theta(|x|))^s. \tag{2.18}$$

We consider the following problem

$$I_A^N = \inf_{\substack{\pi \in \mathcal{P}(\mathcal{X}^M) \\ \forall n, i, \int_{\mathcal{X}^M} \phi_n(x_i) \mathrm{d}\pi(x_1, \ldots, x_M) \\ = \int_{\mathcal{X}} \phi_n(x) \mathrm{d}\mu(x) \\ \int_{\mathcal{X}^M} \sum_{i=1}^M \theta(|x_i|) \mathrm{d}\pi(x_1, \ldots, x_M) \leq A}} \left\{ \int_{\mathcal{X}} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ...x_M) \right\}. \tag{2.19}$$

Then,

$$I_A^N = \inf_{\substack{\pi \in \mathcal{P}(\mathcal{X}^M) \\ \forall n, \int_{\mathcal{X}^M} \left(\frac{1}{M} \sum_{i=1}^M \phi_n(x_i) \right) \mathrm{d}\pi(x_1, \ldots, x_M) \\ = \int_{\mathcal{X}} \phi_n(x) \mathrm{d}\mu(x) \\ \int_{\mathcal{X}^M} \sum_{i=1}^M \theta(|x_i|) \mathrm{d}\pi(x_1, \ldots, x_M) \leq A}} \left\{ \int_{\mathcal{X}} c(x_1, ..., x_M) \mathrm{d}\pi(x_1, ...x_M) \right\}, \tag{2.20}$$

and there exists $A_0 > 0$ such that for all $A \geq A_0$, I_A^N is finite and is a minimum. Moreover, for all $A \geq A_0$, there exists a minimizer π_A^N for the problem (2.20) such that $\pi_A^N = \sum_{k=1}^K p_k \delta_{x_1^k, \ldots, x_M^k}$, for some $0 < K \leq N + 2$, with $p_k \geq 0$ and $x_i^k \in \Sigma_\mu$ for all $1 \leq i \leq M$ and $1 \leq k \leq K$. Besides, the symmetric measure

$$\pi_{\mathrm{sym,A}}^N := \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} \sum_{k=1}^K p_k \delta_{x_{\sigma(1)}^k, \ldots, x_{\sigma(M)}^k} \tag{2.21}$$

is a minimizer to (2.19) and (2.20).

Proof. It is obvious that the right hand side of (2.20) is smaller than the right hand side of (2.19). By using the same arguments as in the proof of Theorem 2.3, there exists $A_0 > 0$ such that for all $A \geq A_0$ the infimum of (2.20) is finite, is a minimum that is attained by some discrete probability measure $\pi_A^N = \sum_{k=1}^K p_k \delta_{x_1^k,...,x_M^k}$, for some $0 < K \leq N+2$ with $x_i^k \in \Sigma_\mu$ for all $1 \leq i \leq M$ and $1 \leq k \leq K$. Then, since c is symmetric and the set of constraints is also symmetric, we get that $\pi^N_{\mathrm{sym,A}}$ also realizes the minimum. Besides, it satisfies $\int_{\mathcal{X}^M} \phi_n(x_i) \mathrm{d}\pi^N_{\mathrm{sym,A}}(x_1, ..., x_M) = \int_{\mathcal{X}} \phi_n(x) \mathrm{d}\mu(x)$ for all n, i, which shows that it is also the minimizer of (2.19). □

Remark 2.3. *Proposition 2.5 is particularly interesting for the design of numerical schemes for the resolution of multimarginal optimal transport problems with Coulomb cost arising in quantum chemistry applications. Indeed, the latter read as (2.19) and the number of charged points of the minimizer π_A^N of (2.20) only scales at most like $N+2$, and that the dimension of the optimal discrete measure is in $dM(N+2)$. This result states that such formulation of the multimarginal optimal transport problem does not suffer from the curse of dimensionality. Let us mention that this result is close in spirit to the recent work [143], where multimarginal optimal transport problems with Coulomb cost are studied on finite state spaces.*

2.3.2.2 Martingale OT problem

In this paragraph, we assume $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ with $d \in \mathbb{N}^*$, and consider two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} |y| \mathrm{d}\nu(y) < \infty$$

and μ is lower than ν for the convex order, i.e.

$$\int_{\mathbb{R}^d} \varphi(x) \mathrm{d}\mu(x) \leq \int_{\mathbb{R}^d} \varphi(y) \mathrm{d}\nu(y), \tag{2.22}$$

for any convex function $\varphi : \mathbb{R}^d \to \mathbb{R}$ non-negative or integrable with respect to μ and ν. This latter condition is equivalent, by Strassen's theorem [311], to the existence of a martingale coupling between μ and ν, i.e.

$$\exists \pi \in \Pi(\mu, \nu),\ \forall x \in \mathbb{R}^d,\ \int_{\mathbb{R}^d} y \mathrm{d}\pi(x, y) = x.$$

The original martingale optimal transport consists then in the minimization problem

$$\inf_{\substack{\pi \in \Pi(\mu,\nu) \\ \forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} y \mathrm{d}\pi(x,y) = x}} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \mathrm{d}\pi(x, y) \right\}, \tag{2.23}$$

with $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ being a l.s.c. cost function. This problem has recently got a great attention in mathematical finance since the work of Beiglböck et al. [27], because it is related to the calculation of model-independent option price bounds on an arbitrage free market.

We consider a set of test functions $(\phi_m)_{1 \leq m \leq N} \in L^1(\mathbb{R}^d, \mu; \mathbb{R})^N$ and $(\psi_n)_{1 \leq n \leq N} \in L^1(\mathbb{R}^d, \nu; \mathbb{R})^N$, and the following problem:

$$I^N = \inf_{\substack{\pi \in \Pi(\mu,\nu;(\phi_m)_{1 \leq m \leq N}, (\psi_n)_{1 \leq n \leq N}) \\ \forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} y \mathrm{d}\pi(x,y) = x}} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \mathrm{d}\pi(x, y) \right\}.$$

Suppose for simplicity that there exist some minimizer to this problem π^*. Then, by using Theorem 5.1 in Beiglböck and Nutz [34] that is an extension of Tchakaloff's theorem to the martingale case, there exists a probability measure $\tilde{\pi}^*$ weighting at most $(d+2N+2)^2$ points such that $\tilde{\pi}^* \in \Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})$,

$$\forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} y \mathrm{d}\tilde{\pi}^*(x,y) = x$$

and

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mathrm{d}\tilde{\pi}^*(x,y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mathrm{d}\pi^*(x,y) = I^N.$$

However, the minimization problem for I^N still has the martingale constraints. To get a problem that is similar to the MCOT, we then relax in addition the martingale constraint. This constraint is equivalent to have

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x)(y-x)\mathrm{d}\pi(x,y) = 0,$$

for all bounded measurable functions $f : \mathbb{R}^d \to \mathbb{R}$, and also for all function $f : \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} |xf(x)|\mathrm{d}\mu(x) < \infty$. Then, it is natural to consider N' test functions $\chi_l : \mathbb{R}^d \to \mathbb{R}$, $1 \leq l \leq N'$, such that

$$\int_{\mathbb{R}^d} |x\chi_l(x)|\mathrm{d}\mu(x) < \infty, \tag{2.24}$$

and then to consider the following minimization problem

$$I^{N,N'} = \inf_{\substack{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N}) \\ \forall l, \int_{\mathbb{R}^d\times\mathbb{R}^d} y\chi_l(x)\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^d} x\chi_l(x)\mathrm{d}\mu(x)}} \left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mathrm{d}\pi(x,y)\right\}. \tag{2.25}$$

We will say that the test functions $(\phi_m)_{1\leq m\leq N}$, $(\psi_n)_{1\leq n\leq N}$ and $(\chi_l)_{1\leq l\leq N'}$ are admissible for the martingale problem of (μ,ν,c) if $I^{N,N'} < \infty$. Similarly to Theorem 2.3, we get the following result.

Theorem 2.6. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$, $\nu \in \mathcal{P}(\mathbb{R}^d)$ and $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{+\infty\}$ a l.s.c. function. Let $\Sigma_\mu, \Sigma_\nu \subset \mathbb{R}^d$ be Borel sets such that $\mu(\Sigma_\mu) = \nu(\Sigma_\nu) = 1$. Let $N \in \mathbb{N}^*$ and let $(\phi_m)_{1\leq m\leq N} \in L^1(\mathbb{R}^d,\mu;\mathbb{R})^N$, $(\psi_n)_{1\leq n\leq N} \in L^1(\mathbb{R}^d,\nu;\mathbb{R})^N$ and $(\chi_l)_{1\leq l\leq N'}$ satisfying (2.24) be an admissible set of test functions for the martingale problem of (μ,ν,c). We make the following assumptions.*

1. *For all $n \in \{1,\dots,N\}$, $l \in \{1,\dots,N'\}$, the functions ϕ_n, ψ_n and χ_l are continuous.*

2. *There exist $\theta_\mu : \mathbb{R}_+ \to \mathbb{R}_+$ and $\theta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ two non-negative non-decreasing continuous functions such that $\theta_\mu(r) \underset{r\to+\infty}{\longrightarrow} +\infty$ and $\theta_\nu(r) \underset{r\to+\infty}{\longrightarrow} +\infty$, and such that there exist $C > 0$ and $0 < s < 1$ such that for all $1 \leq n \leq N$, $1 \leq l \leq N'$, and all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$,*

$$|\phi_n(x)| + |\psi_n(y)| + |y\chi_l(x)| \leq C(1 + \theta_\mu(|x|) + \theta_\nu(|y|))^s. \tag{2.26}$$

For all $A > 0$, let us introduce

$$I_A^{N,N'} = \inf_{\substack{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N}) \\ \forall l, \int_{\mathbb{R}^d\times\mathbb{R}^d} y\chi_l(x)\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^d} x\chi_l(x)\mathrm{d}\mu(x) \\ \int_{\mathbb{R}^d\times\mathbb{R}^d}(\theta_\mu(|x|)+\theta_\nu(|y|))\mathrm{d}\pi(x,y)\leq A}} \left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mathrm{d}\pi(x,y)\right\}. \tag{2.27}$$

Then, there exists $A_0 > 0$ such that for all $A \geq A_0$, $I_A^{N,N'}$ is finite and is a minimum. Moreover, for all $A \geq A_0$, there exists a minimizer $\pi_A^{N,N'}$ for Problem (2.27) such that $\pi_A^{N,N'} = \sum_{k=1}^{K} p_k \delta_{x_k,y_k}$, for some $0 < K \leq 2N + N' + 2$, with $p_k \geq 0$, $x_k \in \Sigma_\mu$ and $y_k \in \Sigma_\nu$ for all $1 \leq k \leq K$.

The proof of Theorem 2.6 follows exactly the same line as the proof of Theorem 2.3, since the relaxation of the martingale moment constraints only brings new moment constraints. Let us stress that the minimizer $\pi_A^{N,N'}$ does not satisfy in general the martingale constraint. Also, we do not impose in Theorem 2.6 to have (2.22), i.e. μ smaller than ν for the convex order. In fact, the admissibility condition already ensures that $I^{N,N'} < \infty$ and thus, by using Theorem 2.1 that $I_A^{N,N'} < \infty$ for A large enough. Nonetheless, if we assume in addition that μ smaller than ν for the convex order and that I, the infimum of Problem 2.23, is finite, then we have $I_A^{N,N'} < \infty$ and $I_A^{N,N'} \leq I$ for any $A \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} (\theta_\mu(|x|) + \theta_\nu(|y|)) \mathrm{d}\pi^1(x,y)$, where $\pi^1 \in \Pi(\mu,\nu)$ is such that $\int_{\mathbb{R}^d} y \mathrm{d}\pi^1(x,y) = x$ and $\int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) \mathrm{d}\pi^1(x,y) \leq I + 1$.

2.4 Convergence of the MCOT problem towards the OT problem

The aim of this section is to prove that when the number of test functions $N \to +\infty$, the minimizer of the MCOT problem converges towards a minimizer of the OT problem, under appropriate assumptions and up to the extraction of a subsequence.

2.4.1 Convergence for two-marginal (or multi-marginal) Optimal Transport problems

Let us consider two sequences of continuous real-valued test functions $(\phi_m)_{m \in \mathbb{N}^*}$ and $(\psi_n)_{n \in \mathbb{N}^*}$ defined on $\mathcal{X}$ (resp. $\mathcal{Y}$) and make the following assumptions.

Let us first assume that there exist continuous non-decreasing non-negative functions $\theta_\mu : \mathbb{R}_+ \to \mathbb{R}_+$ and $\theta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\theta_\mu(|x|) \xrightarrow[|x| \to +\infty]{} +\infty \quad \text{and} \quad \theta_\nu(|y|) \xrightarrow[|y| \to +\infty]{} +\infty \tag{2.28}$$

and

$$\int_{\mathcal{X}} \theta_\mu(|x|) \mathrm{d}\mu(x) < \infty \quad \text{and} \quad \int_{\mathcal{Y}} \theta_\nu(|y|) \mathrm{d}\nu(y) < \infty. \tag{2.29}$$

In the sequel, we set

$$A_0 := \int_{\mathcal{X}} \theta_\mu(|x|) \mathrm{d}\mu(x) + \int_{\mathcal{Y}} \theta_\nu(|y|) \mathrm{d}\nu(y). \tag{2.30}$$

We assume moreover that there exist sequences $(s_m^\mu)_{m \in \mathbb{N}^*}, (s_n^\nu)_{n \in \mathbb{N}^*} \in (0,1)^{\mathbb{N}^*}$ and $(C_m^\mu)_{m \in \mathbb{N}^*}, (C_n^\nu)_{n \in \mathbb{N}^*} \in (\mathbb{R}_+^*)^{\mathbb{N}^*}$ such that

$$\forall m \in \mathbb{N}^*, \ \forall x \in \mathcal{X}, \quad |\phi_m(x)| \leq C_m^\mu (1 + \theta_\mu(|x|))^{s_m^\mu}, \tag{2.31}$$

$$\forall n \in \mathbb{N}^*, \ \forall y \in \mathcal{Y}, \quad |\psi_n(y)| \leq C_n^\nu (1 + \theta_\nu(|y|))^{s_n^\nu}. \tag{2.32}$$

Last, we assume that the probability measures μ and ν are fully characterized by their moments:

$$\forall \eta \in \mathcal{P}(\mathcal{X}), \left(\forall m \in \mathbb{N}^*, \int_{\mathcal{X}} \phi_m(x) \mathrm{d}\eta(x) = \bar{\mu}_m \right) \implies \eta = \mu, \tag{2.33}$$

$$\forall \eta \in \mathcal{P}(\mathcal{Y}), \left(\forall n \in \mathbb{N}^*, \int_{\mathcal{Y}} \psi_n(x) \mathrm{d}\eta(x) = \bar{\nu}_n \right) \implies \eta = \nu. \tag{2.34}$$

We consider the optimal cost for the OT problem (2.4) that we restate here for convenience

$$I = \inf_{\pi \in \Pi(\mu,\nu)} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi(x,y) \right\}, \tag{2.35}$$

and for all $N \in \mathbb{N}^*$, we define the N^{th} MCOT problem,

$$I_{A_0}^N = \min_{\substack{\pi \in \Pi(\mu,\nu; (\phi_m)_{1\leq m\leq N}, (\psi_n)_{1\leq n\leq N}) \\ \int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|)+\theta_\nu(|y|)) \mathrm{d}\pi(x,y) \leq A_0}} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi(x,y) \right\}. \tag{2.36}$$

Theorem 2.7. *Let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$ satisfying (2.29) for some continuous non-decreasing functions $\theta_\mu : \mathbb{R}_+ \to \mathbb{R}_+$ and $\theta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying (2.28). Let $c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ a l.s.c. function. Let $(\phi_m)_{m\in\mathbb{N}^*} \subset L^1(\mathcal{X}, \mu; \mathbb{R})$ and $(\psi_n)_{n\in\mathbb{N}^*} \subset L^1(\mathcal{Y}, \nu; \mathbb{R})$ be continuous functions satisfying (2.31), (2.32), (2.33) and (2.34). Let us finally assume that I, defined by (2.35) is finite.*

Then, for all $N \in \mathbb{N}^*$, there exist at least one minimizer for Problem (2.36) and

$$I_{A_0}^N \underset{N\to+\infty}{\longrightarrow} I.$$

Besides, from every sequence $(\pi^N)_{N\in\mathbb{N}^*}$ such that for all N, $\pi^N \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ is a minimizer for (2.36), one can extract a subsequence which converges towards a minimizer $\pi^\infty \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ to problem (2.35).

Proof. From Theorem 2.3 and Remark 2.1 (i), We know that there exists at least one minimizer $\pi^N \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ to (2.36). Since we have

$$\forall N, \int_{\mathcal{X}\times\mathcal{Y}} (\theta_\mu(|x|) + \theta_\nu(|y|)) \mathrm{d}\pi^N(x,y) \leq A_0,$$

and (2.28), we get that the sequence $(\pi^N)_{N\in\mathbb{N}^*}$ is tight. Thus, up to the extraction of a subsequence, still denoted $(\pi^N)_{N\in\mathbb{N}^*}$ for the sake of simplicity, there exists a measure $\pi^\infty \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ such that $\pi^N \underset{N\to\infty}{\rightharpoonup} \pi^\infty$ tightly. With the same argument as in the proof of Theorem 2.3, we get that for all $m, n \in \mathbb{N}^*$,

$$\int_{\mathcal{X}\times\mathcal{Y}} \phi_m(x) \mathrm{d}\pi^\infty(x,y) = \bar{\mu}_m \quad \text{and} \quad \int_{\mathcal{X}\times\mathcal{Y}} \psi_n(x) \mathrm{d}\pi^\infty(x,y) = \bar{\nu}_n.$$

Then, Properties (2.33) and (2.34) give $\pi^\infty \in \Pi(\mu,\nu)$. Therefore,

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi^\infty(x,y) \geq I. \tag{2.37}$$

On the other hand, note that $(I_{A_0}^N)_{N\in\mathbb{N}}$ is a non-decreasing sequence and that for all $N \in \mathbb{N}^*$, $I_{A_0}^N \leq I$. Thus, there exists $I^\infty \leq I$ such that $I_{A_0}^N \underset{N\to\infty}{\longrightarrow} I^\infty$. Furthermore, since c is a non-negative semi-lower continuous function, using [291][Lemma 1.6], we deduce that

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi^\infty(x,y) \leq \liminf_{N\to+\infty} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi^N(x,y) = I^\infty \leq I.$$

Hence $\int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi^\infty(x,y) = I$ which concludes the proof since $\pi^\infty \in \Pi(\mu,\nu)$. □

Remark 2.4. *Let us make a few remarks:*

(i) A result analogous to Theorem 2.7 can be easily obtained for general multi-marginal optimal transport problems. The extension to martingale optimal transport problems is less obvious and is the object of Section 2.4.2.

(ii) Assuming that $\mathcal{X}$ and $\mathcal{Y}$ are compact subsets of $\mathbb{R}^{d_x}$ and $\mathbb{R}^{d_y}$, a result analogous to Theorem 2.7 that holds without the additional moment constraint and for possibly discontinuous test functions can be proved. More precisely, considering two sequences of bounded measurable real-valued test functions $(\phi_m)_{m\in\mathbb{N}^*} \subset L^\infty(\mathcal{X})$ and $(\psi_n)_{n\in\mathbb{N}^*} \subset L^\infty(\mathcal{Y})$ that satisfy

$$\forall f \in C^0(\mathcal{X}), \inf_{v_N \in \mathrm{Span}\{\phi_m,\, 1\le m\le N\}} \|f - v_N\|_\infty \underset{N\to+\infty}{\longrightarrow} 0 \tag{2.38}$$

and

$$\forall f \in C^0(\mathcal{Y}), \inf_{v_N \in \mathrm{Span}\{\psi_n,\, 1\le n\le N\}} \|f - v_N\|_\infty \underset{N\to+\infty}{\longrightarrow} 0, \tag{2.39}$$

it is easy then to see that the properties (2.33) *and* (2.34) *are satisfied for any $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. The precise statement and proof of this result is given in Section A.2.2 of Appendix A.2.*

(iii) The result of Theorem 2.7 can be seen as a Γ-convergence result, see Braides [55] for an introduction to this theory. Let us define for $\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$, $F_n(\pi) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y)$ if $\pi \in \Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})$ (resp. $F_\infty(\pi) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y)$ if $\pi \in \Pi(\mu,\nu)$) and $F_n(\pi) = +\infty$ (resp. $F(\pi) = +\infty$) otherwise. Let us define $K = \{\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}), \int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|)+\theta_\nu(|y|))\mathrm{d}\pi(x,y) \le A_0\}$. We can then check that on K, the sequence (F_n) Γ-converges to F_∞ by using that $\pi \mapsto \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y)$ is l.s.c. for the weak convergence and properties (2.31), (2.31), (2.33) *and* (2.34), *as in the proof of Theorem 2.7. Then, since K is tight and thus a relatively sequentially compact set, we get the claim by Proposition 1.18 [55].*

2.4.2 Convergence for Martingale Optimal Transport problems

In this subsection, we study the convergence of $I_A^{N,N'}$ defined by (2.25) when the number of test functions for the martingale condition $N' \to +\infty$ towards the following minimization problem:

$$I_A^{N,mg} = \inf_{\substack{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N}) \\ \forall x\in\mathbb{R}^d, \int_{\mathbb{R}^d} y\mathrm{d}\pi(x,y)=x \\ \int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|)+\theta_\nu(|y|))\mathrm{d}\pi(x,y)\le A}} \left\{ \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mathrm{d}\pi(x,y) \right\}. \tag{2.40}$$

This convergence is particularly interesting for the practical application in finance: the marginal laws μ, ν are in general not observed and market data only provide some moments. For $d = 1$, market data give the prices of European put (or call) options that corresponds to $\phi_m(x) = (K_m - x)^+$ and $\psi_n(y) = (K'_m - y)^+$. We consider for simplicity a non-negative underlying asset with zero interest rates. Then, by taking $\theta_\mu(|x|) = \theta_\nu(|x|) = |x|$, we have from the martingale assumption $\int_{\mathcal{X}\times\mathcal{Y}}(|x| + |y|)\mathrm{d}\pi(x,y) = 2S_0$, where $S_0 > 0$ is the current price of the underlying asset. Then, a natural choice would be to take $A_0 = 2S_0$. Therefore, the convergence stated in Proposition 2.8 gives a way to approximate option price bounds

by taking into account that only some moments are known, while the few existing numerical methods for Martingale Optimal Transport in the literature assume that the marginal laws are known [5, 6, 171].

Proposition 2.8. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ lower than $\nu \in \mathcal{P}(\mathbb{R}^d)$ for the convex order and $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{+\infty\}$ a l.s.c. function. We assume $|x| \leq \theta_\mu(|x|)$, $|y| \leq \theta_\nu(|y|)$ and suppose $A_0 < \infty$ with A_0 defined by (2.30). We assume that the test functions $(\chi_l, l \in \mathbb{N}^*)$ are bounded and such that for any function $f : \mathbb{R}^d \to \mathbb{R}$ continuous with compact support, we have*

$$\inf_{g \in \mathrm{Span}\{\chi_l,\, 1 \leq l \leq N'\}} \|f - g\|_\infty \underset{N' \to +\infty}{\longrightarrow} 0. \tag{2.41}$$

Let the assumptions of Theorem 2.6 hold for any $N' \geq 1$. Then, we have $I_{A_0}^{N,N'} \underset{N' \to +\infty}{\longrightarrow} I_{A_0}^{N,mg} < \infty$.

Proof. Since $A_0 < \infty$, any martingale coupling between μ and ν satisfies the constraints of $I_{A_0}^{N,N'}$. By using Tchakaloff's theorem and the fact that c is finite-valued, we get that $I_{A_0}^{N,N'}$ is finite for any N' and is attained by a measure denoted by $\pi^{N'}$ according to Theorem 2.6. Similarly, using Tchakaloff's theorem for the martingale case, Theorem 5.1 [34], we get that $I_{A_0}^{N,mg} < \infty$. Note that from the inclusion of the constraints, we clearly have $I_{A_0}^{N,N'_1} \leq I_{A_0}^{N,N'_2} \leq I_{A_0}^{N,mg}$ for $N'_1 \leq N'_2$. We can then repeat the arguments in the proof of Theorem 2.7 to get that $(\pi^{N'})$ is tight and any limit π^∞ of a weakly converging subsequence satisfies $I_{A_0}^{N,mg} = \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) \mathrm{d}\pi^\infty(x,y)$.

The only thing to prove is that $\int_{\mathbb{R}^d \times \mathbb{R}^d} (y-x) f(x) d\pi^\infty(x,y) = 0$ for any function $f : \mathbb{R}^d \to \mathbb{R}$ continuous with compact support. Let $\epsilon > 0$. By assumption, there exists $M \in \mathbb{N}^*$ and $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$ such that $\sup_{x \in \mathbb{R}^d} |f(x) - \sum_{l=1}^M \lambda_l \chi_l(x)| \leq \epsilon$. Therefore, for $N' \geq M$, we have

$$\begin{aligned}\left| \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x)(y-x) d\pi^{N'}(x,y) \right| &= \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(f(x) - \sum_{l=1}^M \lambda_l \chi_l(x) \right) (y-x) d\pi^{N'}(x,y) \right| \\ &\leq \epsilon \int_{\mathbb{R}^d \times \mathbb{R}^d} |y-x| d\pi^{N'}(x,y) \leq \epsilon A_0,\end{aligned}$$

by using the triangle inequality and the fact that $|x| \leq \theta_\mu(|x|)$, $|y| \leq \theta_\nu(|y|)$. We conclude then easily letting $N' \to \infty$. □

Let us mention that we can obtain using similar arguments that $I_{A_0}^{N,mg}$ and $I_{A_0}^{N,N'}$ converge towards (2.23) as N and N' go to infinity. Note that the convergence of $I_{A_0}^{N,mg}$ is implictly used in the literature on robust finance: it is usually assumed to know marginal laws while in practice market data only provide some moments.

2.5 Rates of convergence for particular sets of test functions

Throughout this section, we assume that

$$\mathcal{X} = \mathcal{Y} = [0,1]$$

and for all $N \in \mathbb{N}^*$, we define the intervals

$$T_1^N = \left[0, \frac{1}{N}\right], \ \forall 2 \leq m \leq N, \ T_m^N = \left(\frac{m-1}{N}, \frac{m}{N}\right]. \tag{2.42}$$

We investigate in this section the rate of convergence of I^N defined by

$$I^N = \inf_{\pi \in \Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\pi(x,y) \right\} \tag{2.43}$$

towards I defined by (2.35), when the test functions are piecewise constant (resp. piecewise linear) on T_m^N. We obtain, under suitable assumptions a convergence rate of $O(1/N)$ (resp. $O(1/N^2)$). This shows, as one may expect, the importance of the choice of test functions to approximate the Optimal Transport problem.

Note that, as studied in Appendix A.2, the compactness of $\mathcal{X}$ and $\mathcal{Y}$ allows to define Problem (2.43) with no inequality constraint (contrary to (2.10)), and that, despite non-continuous test functions, such MCOT problems are well defined and under appropriate assumptions converge towards the OT problem.

2.5.1 Piecewise constant test functions on compact sets

In this section, we assume that the cost function $c : [0,1]^2 \to \mathbb{R}_+$ is Lipschitz:

$$|c(x,y) - c(x',y')| \leq K \max(|x-x'|, |y-y'|). \tag{2.44}$$

We define, for $\pi \in \mathcal{P}([0,1]^2)$, $I(\pi) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y)$ and

$$I = \inf_{\pi\in\Pi(\mu,\nu)} I(\pi). \tag{2.45}$$

We introduce the piecewise constant test functions

$$\forall N \geq 1, 1 \leq m \leq N, \quad \phi_m^N = \mathbf{1}_{T_m^N},$$

and consider the MCOT problem:

$$I^N = \inf_{\pi \in \Pi(\mu,\nu;(\phi_m^N)_{1\leq m\leq N},(\phi_n^N)_{1\leq n\leq N})} \left\{ \int_{[0,1]^2} c(x,y) \mathrm{d}\pi(x,y) \right\}. \tag{2.46}$$

Then, Theorem 2.9 establishes the rate of convergence of the sequence $(I^N)_{N\in\mathbb{N}^*}$ to I as N increases.

Theorem 2.9. *Let $\mu, \nu \in \mathcal{P}([0,1])$ and $c : [0,1]^2 \to \mathbb{R}_+$ a Lipschitz function with Lipschitz constant $K > 0$. Then, for all $N \in \mathbb{N}^*$,*

$$I^N \leq I \leq I^N + \frac{K}{N}. \tag{2.47}$$

Remark 2.5. *Let us note that we are not exactly in the framework of Section 2.4, since the test functions depends on N. However, we have*

$$\mathrm{Span}\left\{\phi_m^N,\, 1 \leq m \leq N\right\} \subset \mathrm{Span}\left\{\phi_m^{2N},\, 1 \leq m \leq 2N\right\}$$

and thus Proposition A.4 gives for any $L \in \mathbb{N}^*$,

$$I^{L2^k} \underset{k\to+\infty}{\longrightarrow} I.$$

Before proving Theorem 2.9, we state a result which bounds the distance between an MCOT optimizer and the minimizer of the OT problem (2.45). We define, for $p \geq 1$, the W_p-Wasserstein distance between $\eta_1, \eta_2 \in \mathcal{P}(\mathbb{R}^d)$ as $W_p^p(\eta_1,\eta_2) = \inf_{\pi\in\Pi(\eta_1,\eta_2)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x_1 - x_2\|_p^p \mathrm{d}\pi(x_1,x_2)$, i.e. we take the $\|\|_p$-norm for W_p.

Proposition 2.10. *Let $p > 1$. Let $\mu \in \mathcal{P}([0,1])$. If $\mu^N \in \mathcal{P}([0,1])$ is such that $\int_0^1 \phi_m^N(x)\mathrm{d}\mu^N(x) = \int_0^1 \phi_m^N(x)\mathrm{d}\mu(x)$ for all $m \in \{1,\dots,N\}$, then*

$$W_p(\mu,\mu^N) \leq \frac{1}{N}.$$

Let us assume besides that the cost function satisfies $c(x,y) = H(y-x)$ with $H : \mathbb{R} \to \mathbb{R}_+$ strictly convex. There exists then a unique minimizer of (2.45) which we denote by π^*.
Let $\pi^N \in \Pi(\mu,\nu;(\phi_m^N)_{1\leq m\leq N},(\phi_n^N)_{1\leq n\leq N})$, μ^N and ν^N the marginal laws of π^N and assume that

$$\int_{[0,1]^2} c(x,y)\mathrm{d}\pi^N(x,y) = \min_{\pi\in\Pi(\mu^N,\nu^N)} \int_{[0,1]^2} c(x,y)\mathrm{d}\pi(x,y).$$

Then, we have $W_p(\pi^N,\pi^*) \leq \frac{2^{1/p}}{N}$, where W_p is defined using the $\|\|_p$ norm on $\mathbb{R}^2$.

Proof. For $\eta \in \mathcal{P}(\mathbb{R})$, we define $F_\eta^{-1}(u) = \inf\{x \in \mathbb{R} : \eta((-\infty,x]) \geq u\}$, that coincides with the usual inverse when F_η is increasing continuous. Let $p > 1$. By Theorem 2.9 [291], we have

$$\begin{aligned} W_p^p(\mu,\mu^N) &= \int_0^1 |F_\mu^{-1}(u) - F_{\mu^N}^{-1}(u)|^p du \\ &= \int_0^{F_\mu(0)} |F_\mu^{-1}(u) - F_{\mu^N}^{-1}(u)|^p du + \sum_{m=1}^N \int_{F_\mu\left(\frac{m-1}{N}\right)}^{F_\mu\left(\frac{m}{N}\right)} |F_\mu^{-1}(u) - F_{\mu^N}^{-1}(u)|^p du. \end{aligned}$$

If $F_\mu\left(\frac{m}{N}\right) = F_\mu\left(\frac{m-1}{N}\right)$, we clearly have $\int_{F_\mu\left(\frac{m-1}{N}\right)}^{F_\mu\left(\frac{m}{N}\right)} |F_\mu^{-1}(u) - F_{\mu^N}^{-1}(u)|^p du = 0$. Otherwise, we have $F_{\mu^N}\left(\frac{m-1}{N}\right) = F_\mu\left(\frac{m-1}{N}\right) < F_\mu\left(\frac{m}{N}\right) = F_{\mu^N}\left(\frac{m}{N}\right)$, and therefore

$$\forall u \in \left(F_\mu\left(\frac{m-1}{N}\right), F_\mu\left(\frac{m}{N}\right)\right), \quad F_\mu^{-1}(u), F_{\mu^N}^{-1}(u) \in \left[\frac{m-1}{N}, \frac{m}{N}\right].$$

This gives $|F_\mu^{-1}(u) - F_{\mu^N}^{-1}(u)| \leq 1/N$. Since $F_\mu(0) = F_{\mu^N}(0)$, we get that $F_\mu^{-1}(u) = F_{\mu^N}^{-1}(u) = 0$ for $u \in (0, F_\mu(0))$. We finally get $W_p^p(\mu,\mu^N) \leq N^{-p}$.

Now, let $U \sim \mathcal{U}([0,1])$ be a uniform random variable on $[0,1]$. Still by Theorem 2.9 [291], we have $(F_\mu^{-1}(U), F_\nu^{-1}(U)) \sim \pi^*$ and $(F_{\mu^N}^{-1}(U), F_{\nu^N}^{-1}(U)) \sim \pi^N$. This gives a coupling between π^* and π^N, and thus

$$W_p^p(\pi^N,\pi^*) \leq \mathbb{E}[|F_{\mu^N}^{-1}(U) - F_\mu^{-1}(U)|^p] + \mathbb{E}[|F_{\nu^N}^{-1}(U) - F_\nu^{-1}(U)|^p] \leq \frac{2}{N^p}.$$

□

In order to prove Theorem 2.9, let us introduce the following auxiliary problem. For all $N \in \mathbb{N}^*$, let us define

$$\begin{aligned} \bar{\Pi}^N(\mu,\nu) := \Big\{ &(\bar\pi_{m,n})_{1\leq m,n\leq N} \,|\, \forall 1 \leq m,n \leq N,\ \bar\pi_{m,n} \geq 0, \\ &\forall m, \sum_{n=1}^N \bar\pi_{m,n} = \mu(T_m^N),\ \forall n, \sum_{m=1}^N \bar\pi_{m,n} = \nu(T_n^N) \Big\} \end{aligned}$$

and

$$J^N := \inf_{\bar\pi\in\bar\Pi^N(\mu,\nu)} \sum_{m,n=1}^N c\left(\frac{m-\frac12}{N}, \frac{n-\frac12}{N}\right)\bar\pi_{m,n}. \tag{2.48}$$

Let us introduce the following applications:

$$\begin{array}{rccc} D: & \Pi(\mu,\nu) & \to & \bar{\Pi}^N(\mu,\nu) \\ & \pi & \mapsto & (\pi(T_m^N \times T_n^N))_{1\le m,n\le N} \end{array} \tag{2.49}$$

and

$$\begin{array}{rccc} J: & \bar{\Pi}^N(\mu,\nu) & \to & \mathbb{R}^+ \\ & \bar{\pi} & \mapsto & \sum_{m,n=1}^N c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)\bar{\pi}_{m,n}. \end{array} \tag{2.50}$$

Lemma 2.11. *Let $N \in \mathbb{N}^*$. We have*

$$\forall \pi \in \mathcal{P}([0,1]),\ |I(\pi) - J(D(\pi))| \le \frac{K}{2N}. \tag{2.51}$$

Besides, we have

$$I^N \le J^N \le I^N + \frac{K}{2N}. \tag{2.52}$$

Proof of Lemma 2.11. Let $\pi \in \mathcal{P}([0,1]^2)$. Then, we write

$$\begin{aligned} I(\pi) &= \int_{[0,1]^2} c(x,y)\mathrm{d}\pi(x,y) = \sum_{m,n=1}^N \int_{T_m^N \times T_n^N} c(x,y)\mathrm{d}\pi(x,y) \\ &= \sum_{m,n=1}^N c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right) D_{mn}(\pi) \\ &\quad + \sum_{m,n=1}^N \int_{T_m^N \times T_n^N} \left(c(x,y) - c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)\right)\mathrm{d}\pi(x,y), \end{aligned}$$

and get $|I(\pi) - J(D(\pi))| \le \frac{K}{2N}$ since $|c(x,y) - c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)| \le \frac{K}{2N}$ for $(x,y) \in T_m^N \times T_n^N$.

Let $N \in \mathbb{N}^*$. For all $\bar{\pi} \in \bar{\Pi}(\mu,\nu)$, defining $\pi := \sum_{m,n=1}^N \bar{\pi}_{mn}\delta_{\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}}$, one obtains that $\pi \in \mathcal{P}([0,1]^2)$, $D(\pi) = \bar{\pi}$ and $I(\pi) = J(\bar{\pi})$; this implies that $I^N \le J^N$.

Conversely, if $\pi \in \Pi(\mu,\nu; (\phi_m^N)_{1\le m\le N}, (\phi_n^N)_{1\le n\le N})$ is chosen to satisfy $I(\pi) \le I^N + \epsilon$ for some $\epsilon > 0$, one gets $J^N \le J(D(\pi)) \le I(\pi) + \frac{K}{2N} = I^N + \frac{K}{2N} + \epsilon$. Letting $\epsilon \to 0$ provides the wanted result. □

We also need the following auxiliary lemma.

Lemma 2.12. *For all $\bar{\pi} \in \bar{\Pi}^N(\mu,\nu)$, there exists $\bar{\pi}^* \in \Pi(\mu,\nu)$ such that $\bar{\pi} = D(\bar{\pi}^*)$.*[i]

Proof of Lemma 2.12. Let $\bar{\pi} \in \bar{\Pi}(\mu,\nu)$. We define $\bar{\pi}^*$ by

$$\mathrm{d}\bar{\pi}^*(x,y) = \mathrm{d}\mu(x)\sum_{m=1}^N \mathbf{1}_{T_m^N}(x)\sum_{n=1}^N \frac{\bar{\pi}_{m,n}}{\sum_{n'=1}^N \bar{\pi}_{m,n'}}\frac{\mathbf{1}_{T_n^N}(y)\mathrm{d}\nu(y)}{\nu(T_n^N)}.$$

Since $\sum_{n'=1}^N \bar{\pi}_{m,n'} = \mu(T_m^N)$ and $\sum_{m=1}^N \bar{\pi}_{m,n} = \nu(T_n^N)$, we have

$$\begin{aligned} \int_{\mathcal{X}} \mathrm{d}\bar{\pi}^*(x,y) &= \sum_{m=1}^N \mu(T_m^N)\sum_{n=1}^N \frac{\bar{\pi}_{m,n}}{\mu(T_m^N)}\frac{\mathbf{1}_{T_n^N}(y)\mathrm{d}\nu(y)}{\nu(T_n^N)} \\ &= \sum_{n=1}^N \left(\sum_{m=1}^N \bar{\pi}_{m,n}\right)\frac{\mathbf{1}_{T_n^N}(y)\mathrm{d}\nu(y)}{\nu(T_n^N)} = \sum_{n=1}^N \mathbf{1}_{T_n^N}(y)\mathrm{d}\nu(y) = \mathrm{d}\nu(y). \end{aligned}$$

[i] In the literature, $\bar{\pi}^*$ is called the *block approximation* of $\bar{\pi}$ [83, Definition 2.9].

Also, we have $\int_{\mathcal{Y}} \mathrm{d}\bar{\pi}^*(x,y) = \mathrm{d}\mu(x) \sum_{m=1}^N \mathbf{1}_{T_m^N}(x) \sum_{n=1}^N \frac{\bar{\pi}_{m,n}}{\sum_{n'=1}^N \bar{\pi}_{m,n'}} = \mathrm{d}\mu(x)$, which gives $\bar{\pi}^* \in \Pi(\mu,\nu)$. Last, we have

$$\int_{T_m^N \times T_n^N} \mathrm{d}\bar{\pi}^*(x,y) = \mu(T_m^N) \frac{\bar{\pi}_{m,n}}{\sum_{n'=1}^N \bar{\pi}_{m,n'}} = \bar{\pi}_{m,n},$$

which precisely gives $\bar{\pi} = D(\bar{\pi}^*)$. □

We are now in position to give the proof of Theorem 2.9.

Proof of Theorem 2.9. The inclusion $\Pi(\mu,\nu) \subset \Pi(\mu,\nu;(\phi_m^N)_{1\le m\le N},(\phi_n^N)_{1\le n\le N})$ gives $I^N \le I$.

Lemma 2.12 implies that for all $\bar{\pi} \in \bar{\Pi}^N(\mu,\nu)$, there exists $\bar{\pi}^* \in \Pi(\mu,\nu)$ such that $D(\bar{\pi}^*) = \bar{\pi}$, and we get by Lemma 2.11 $|J(\bar{\pi}) - I(\bar{\pi}^*)| \le \frac{K}{2N}$. Let now $\bar{\pi} \in \bar{\Pi}^N(\mu,\nu)$ such that $J(\bar{\pi}) \le J^N + \epsilon$ for some $\epsilon > 0$. Then one gets that $J^N + \frac{K}{2N} + \epsilon \ge I(\bar{\pi}^*) \ge I$. Letting ϵ go to zero yields that

$$I \le J^N + \frac{K}{2N}. \tag{2.53}$$

Furthermore, Lemma 2.11 gives $J^N \le I^N + \frac{K}{N}$ and thus $I \le I^N + \frac{K}{N}$. □

Remark 2.6. *Theorem 2.9 can be easily extended to higher dimensions and to the multi-marginal case. Let us assume that $c : ([0,1]^d)^M \to \mathbb{R}_+$ is such that*

$$|c(x_1,\dots,x_M) - c(x_1',\dots,x_M')| \le K \max_{i\in\{1,\dots,M\}} \|x_i - x_i'\|_\infty.$$

For $N \in \mathbb{N}^*$ and $\mathbf{m} \in \{1,\dots,N\}^d =: \mathcal{E}_N$, we consider the test function $\phi_{\mathbf{m}}^N(x) = \prod_{i=1}^d \phi_{m_i}^N(x_i)$ for $x \in [0,1]^d$. Then, with

$$I = \inf_{\pi\in\Pi(\mu_1,\dots,\mu_M)} \int_{([0,1]^d)^M} c(x_1,\dots,x_M)\mathrm{d}\pi(x_1,\dots,x_M)$$

and

$$I^N = \inf_{\pi:\forall \mathbf{m},k,\int_{([0,1]^d)^M} \phi_{\mathbf{m}}^N(x_k)\mathrm{d}\pi(x_1,\dots,x_M)=\overline{\mu}_k^{\mathbf{m}}} \left\{ \int_{([0,1]^d)^M} c(x_1,\dots,x_M)\mathrm{d}\pi(x_1,\dots,x_M) \right\},$$

where $\forall \mathbf{m},k,\overline{\mu}_k^{\mathbf{m}} = \int_{[0,1]^d} \phi_{\mathbf{m}}^N(x)\mathrm{d}\mu_k(x)$, we get similarly (it is straightforward to generalize Proposition 2.10, and we can extend the result of Lemma 2.12 by induction on M)

$$I^N \le I^* \le I^N + \frac{K}{N}.$$

Since the number of moments (i.e. of test functions) involved in the computation of I^N is MN^d, we see that the storage complexity of getting an approximation of I^* with a given error is exponential in d but, in view of Remark 2.2 (resp. 2.3), only quadratically (resp. linearly) dependant on M.

2.5.2 Piecewise affine test functions in dimension 1 on a compact set

The test functions considered are discontinuous piecewise affine functions, identical on each space. For all $N \in \mathbb{N}^*$ and all $1 \le m \le N$, let us define the following

discontinuous piecewise affine functions

$$\phi_{m,1}^N(x) = \begin{cases} N\left(x - \frac{m-1}{N}\right) & \text{if} \quad x \in T_m^N, \\ 0 & \text{otherwise,} \end{cases}$$
$$\phi_{m,2}^N(x) = \begin{cases} 1 - N\left(x - \frac{m-1}{N}\right) & \text{if} \quad x \in T_m^N, \\ 0 & \text{otherwise,} \end{cases}$$

and for all $i = 1, 2$,

$$\bar{\mu}_{m,i}^N := \int_{\mathcal{X}} \phi_{m,i}^N \, d\mu \quad \text{and } \bar{\nu}_{m,i}^N := \int_{\mathcal{Y}} \phi_{m,i}^N \, d\nu.$$

Lemma 2.13. *Let $\mu_1, \mu_2 \in \mathcal{P}([0,1])$. Let $N \in \mathbb{N}^*$ and let us assume that for all $1 \leq m \leq N$ and $i = 1, 2$,*

$$\int_{[0,1]} \phi_{m,i}^N(x) \mathrm{d}\mu_1(x) = \int_{[0,1]} \phi_{m,i}^N(x) \mathrm{d}\mu_2(x).$$

Then, denoting by $F_1 : [0,1] \to [0,1]$ (resp. $F_2 : [0,1] \to [0,1]$) the cumulative distribution function of μ_1 (resp. μ_2), one gets that

$$\forall 1 \leq m \leq N, \quad \int_{T_m^N} F_1(x) \mathrm{d}x = \int_{T_m^N} F_2(x) \mathrm{d}x, \tag{2.54}$$

and

$$\forall 1 \leq m \leq N, \quad F_1\left(\frac{m}{N}\right) = F_2\left(\frac{m}{N}\right). \tag{2.55}$$

Proof. We have $\phi_{m,1} + \phi_{m,2} = \mathbf{1}_{T_m^N}$ and thus, for $2 \leq m \leq N$, $F_1\left(\frac{m}{N}\right) - F_1\left(\frac{m-1}{N}\right) = F_2\left(\frac{m}{N}\right) - F_2\left(\frac{m-1}{N}\right)$. Since $F_1(1) = F_2(1) = 1$, this gives (2.55). Now, let $l = 1, 2$. An integration by parts yields for $1 \leq m \leq N$

$$\begin{aligned} \int_{[0,1]} \phi_{m,1}^N(x) \mathrm{d}\mu_l(x) &= \int_{\frac{m-1}{N}}^{\frac{m}{N}} (x - \frac{m-1}{N}) \mathrm{d}\mu_l(x) \\ &= \frac{1}{N} F_l\left(\frac{m}{N}\right) - \int_{\frac{m-1}{N}}^{\frac{m}{N}} F_l(x) \mathrm{d}x \end{aligned}$$

Using (2.55), this gives (2.54). □

Let us remark that we may have $F_1(0) \neq F_2(0)$ under the assumptions of Lemma 2.13, since μ_1 and μ_2 may charge differently 0.

Let us now explain with a rough calculation why considering these test functions may lead to a convergence rate of $O(1/N^2)$ when c is C^1 with a Lipschitz gradient. Let

$$I^N = \inf_{\pi \in \Pi(\mu, \nu; (\phi_{m,i}^N), (\phi_{n,i}^N))} \left\{ \int_{\mathcal{X} \times \mathcal{Y}} c(x,y) \mathrm{d}\pi(x,y) \right\}. \tag{2.56}$$

We have $I^N \leq I$ and, for any $\pi \in \Pi(\mu, \nu; (\phi_{m,i}^N), (\phi_{n,i}^N))$,

$$\begin{aligned} I(\pi) = \sum_{m,n=1}^N \int_{T_m^N \times T_n^N} & c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right) + \partial_x c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)(x - \frac{m-\frac{1}{2}}{N}) \\ & + \partial_y c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)(y - \frac{m-\frac{1}{2}}{N}) \mathrm{d}x\mathrm{d}y + O(1/N^2) \end{aligned}$$

Thus, we have

$$I(\pi) = \sum_{m,n=1}^{N} \left(c - \frac{1}{2}\partial_x - \frac{1}{2}\partial_y\right)\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)\pi^1_{mn} \tag{2.57}$$
$$+ \partial_x c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)\pi^2_{mn} + \partial_y c\left(\frac{m-\frac{1}{2}}{N}, \frac{n-\frac{1}{2}}{N}\right)\pi^3_{mn} + O(1/N^2),$$

with the notations $\pi^1_{mn} = \pi(T^N_m \times T^N_n)$, $N\pi^2_{mn} = \int_{T^N_m \times T^N_n} \phi^N_{m,1}(x)\mathrm{d}\pi(x,y)$ and $N\pi^3_{mn} = \int_{T^N_m \times T^N_n} \phi^N_{m,1}(y)\mathrm{d}\pi(x,y)$. We can thus consider the linear programming problem of minimizing the right-hand-side of (2.57) under the constraints $\sum_n \pi^1_{mn} = \bar{\mu}^N_{m,1} + \bar{\mu}^N_{m,2}$, $\sum_m \pi^1_{mn} = \bar{\nu}^N_{m,1} + \bar{\nu}^N_{m,2}$, $\sum_n \pi^2_{mn} = \bar{\mu}^N_{m,1}/N$, $\sum_m \pi^3_{mn} = \bar{\nu}^N_{m,1}/N$ and $\pi^i_{mn} \geq 0$. Suppose for simplicity that we can find a minimum $(\pi^{*,i}_{mn})$ to this discrete problem. If we could find (similarly as Lemma 2.12) $\pi^* \in \Pi(\mu,\nu)$ such that $\pi^{*,1}_{mn} = \pi^*(T^N_m \times T^N_n)$, $N\pi^{*,2}_{mn} = \int_{T^N_m \times T^N_n} \phi^N_{m,1}(x)\mathrm{d}\pi^*(x,y)$ and $N\pi^{*,3}_{mn} = \int_{T^N_m \times T^N_n} \phi^N_{m,1}(y)\mathrm{d}\pi^*(x,y)$, we would get then

$$I \leq I^N + O(1/N^2).$$

Unfortunately, such a result is not obvious. Besides, we see from this derivation that the smoothness of the cost function plays an important role.

Let us recall that for $p \geq 1$, the W_p-Wasserstein distance at the power p, $W^p_p(\mu,\nu)$, corresponds to the cost function $c(x,y) = |x-y|^p$. In the following, we prove convergence results with rate $O(1/N^2)$ for $c(x,y) = |x-y|$ and $c(x,y) = |x-y|^2$. In the first case, the cost function is not smooth on the diagonal, and we need to impose an extra condition on μ and ν to get this rate. We first state a first result, which is already interesting, but will be not sufficient to prove the desired convergence. Its proof is postponed to Appendix A.1.

Proposition 2.14. *Let $\mu_1, \mu_2 \in \mathcal{P}([0,1])$ be two probability measures with cumulative distribution functions F_1 and F_2, respectively, such that $F_1, F_2 \in C^2([0,1])$. Let us assume that for all $1 \leq m \leq N$ and $i = 1, 2$,*

$$\int_{[0,1]} \phi^N_{m,i}(x)\mathrm{d}\mu_1(x) = \int_{[0,1]} \phi^N_{m,i}(x)\mathrm{d}\mu_2(x).$$

Then,

$$W_1(\mu_1,\mu_2) \leq \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{3N^2}. \tag{2.58}$$

In addition, let $m_1 := \min_{u\in[0,1]} F_1'(u)$ and $m_2 = \min_{u\in[0,1]} F_2'(u)$ and let us assume that $m_1 > 0$ and $m_2 > 0$. Then, for all $p > 1$, we have

$$W_p(\mu_1,\mu_2) \leq \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{3N^2}(p!)^{\frac{1}{p}}\left(\frac{5}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\right)^{\frac{p-1}{p}}. \tag{2.59}$$

Remark 2.7. *The result of Proposition 2.14 can be extended through the triangle inequality in order to treat regular measures with different piecewise affine moments. Indeed, for $p \geq 1$:*

$$W_p(\mu,\nu) \leq W_p(\mu,\tilde{\mu}) + W_p(\tilde{\mu},\tilde{\nu}) + W_p(\tilde{\nu},\nu),$$

thus

$$|W_p(\mu,\nu) - W_p(\tilde{\mu},\tilde{\nu})| \leq W_p(\mu,\tilde{\mu}) + W_p(\tilde{\nu},\nu). \tag{2.60}$$

Thus, using Proposition 2.14, one gets that for μ, ν two measures with cumulative distribution functions F and G, respectively, such that $F, G \in C^2([0,1])$ and $\tilde{\mu}$, $\tilde{\nu}$

two measures with cumulative distribution functions $\tilde{F}$ and $\tilde{G}$, respectively, such that $\tilde{F}, \tilde{G} \in C^2([0,1])$; If μ and $\tilde{\mu}$ (respectively ν and $\tilde{\nu}$) have the same $2N$ piecewise affine moments of step $1/N$, then

$$|W_1(\mu,\nu) - W_1(\tilde{\mu},\tilde{\nu})| \leq \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty + \|G''\|_\infty + \|\tilde{G}''\|_\infty}{3N^2}. \tag{2.61}$$

Besides, if $m_\mu = \min_{u\in[0,1]} F'(u)$, $m_{\tilde{\mu}} = \min_{u\in[0,1]} \tilde{F}'(u)$, $m_\nu = \min_{u\in[0,1]} G'(u)$ and $m_{\tilde{\nu}} = \min_{u\in[0,1]} \tilde{G}'(u)$, are positive, one has for all $p \in \mathbb{N}^*$,

$$\begin{aligned}|W_p(\mu,\nu) - W_p(\tilde{\mu},\tilde{\nu})| \\ &\leq \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{3N^2} \left(\frac{5}{2}\left(\frac{1}{m_\mu} + \frac{1}{m_{\tilde{\mu}}}\right)\right)^{\frac{p-1}{p}} (p!)^{\frac{1}{p}} \\ &\quad + \frac{\|G''\|_\infty + \|\tilde{G}''\|_\infty}{3N^2} \left(\frac{5}{2}\left(\frac{1}{m_\nu} + \frac{1}{m_{\tilde{\nu}}}\right)\right)^{\frac{p-1}{p}} (p!)^{\frac{1}{p}}. \end{aligned} \tag{2.62}$$

Unfortunately, Proposition 2.14 cannot be extended to non-smooth measures, as Example 2.3 below shows. However, the $O(1/N^2)$ convergence obtained in Remark 2.7 may stay true even for non-smooth measures $\tilde{\mu}$ and $\tilde{\nu}$. This is important in our context to treat the case where $\tilde{\mu}$ and $\tilde{\nu}$ are not smooth since the solution of the MCOT problem may typically be a discrete measure that match respectively the moments of μ and ν. We tackle this issue for W_1 and W_2 in the two following paragraphs.

Example 2.3. *In Proposition 2.14, if one of the measures (let us say $\tilde{\mu}$) is not regular enough, then the convergence in $O(1/N^2)$ may not be true, as shown thereafter.*

We consider $\mu \sim \mathcal{U}([0,1])$ and

$$\tilde{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\frac{1}{2N}+\frac{i-1}{N}}.$$

Then, for all $1 \leq m \leq N$, we have

$$\tilde{F}\left(\frac{m}{N}\right) = \frac{m}{N} = F\left(\frac{m}{N}\right),$$

and

$$\int_{T_m^N} \tilde{F} = \frac{m-1}{N^2} + \frac{1}{2N}\frac{1}{N} = \int_{T_m^N} F.$$

However, we have

$$W_1(\mu,\tilde{\mu}_N) = N\int_0^{1/N} \left|u - \frac{1}{2N}\right| \mathrm{d}u = 2N\left(\frac{1}{2N}\right)^2 \frac{1}{2} = \frac{1}{4N}.$$

2.5.2.1 Convergence speed for W_1

Proposition 2.15. *Let $\mu, \nu, \tilde{\mu}, \tilde{\nu} \in \mathcal{P}([0,1])$. Let us assume that μ and ν are absolutely continuous with respect to the Lebesgue measure and let us denote by ρ_μ and ρ_ν their density probability functions. Let us denote by F_μ, F_ν, $F_{\tilde{\mu}}$ and $F_{\tilde{\nu}}$ the cumulative distribution functions of μ, ν, $\tilde{\mu}$ and $\tilde{\nu}$ respectively. Let $N \in \mathbb{N}^*$. Let us assume that*

$$\forall 1 \leq m \leq N, \quad \int_{T_m^N} F_\mu = \int_{T_m^N} F_{\tilde{\mu}} \quad \text{and} \quad \int_{T_m^N} F_\nu = \int_{T_m^N} F_{\tilde{\nu}}. \tag{2.63}$$

Let us assume in addition that the function $F_\mu - F_\nu$ changes sign at most Q times for some $Q \in \mathbb{N}$. More precisely, denoting by $G := F_\mu - F_\nu$, we assume that there exist $x_0 = 0 < x_1 < x_2 < \cdots < x_Q < x_{Q+1} = 1 \in [0,1]$ such that for all $1 \leq q \leq Q+1$,

$$\forall x, y \in [x_{q-1}, x_q],\ G(x)G(y) \geq 0, \tag{2.64}$$

and for all $1 \leq q \leq Q$,

$$\forall x \in [x_{q-1}, x_q],\ \forall z \in [x_q, x_{q+1}],\ G(x)G(z) \leq 0. \tag{2.65}$$

Let us also assume that $\rho_\mu - \rho_\nu \in L^\infty([0,1], \mathrm{d}x; \mathbb{R})$. Then,

$$W_1(\mu, \nu) \leq W_1(\tilde{\mu}, \tilde{\nu}) + 2\|\rho_\mu - \rho_\nu\|_\infty \frac{Q}{N^2}. \tag{2.66}$$

Note that we only assume regularity of the measures μ, ν, not of $\tilde{\mu}, \tilde{\nu}$. The assumption that $F_\mu - F_\nu$ changes sign at most Q times is related to the fact that $c(x,y) = |x-y|$ is not smooth on the diagonal: an optimal coupling is given by the inverse transform coupling, and $F_\mu^{-1} - F_\nu^{-1}$ changes sign at most Q times as well. Last, remarkably, we do not need for this result to assume $F_\mu(m/N) = F_{\tilde{\mu}}(m/N)$ and $F_\nu(m/N) = F_{\tilde{\nu}}(m/N)$. Thus, it is sufficient to work with continuous piecewise affine test functions.

More precisely, for all $N \in \mathbb{N}^*$, let us define

$$\forall x \in [0,1], \quad \psi_1^N(x) = \begin{cases} 1 - Nx & \text{if} \quad x \in T_1^N \\ 0 & \text{elsewhere,} \end{cases}$$

and for all $2 \leq m \leq N$,

$$\psi_m^N(x) = \begin{cases} N\left(x - \frac{m-2}{N}\right) & \text{if} \quad x \in T_{m-1}^N \\ 1 - N\left(x - \frac{m-1}{N}\right) & \text{if} \quad x \in T_m^N \\ 0 & \text{elsewhere.} \end{cases}$$

We can check by integration by parts that $\int_{[0,1]} \psi_1^N(x)\mathrm{d}\mu(x) = N\int_{T_1^N} F_\mu(x)\mathrm{d}x$ and $\int_{[0,1]} \psi_m^N(x)\mathrm{d}\mu(x) = N\int_{T_m^N} F_\mu(x)\mathrm{d}x - N\int_{T_{m-1}^N} F_\mu(x)\mathrm{d}x$ for $2 \leq m \leq N$. Therefore, we get

$$\begin{aligned} &\forall m \in \{1, \ldots, N\}, \int_{[0,1]} \psi_m^N(x)\mathrm{d}\mu(x) = \int_{[0,1]} \psi_m^N(x)\mathrm{d}\tilde{\mu}(x) \\ \iff &\forall m \in \{1, \ldots, N\}, \int_{T_m^N} F_\mu(x)\mathrm{d}x = \int_{T_m^N} F_{\tilde{\mu}(x)}\mathrm{d}x. \end{aligned} \tag{2.67}$$

Last, let us remark that $\psi_1^N = \phi_{1,2}^N$ and $\psi_m^N = \phi_{m-1,1}^N - \phi_{m,2}^N$ for $2 \leq m \leq N$ so that $\mathrm{Span}\left\{\psi_n^N,\ 1 \leq n \leq N\right\} \subset \mathrm{Span}\left\{\phi_{n,1}^N, \phi_{n,2}^N,\ 1 \leq n \leq N\right\}$ and

$$\Pi(\mu, \nu; (\phi_{n,l}^N), (\phi_{n,l}^N)) \subset \Pi(\mu, \nu; (\psi_n^N), (\psi_n^N)).$$

Corollary 2.16. *Let $\mu, \nu \in \mathcal{P}([0,1])$. Let us assume that μ and ν are absolutely continuous with respect to the Lebesgue measure and let us denote by ρ_μ and ρ_ν their density probability functions. Let F_μ and F_ν be their cumulative distribution functions. For all $N \in \mathbb{N}^*$, let us define*

$$I^N = \inf_{\pi \in \Pi(\mu, \nu; (\psi_m^N)_{1 \leq m \leq N}, (\psi_n^N)_{1 \leq n \leq N})} \left\{ \int_{[0,1]^2} |x-y| \mathrm{d}\pi(x,y) \right\}. \tag{2.68}$$

There exists a minimizer for (2.68). Let us assume in addition that the function $F_\mu - F_\nu$ changes sign at most Q times for some $Q \in \mathbb{N}$ (in the sense of Proposition 2.15) and that $\rho_\mu - \rho_\nu \in L^\infty([0,1], \mathrm{d}x; \mathbb{R})$. Then,

$$I^N \leq W_1(\mu, \nu) \leq I^N + 2\|\rho_\mu - \rho_\nu\|_\infty \frac{Q}{N^2} \tag{2.69}$$

In fact, looking at the proof of Proposition 2.15, it even is sufficient to assume that $\rho_\mu - \rho_\nu$ is bounded on a neighborhood of the points at which $F_\mu - F_\nu$ changes sign. For simplicity of statements, we have assumed in Proposition 2.15 and Corollary 2.16 that $\rho_\mu - \rho_\nu$ is bounded on $[0,1]$.

Proof of Corollary 2.16. From the inclusion $\Pi(\mu,\nu) \subset \Pi(\mu,\nu;(\psi_m^N)_{1\le m\le N},(\psi_n^N)_{1\le n\le N})$, we clearly have $I^N \le W_1(\mu,\nu)$. Using Theorem 2.3 together with Remark 2.1 (ii)-(iii), since the functions ψ_m^N are continuous on $[0,1]$ for all $1 \le m \le N$, there exists $\pi^N \in \Pi(\mu,\nu;(\psi_m^N)_{1\le m\le N},(\psi_n^N)_{1\le n\le N})$ which is a minimizer to Problem (2.68). Let us denote by $\tilde\mu$ and $\tilde\nu$ the marginal laws of π^N. First, we remark that

$$I^N = \int_0^1 |x-y| \mathrm{d}\pi^N(x,y) \ge \min_{\pi\in\Pi(\tilde\mu,\tilde\nu)} \left\{ \int_0^1 |x-y| \mathrm{d}\pi(x,y) \right\} = W_1(\tilde\mu,\tilde\nu).$$

Second, using the fact that

$$\Pi(\tilde\mu,\tilde\nu) \subset \Pi(\tilde\mu,\tilde\nu;(\psi_m^N)_{1\le m\le N},(\psi_n^N)_{1\le n\le N}) = \Pi(\mu,\nu;(\psi_m^N)_{1\le m\le N},(\psi_n^N)_{1\le n\le N}),$$

we obtain

$$\begin{aligned} I^N = \int_0^1 |x-y| \mathrm{d}\pi^N(x,y) &= \min_{\pi\in\Pi(\tilde\mu,\tilde\nu;(\psi_m^N)_{1\le m\le N},(\psi_n^N)_{1\le n\le N})} \left\{ \int_0^1 |x-y| \mathrm{d}\pi(x,y) \right\} \\ &\le \min_{\pi\in\Pi(\tilde\mu,\tilde\nu)} \left\{ \int_0^1 |x-y| \mathrm{d}\pi(x,y) \right\} = W_1(\tilde\mu,\tilde\nu). \end{aligned}$$

Thus, $I^N = W_1(\tilde\mu,\tilde\nu)$. Besides, we have for all $1 \le m \le N$,

$$\int_{[0,1]} \psi_m^N(x) \mathrm{d}\tilde\mu(x) = \int_{[0,1]} \psi_m^N(x) \mathrm{d}\mu(x), \quad \int_{[0,1]} \psi_m^N(y) \mathrm{d}\tilde\nu(y) = \int_{[0,1]} \psi_m^N(y) \mathrm{d}\nu(y),$$

and we therefore get (2.63) from (2.67). We can thus apply Proposition 2.15 and get the desired result. □

Proof of Proposition 2.15. Let $1 \le m \le N$. If for all $1 \le q \le Q$, $x_q \notin T_m^N$, then $F_\mu - F_\nu$ remains non-negative or non-positive on T_m^N. Thus, using (2.63), we deduce that

$$\begin{aligned} \int_{T_m^N} |F_\mu - F_\nu| &= \epsilon \int_{T_m^N} (F_\mu - F_\nu) \\ &= \epsilon \int_{T_m^N} (F_{\tilde\mu} - F_{\tilde\nu}) = \left| \int_{T_m^N} (F_{\tilde\mu} - F_{\tilde\nu}) \right| \le \int_{T_m^N} |F_{\tilde\mu} - F_{\tilde\nu}|, \end{aligned}$$

where $\epsilon = 1$ if $F_\mu - F_\nu \ge 0$ on T_m^N and $\epsilon = -1$ if $F_\mu - F_\nu \le 0$ on T_m^N. On the other hand, if there exists $1 \le q \le Q$, such that $x_q \in T_m^N$, one gets

$$\begin{aligned} \int_{T_m^N} |F_\mu - F_\nu| &= \int_{T_m^N} (F_\mu - F_\nu) + 2 \int_{T_m^N} (F_\mu - F_\nu)^- \\ &= \int_{T_m^N} (F_{\tilde\mu} - F_{\tilde\nu}) + 2 \int_{T_m^N} (F_\mu - F_\nu)^- \\ &\le \int_{T_m^N} |F_{\tilde\mu} - F_{\tilde\nu}| + 2 \int_{T_m^N} (F_\mu - F_\nu)^- \\ &\le \int_{T_m^N} |F_{\tilde\mu} - F_{\tilde\nu}| + 2\|\rho_\mu - \rho_\nu\|_\infty \frac{1}{N^2}, \end{aligned}$$

since for $x \in T_m^N$, $F_\mu(x) - F_\nu(x) = \int_{x_q}^x \rho_\mu - \rho_\nu$ and $|x - x_q| \leq 1/N$.

Thus, as there are at most Q intervals of that last type, we get

$$\int_0^1 |F_\mu - F_\nu| \leq \int_0^1 |F_{\tilde{\mu}} - F_{\tilde{\nu}}| + 2\|\rho_\mu - \rho_\nu\|_\infty \frac{Q}{N^2},$$

i.e. $W_1(\mu,\nu) \leq W_1(\tilde{\mu},\tilde{\nu}) + 2\|\rho_\mu - \rho_\nu\|_\infty \frac{Q}{N^2}$. □

2.5.2.2 Convergence speed for W_2

Proposition 2.17. *Let $\mu, \nu, \tilde{\mu}, \tilde{\nu} \in \mathcal{P}([0,1])$. Let us assume that $\mu(\mathrm{d}x) = \rho_\mu(x)\mathrm{d}x$ and $\nu(\mathrm{d}x) = \rho_\nu(x)\mathrm{d}x$ with $\rho_\mu, \rho_\nu \in L^\infty([0,1], \mathrm{d}x; \mathbb{R}_+)$. Let us denote by F_μ, F_ν, $F_{\tilde{\mu}}$ and $F_{\tilde{\nu}}$ the cumulative distribution functions of μ, ν, $\tilde{\mu}$ and $\tilde{\nu}$ respectively. Let $N \in \mathbb{N}^*$. Let us assume that*

$$\forall 1 \leq m \leq N, F_\mu\left(\frac{m}{N}\right) = F_{\tilde{\mu}}\left(\frac{m}{N}\right) \quad \text{and} \quad F_\nu\left(\frac{m}{N}\right) = F_{\tilde{\nu}}\left(\frac{m}{N}\right), \tag{2.70}$$

$$\forall 1 \leq m \leq N, \quad \int_{T_m^N} F_\mu = \int_{T_m^N} F_{\tilde{\mu}} \quad \text{and} \quad \int_{T_m^N} F_\nu = \int_{T_m^N} F_{\tilde{\nu}}. \tag{2.71}$$

Then,

$$W_2^2(\mu,\nu) \leq W_2^2(\tilde{\mu},\tilde{\nu}) + \frac{7}{3}\frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{N^2}. \tag{2.72}$$

This proposition plays the same role as Proposition 2.15 for W_1. Again, the important point is that no regularity assumption is made on $\tilde{\mu}$ and $\tilde{\nu}$. We note that we no longer have restriction on the number of points where $F_\mu - F_\nu$ changes sign, which is related to the fact that $c(x,y) = (x-y)^2$ is smooth. Contrary to Proposition 2.15, we need here the condition (2.70).

Corollary 2.18. *Let $\mu, \nu \in \mathcal{P}([0,1])$. Let us assume that $\mu(\mathrm{d}x) = \rho_\mu(x)\mathrm{d}x$ and $\nu(\mathrm{d}x) = \rho_\nu(x)\mathrm{d}x$ with $\rho_\mu, \rho_\nu \in L^\infty([0,1], \mathrm{d}x; \mathbb{R}_+)$. Let F_μ and F_ν be their cumulative distribution functions. For all $N \in \mathbb{N}^*$, let us define*

$$I^N = \inf_{\pi \in \Pi(\mu,\nu;(\phi^N_{m,l})_{\substack{1\leq m\leq N\\1\leq l\leq 2}},(\phi^N_{m,l})_{\substack{1\leq m\leq N\\1\leq l\leq 2}})} \left\{ \int_{[0,1]^2} (x-y)^2 \mathrm{d}\pi(x,y) \right\}. \tag{2.73}$$

Then,

$$I^N \leq W_2^2(\mu,\nu) \leq I^N + \frac{7}{3}\frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{N^2}. \tag{2.74}$$

We omit the proof of Corollary 2.18 since it follows the same line as the one of Corollary 2.16. The only difference is that we do not know here if the infimum is a minimum and have to work for an arbitrary $\epsilon > 0$ with the probability measure $\pi \in \Pi(\mu,\nu;(\phi^N_{m,l})_{\substack{1\leq m\leq N\\1\leq l\leq 2}},(\phi^N_{m,l})_{\substack{1\leq m\leq N\\1\leq l\leq 2}})$ such that $\int_{[0,1]^2}(x-y)^2\mathrm{d}\pi(x,y) \leq I^N + \epsilon$. Let us also mention here that we can use Proposition 2.10 to bound the distance between an MCOT minimizer and an OT minimizer since $\phi_{m,1} + \phi_{m,2} = \mathbf{1}_{T_m^N}$.

Proof of Proposition 2.17. From Lemma B.3 [188], we have

$$\begin{aligned}
W_2^2(\mu,\nu) &= \int_0^1 \int_0^1 \mathbf{1}_{x<y}([F_\mu(x) - F_\nu(y)]^+ + [F_\nu(x) - F_\mu(y)]^+)\mathrm{d}x\mathrm{d}y \\
&= \sum_{k=1}^N \sum_{l=k+1}^N \int_{T_k^N}\int_{T_l^N} ([F_\mu(x) - F_\nu(y)]^+ + [F_\nu(x) - F_\mu(y)]^+)\mathrm{d}x\mathrm{d}y \\
&\quad + \sum_{k=1}^N \int_{T_k^N}\int_{T_k^N} \mathbf{1}_{x<y}([F_\mu(x) - F_\nu(y)]^+ + [F_\nu(x) - F_\mu(y)]^+)\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

The two terms $[F_\mu(x) - F_\nu(y)]^+$ and $[F_\nu(x) - F_\mu(y)]^+$ can be analyzed in the same way by exchanging μ and ν, and we focus on the first one. Thus, we consider for $k \leq l$ the term $\alpha_{kl} := \int_{T_k^N}\int_{T_l^N} \mathbf{1}_{x<y}[F_\mu(x) - F_\nu(y)]^+ \mathrm{d}x\mathrm{d}y$ and denote $\tilde{\alpha}_{kl} = \int_{T_k^N}\int_{T_l^N} \mathbf{1}_{x<y}[F_{\tilde{\mu}}(x) - F_{\tilde{\nu}}(y)]^+ \mathrm{d}x\mathrm{d}y$.

• If $F_\mu(k/N) \leq F_\nu((l-1)/N)$, then from (2.70), we have also $F_{\tilde{\mu}}(k/N) \leq F_{\tilde{\nu}}((l-1)/N)$ (note that if $l = 1$, $F_{\tilde{\nu}}(0) \geq 0 = F_\nu(0)$). Thus, $\alpha_{kl} = \tilde{\alpha}_{kl} = 0$.

• If $F_\nu(l/N) \leq F_\mu((k-1)/N)$, then from (2.70), we have also $F_{\tilde{\nu}}(l/N) \leq F_{\tilde{\mu}}((k-1)/N)$, and using (2.71) we get for $k < l$

$$\alpha_{kl} = \int_{T_k^N}\int_{T_l^N} F_\mu(x) - F_\nu(y)\mathrm{d}x\mathrm{d}y = \int_{T_k^N}\int_{T_l^N} F_{\tilde{\mu}}(x) - F_{\tilde{\nu}}(y)\mathrm{d}x\mathrm{d}y = \tilde{\alpha}_{kl}.$$

For $k = l$, we have by using (2.71) and Lemma A.1 for the inequality

$$\begin{aligned}
\alpha_{kk} &= \int_{T_k^N} \left(\int_{\frac{k-1}{N}}^{x} F_\mu - \int_{x}^{\frac{k}{N}} F_\nu \right) \mathrm{d}x \\
&= \int_{T_k^N} \left(\int_{\frac{k-1}{N}}^{x} F_\mu + \int_{\frac{k-1}{N}}^{x} F_\nu \right) \mathrm{d}x - \frac{1}{N}\int_{T_k^N} F_\nu \\
&\leq \int_{T_k^N} \left(\int_{\frac{k-1}{N}}^{x} F_{\tilde{\mu}} + \int_{\frac{k-1}{N}}^{x} F_{\tilde{\nu}} \right) \mathrm{d}x - \frac{1}{N}\int_{T_k^N} F_{\tilde{\nu}} + \frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{6N^3} \\
&= \tilde{\alpha}_{kk} + \frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{6N^3}.
\end{aligned}$$

• We now consider the case where $F_\mu(k/N) > F_\nu((l-1)/N)$ and $F_\nu(l/N) > F_\mu((k-1)/N)$. We can thus find $x_0 \in T_k^N$ and $y_0 \in T_l^N$ such that $F_\mu(x_0) = F_\nu(y_0)$. We then have $\forall x \in T_k^N, y \in T_l^N, |F_\mu(x) - F_\nu(y)| \leq |F_\mu(x) - F_\mu(x_0)| + |F_\nu(y_0) - F_\nu(y)| \leq \|\rho_\mu\|_\infty |x - x_0| + \|\rho_\nu\|_\infty |y - y_0|$, and thus using that $\int_{T_k^N} |x - x_0|\mathrm{d}x \leq \frac{1}{2N^2}$,

$$\alpha_{kl} \leq \frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{2N^3} \leq \tilde{\alpha}_{kl} + \frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{2N^3}.$$

For $1 \leq k \leq N$, we note $\{l_k, l_k + 1, , \ldots, l_k + n_k - 1\} \subset \{1, \ldots, N\}$ the set of l such that $F_\mu((k-1)/N) < F_\nu(l/N)$ and $F_\mu(k/N) > F_\nu((l-1)/N)$. We necessarily have $l_{k+1} \geq l_k + n_k - 1$ since $F_\nu((l_k + n_k - 2)/N) < F_\mu(k/N) < F_\nu(l_{k+1}/N)$. Therefore, there is at most one element overlap between two consecutive sets, and thus $\sum_{k=1}^N n_k \leq 2N$.

Combining all cases, and taking into account the contribution of the symmetric term $[F_\nu(x) - F_\mu(y)]^+$ in the integral, we finally get

$$W_2^2(\mu, \nu) \leq W_2^2(\tilde{\mu}, \tilde{\nu}) + 2\left(N\frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{6N^3} + 2N\frac{\|\rho_\mu\|_\infty + \|\rho_\nu\|_\infty}{2N^3} \right),$$

which gives (2.72) □

2.6 Numerical algorithms to approximate optimal transport problems

This section presents the implementation of two algorithms for the approximation of the Optimal Transport cost. Both algorithms rely on Theorem 2.3, i.e. that the optimum of the MCOT problem is attained by a finite discrete measure $\sum_{k=1}^{2N+2} p_k \delta_{(x_k, y_k)}$. The two algorithms corresponds to the following choices:

1. piecewise constant test functions,

2. (regularized) piecewise linear test functions.

In the first case, the precise positions (x_k, y_k) are useless to satisfy the moment constraints: only matters in which cell (x_k, y_k) belongs. Thus, the optimization problem is essentially discrete on a (large) finite space, for which Metropolis-Hastings algorithms are relevant. In the second case, we implement a penalized gradient algorithm to optimize the positions (x_k, y_k) and the weights p_k.

The goal of these numerical tests is only illustrative to see the potential relevance of this approach. We do not claim that these algorithms are more efficient than other existing methods in the literature, and the improvement of our algorithms is left for future research.

2.6.1 Metropolis-Hastings algorithm on a finite state space

We expose in the following the principles of the Metropolis-Hastings algorithm used to compute an approximation of the OT cost. For simplicity, we do so in the case of two one-dimensional marginal laws. However, the algorithm principles can be adapted to solve a Multimarginal MCOT problem with marginal laws defined on spaces of any finite dimension.

2.6.1.1 Description of the algorithm

For this algorithm, we consider the framework of Subsection 2.5.1, i.e. N piecewise constant functions $\phi_m^N = \mathbf{1}_{T_m^N}$, $1 \leq m \leq N$, and the MCOT problem (2.46). As mentioned above, if (x_k, y_k) belongs to the cell $T_i^N \times T_j^N$, its position in this cell does not matter for the moment constraints. We can therefore assume that the position minimizes the cost in this cell. For $c(x, y) = |y - x|^2$, this amounts to take

$$c\left(x_k, y_k\right) = \tilde{c}(i, j) \text{ with } \tilde{c}(i, j) = \begin{cases} c\left(\frac{i}{N}, \frac{j+1}{N}\right) & \text{if} \quad i > j \\ c\left(\frac{i}{N}, \frac{j}{N}\right) & \text{if} \quad i = j \\ c\left(\frac{i+1}{N}, \frac{j}{N}\right) & \text{if} \quad i < j. \end{cases}$$

We consider then $2N + 2$ distinct cells $T_{i_k}^N \times T_{j_k}^N$, $k \in \{1, \ldots, 2N + 2\}$. The weights associated to each cell is determined as the solution of the linear optimization of the cost associated under the constraint that the weights satisfy the moments constraints:

$$\left(p_1, ..., p_{2N+2}\right) = \underset{\substack{p_k \geq 0,\ \sum_{k=1}^{2N+2} p_k = 1 \\ \forall 1 \leq m \leq N, \sum_{k=1}^{2N+2} p_k \mathbf{1}_{i_k = m} = \bar{\mu}_m \\ \forall 1 \leq n \leq N, \sum_{k=1}^{2N+2} p_k \mathbf{1}_{j_k = n} = \bar{\nu}_n}}{\arg\min} \sum_{k=1}^{2N+2} p_k \tilde{c}\left(i_k, j_k\right). \tag{2.75}$$

Note that this set of constraints may be void. To start with an initial configuration $(i_k, j_k)_{1 \leq k \leq 2N+2}$ that allows the existence of weights which satisfy the constraints, we use the inverse transform sampling between the distributions given by $(\bar{\mu}_k)_{1 \leq k \leq N}$ and $(\bar{\nu}_k)_{1 \leq k \leq N}$ on $\{1, \ldots, N\}$. This gives in fact the optimal solution $(p_k, (i_k, j_k))_{1 \leq k \leq 2N+2}$ for (2.46) satisfying in particular the constraints. Since we want here to test the relevance of the Metropolis-Hastings algorithm in this framework, we do not want to start from the optimal solution: thus, we consider a random permutation σ on $\{1, \ldots, N\}$ and then the inverse transform sampling between the distributions given by $(\bar{\mu}_k)_{1 \leq k \leq N}$ and $(\bar{\nu}_{\sigma(k)})_{1 \leq k \leq N}$ on $\{1, \ldots, N\}$. This gives a configuration that satisfy the constraints and is not a priori optimal.

We now have to specify how the Markov chain defining the algorithm moves from one state $(i_k, j_k)_{1\leq k\leq 2N+2}$ to another. Let us denote by $N(i_k, j_k) = \{(i_k+1, j_k), (i_k - 1, j_k), (i_k, j_k + 1), (i_k, j_k - 1)\}$ the neighboring cells of (i_k, j_k) and

$$FN(i_k, j_k) = N(i_k, j_k) \cap (\{1, \dots, N\}^2 \setminus (\cup_{k'\neq k}\{(i_{k'}, j_{k'})\}),$$

the neighboring cells that are free, i.e. that are not in the current configuration. We choose randomly and uniformly a cell $l \in \{1, \dots, 2N + 2\}$. If $FN(i_l, j_l) = \emptyset$, we pick randomly another one. This rejection method amounts to choose randomly and uniformly a cell l among those such that $FN(i_l, j_l) \neq \emptyset$. Then, we select (i'_l, j'_l) uniformly on $FN(i_l, j_l)$ and set $(i'_k, j'_k) = (i_k, j_k)$ for $k \neq l$, and we accept the new configuration $(i'_k, j'_k)_{1\leq k\leq 2N+2}$ only if it allows to satisfy the constraints and with an acceptance ratio described in Algorithm 1. In practice, we run this algorithm with $K \geq 2N + 2$ cells, in order to increase the chance that the new configuration is compatible with the constraints.

Algorithm 1 Metropolis-Hastings algorithm

Fix a temperature $\beta \in \mathbb{R}^+$ and take $2N + 2 \leq K \leq N^2$.
Initialize cells $(i_k, j_k)_{1\leq k\leq K}$ and compute the actual optimal cost $c_{\text{actual}} = \sum_{k=1}^{K} p_k \tilde{c}\,(i_k, j_k)$.
for a given number of steps **do**
  Choose randomly a particle $1 \leq l \leq K$ such that $FN(i_l, j_l) \neq \emptyset$.
  Compute $n_{\text{actual}} = \text{Card}(FN(i_l, j_l))$ the number of free cells near (i_l, j_l).
  Choose randomly a new cell (i'_l, j'_l) in $FN(i_l, j_l)$.
  if the configuration $(i'_k, j'_k)_{1\leq k\leq 2N+2}$ allows to satisfy the constraints **then**
    Compute c_{newpos} the optimal cost associated to the configuration $(i'_k, j'_k)_{1\leq k\leq 2N+2}$.
    Compute n_{newpos}, the number of free cells near (i'_l, j'_l) in the new configuration.
    Move the particle l with probability $\min\left(1, \dfrac{e^{-c_{\text{newpos}}/\beta}}{e^{-c_{\text{actual}}/\beta}} \dfrac{n_{\text{actual}}}{n_{\text{newpos}}}\right)$. This probability is the acceptance ratio of the Metropolis-Hastings algorithm, as explained in Section 2.2 of [122].
    Update the value of c_{actual} to c_{newpos} if the move is accepted.
  end if
end for
return the lowest cost encountered throughout the loop.

The state space of the Markov Chain describing Algorithm 1 is the set of K distinct elements of $\{1, \dots, N\}^2$. Note that we can go from any points (i, j) to (i', j') with at most $2N - 2$ moves (a move consists in adding or removing one to one of the coordinate). If we ignore the problem of satisfying the constraints, we can therefore go from a configuration $(i_k, j_k)_{1\leq k\leq 2N+2}$ to another one $(i'_k, j'_k)_{1\leq k\leq 2N+2}$ with at most $K(2N - 2)$ moves, which let think that the Doeblin condition may be satisfied. This would ensure theoretically the convergence of the algorithm converges towards the infimum

$$\inf_{\pi\in\Pi(\mu,\nu;(\phi_m^N)_{1\leq m\leq N},(\phi_n^N)_{1\leq n\leq N})} \int_0^1 \int_0^1 c(x, y)\mathrm{d}\pi(x, y), \tag{2.76}$$

and that the convergence is exponentially fast (see e.g. Section 2 of [122]).

2.6.1.2 Numerical examples

We tested the algorithm for the marginal laws with probability density functions

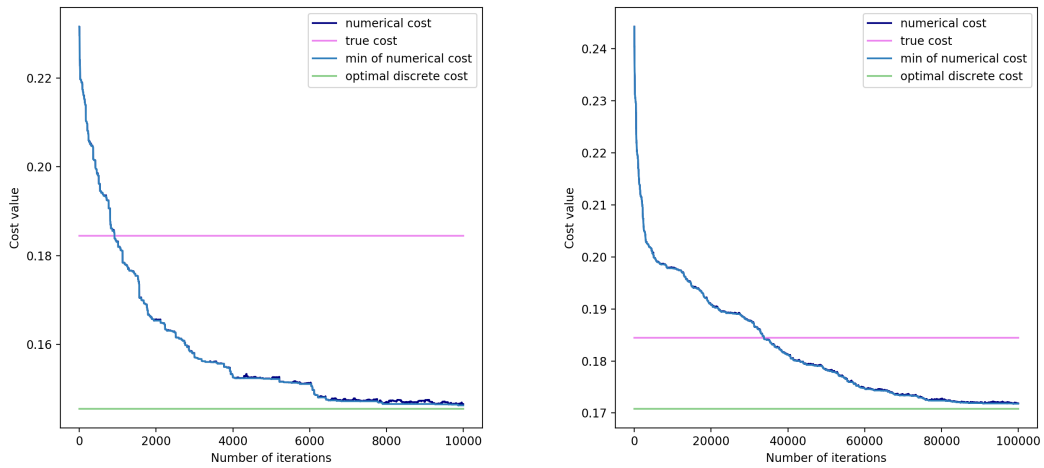
$$\rho_\mu(x) = 3x^2 \mathbf{1}_{[0,1]}(x), \quad \rho_\nu(y) = (2-2y)\mathbf{1}_{[0,1]}(y), \tag{2.77}$$

and the quadratic cost $c(x,y) = |x-y|^2$.

We consider a number of particles $K = 3N+2$ in order at each step to have more freedom among the configurations which satisfy the constraints. We present two minimizations:

- $N = 20$ and $\beta = 0.000075$
- $N = 60$ and $\beta = 0.00002$.

The evolution of the configurations through the iterations are represented for $N = 20$ and $N = 60$ in Figure 2.2. The darker the cell, the more weight it has. In green (Figures 2.2.6 and 2.2.12) are represented the optimal configuration for the given number of moment constraints. The convergence of the numerical cost for each minimization is represented in Figure 2.1. The pink line represents the cost of the Optimal Transport problem we approximate, the dark blue line the one of the cost of the current configuration and the light blue one the minimum numerical cost encountered during the minimization. The green line is the cost of the optimal configuration for the given number of moment constraints, that we aim to compute.



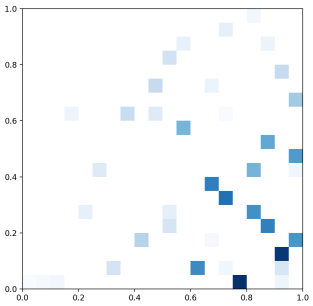
(1) $N = 20$. (2) $N = 60$.

Figure 2.1: Cost in function of the number of iterations for Metropolis-Hastings algorithm and piecewise constant test functions ($N = 20$ and $N = 60$).

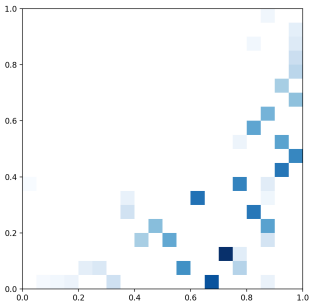
2.6.2 Gradient on a penalized functional

2.6.2.1 Principles

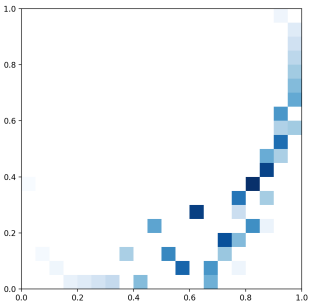
We make use of Theorem 2.3 by searching optima of the MCOT problem with N test functions on each space by looking for an optimal probability measure which is finitely supported on at most $2N+2$ points (note that in the multimarginal case, we can look similarly for measures supported on $DN+2$ points). This algorithm consists in penalizing moments constraints of the MCOT problem for N differentiable test



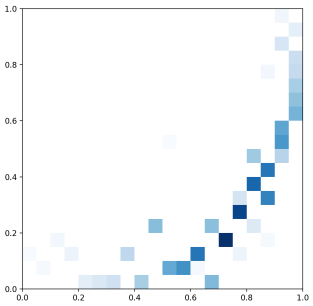
(1) $N = 20$, iteration 0.



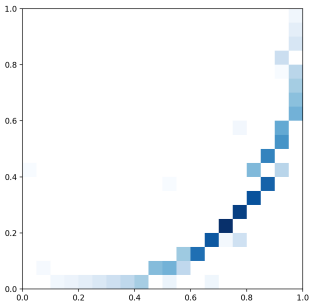
(2) $N = 20$, iteration 2000.



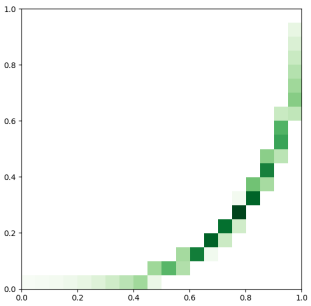
(3) $N = 20$, iteration 4000.



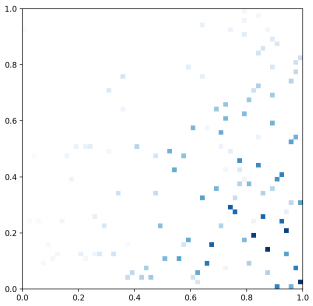
(4) $N = 20$, iteration 6000.



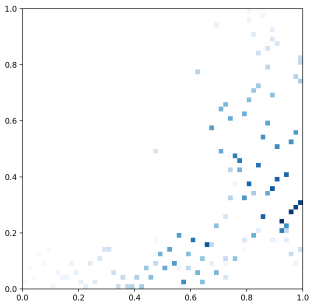
(5) $N = 20$, iteration 10000.



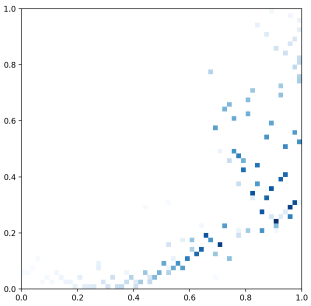
(6) $N = 20$, optimal config.



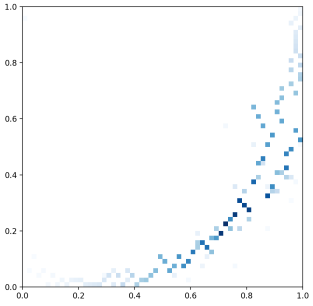
(7) $N = 60$, iteration 0.



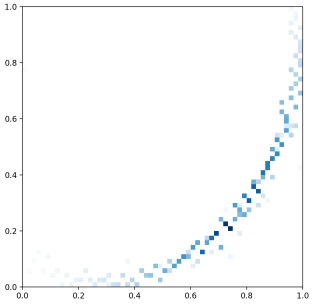
(8) $N = 60$, iteration 3000.



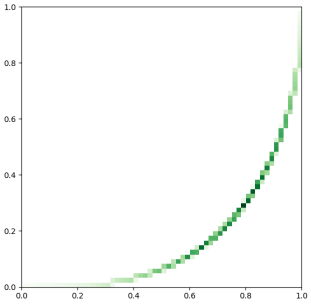
(9) $N = 60$, iteration 20000.



(10) $N = 60$, iter. 60000.



(11) $N = 60$, iter. 100000.



(12) $N = 60$, optimal config.

Figure 2.2: Particles and weights configurations during the optimization using a Metropolis-Hastings algorithm with piecewise constant test functions ($N = 20$ and $N = 60$) for the marginal laws of (2.77) and a quadratic cost.

functions on each space ($(\phi_m)_{1\leq m\leq N}$ and $(\psi_n)_{1\leq n\leq N}$) and then using a gradient-type algorithm to compute the optimum.

For the sake of simplicity, we consider the case of two marginal laws where the cost function c is assumed to be differentiable. Let us write the position of the $2N+2$ particles by $((x_k,y_k))_{1\leq k\leq 2N+2}$ and their weights by $(p_k)_{1\leq k\leq 2N+2}$. Then, it is natural to consider the minimization of

$$\sum_{k=1}^{2N+2} p_k c(x_k,y_k)+\frac{1}{\eta}\left(\sum_{m=1}^{N}\left(\sum_{k=1}^{2N+2} p_k\phi_m(x_k)-\bar{\mu}_m\right)^2+\sum_{n=1}^{N}\left(\sum_{k=1}^{2N+2} p_k\psi_n(y_k)-\bar{\nu}_n\right)^2\right),$$

for some small parameter $\eta>0$ and under the constraints $p_k\geq 0$, $\sum_{k=1}^{2N+2} p_k=1$. To avoid the handling of these latter constraints, we prefer to consider weights $p_k=\frac{e^{a_k}}{\sum_{k=1}^{2N+2}e^{a_k}}$ for some $a_k\in\mathbb{R}$. Although the latter weights cannot be equal to zero, the previous minimization problem is equivalent to minimize

$$\begin{aligned}
&F\left(x_1,...,x_{2N+2},y_1,...,y_{2N+2},a_1,...,a_{2N+2}\right)\\
&\quad=\sum_{k=1}^{2N+2}\frac{e^{a_k}}{\sum_{l=1}^{2N+2}e^{a_l}}c(x_k,y_k)+\frac{1}{\eta}\left(\sum_{m=1}^{N}\left(\sum_{k=1}^{2N+2}\frac{e^{a_k}}{\sum_{l=1}^{2N+2}e^{a_l}}\phi_m(x_k)-\bar{\mu}_m\right)^2\right.\\
&\qquad\left.+\sum_{n=1}^{N}\left(\sum_{k=1}^{2N+2}\frac{e^{a_k}}{\sum_{l=1}^{2N+2}e^{a_l}}\psi_n(y_k)-\bar{\nu}_n\right)^2\right), \quad (2.78)
\end{aligned}$$

since some particles can have the same positions as other ones. For a fixed value of $\eta>0$, we use a projected gradient algorithm (see e.g. Algorithm 1.3.16 of [279]), to ensure that $(x_k,y_k)\in[0,1]^2$ for all k, together with a line search method. We implement alternated gradient steps as follows: first, a gradient step is performed on the coefficients $(a_k)_{1\leq k\leq 2N+2}$ with $(x_k,y_k)_{1\leq k\leq 2N+2}$ fixed; second, a gradient step is done on the positions $(x_k)_{1\leq k\leq 2N+2}$ with the other variables fixed; lastly, a gradient step is done on the positions $(y_k)_{1\leq k\leq 2N+2}$ with the other variables fixed. This procedure is repeated until the norm of the projected gradient is below some error threshold. The convergence of this algorithm is ensured by Wolfe theorem (see Theorem 1.2.21 of [279]).

The example computations exposed thereafter use two sets of test functions: regularized continuous piecewise affine functions and Gaussian test functions. Remark that we do not use discontinuous piecewise affine test functions, for which we have rates of convergence for W_1 and W_2. We make this choice because the gradient algorithm that we describe above has better numerical properties for continuously differentiable test functions.

In the MCOT formulation (2.6) with $M=N$, minimizers of MCOT problems are the same if we consider test functions $(\bar{\phi}_m)_{1\leq m\leq N}$ and $(\bar{\psi}_m)_{1\leq m\leq N}$ such that $\mathrm{Span}((\bar{\phi}_m)_{1\leq m\leq N})=\mathrm{Span}((\phi_m)_{1\leq m\leq N})$ and $\mathrm{Span}((\bar{\psi}_m)_{1\leq m\leq N})=\mathrm{Span}((\psi_m)_{1\leq m\leq N})$. However, in the penalized version of the problem (2.78), the choice of the test functions has a strong impact on the convergence of the gradient algorithms. It appears that considering positive part functions (which are convex functions) greatly improves the efficiency of the procedure with respect to classical hat functions, even if both spans are identical.

Thus, for the numerical examples in 1D, we use the functions for $\epsilon>0$ and for all $N\in\mathbb{N}^*$,

$$\forall x\in[0,1],\quad \varphi_0^N(x)=\begin{cases}-\left(x-\frac{1}{N}\right) & \text{if} \quad x-\frac{1}{N}\leq-\epsilon\\ \frac{1}{4\epsilon}\left(x-\frac{1}{N}-\epsilon\right)^2 & \text{if} \quad -\epsilon\leq x-\frac{1}{N}\leq\epsilon\\ 0 & \text{if} \quad x-\frac{1}{N}\geq\epsilon,\end{cases} \quad (2.79)$$

and for all $1 \leq m \leq N$,

$$\forall x \in [0,1], \quad \varphi_m^N(x) = \begin{cases} 0 & \text{if} \quad x - \frac{m-1}{N} \leq -\epsilon \\ \frac{1}{4\epsilon}\left(x - \frac{m-1}{N} + \epsilon\right)^2 & \text{if} \quad -\epsilon \leq x - \frac{m-1}{N} \leq \epsilon \\ x - \frac{m-1}{N} & \text{if} \quad x - \frac{m-1}{N} \geq \epsilon; \end{cases} \tag{2.80}$$

which are a regularization of the functions, for all $N \in \mathbb{N}^*$, and $1 \leq m \leq \mathbb{N}$,

$$\left(\cdot - \frac{1}{N}\right)^- \quad \text{and} \quad \left(\cdot - \frac{m-1}{N}\right)^+. \tag{2.81}$$

The vector space spanned by the restriction to $[0,1]$ of the latter functions (defined in (2.81)) is the same as the one spanned by the classical continuous piecewise affine functions (i.e. the functions ψ_m^N introduced in Section 2.5.2.1). We also use Gaussian test functions in the 1D numerical examples defined by, for all $0 \leq m \leq N$,

$$\forall x \in [0,1], \quad \varphi_m^{G,N}(x) = \exp\left(-\frac{(x - \frac{m}{N})^2}{2\left(\frac{1}{1.8N}\right)^2}\right). \tag{2.82}$$

For the example in dimension 2, for $N \in \mathbb{N}^*$, we use the following $(N+1)^2$ test functions defined as follows: for all $1 \leq m, n \leq N$ and $(x,y) \in [0,1]^2$,

$$\varphi_{m,n}^N(x,y) = \varphi_{m+n-1}^{2N}\left(\frac{x + y - \tilde{\varphi}_{m-n+1}^N(x-y) - \tilde{\varphi}_{n-m+1}^N(y-x)}{2}\right) \tag{2.83}$$

where for all $q \in \mathbb{Z}$,

$$\forall x \in [0,1], \quad \tilde{\varphi}_q^N(x) = \begin{cases} \varphi_q^N(x) & \text{if} \quad 1 \leq q \leq N, \\ 0 & \text{otherwise,} \end{cases} \tag{2.84}$$

and

$$\varphi_{0,0}^N(x,y) = \varphi_{1,1}^N\left(\frac{1}{N} - x, \frac{1}{N} - y\right). \tag{2.85}$$

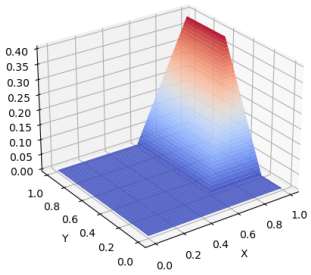
For $1 \leq m, n \leq N$, we set

$$\varphi_{m,0}^N(x,y) = \varphi_{m,1}^N\left(x, \frac{1}{N} - y\right) \quad \text{and} \quad \varphi_{0,n}^N(x,y) = \varphi_{1,n}^N\left(\frac{1}{N} - x, y\right). \tag{2.86}$$

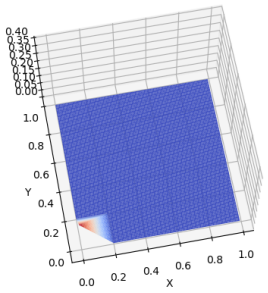
Those functions (plotted in Figure 2.3) are a regularization of the continuous non-negative functions $G_{m,n}^N(x,y) = \left(\min\left(x - \frac{m-1}{N}, y - \frac{n-1}{N}\right)\right)^+$ with $1 \leq m, n \leq N$, $G_{0,0}^N(x,y) = \left(\min\left(\frac{1}{N} - x, \frac{1}{N} - y\right)\right)^+$, $G_{0,m}^N(x,y) = \left(\min\left(x - \frac{m-1}{N}, \frac{1}{N} - y\right)\right)^+$, and $G_{n,0}^N(x,y) = \left(\min\left(\frac{1}{N} - x, y - \frac{n-1}{N}\right)\right)^+$. The vector space spanned by the restriction to $[0,1]^2$ of functions $(G_{m,n}^N)_{0 \leq m,n \leq N}$ is the same as the one spanned by the classical continuous piecewise affine functions associated to the mesh illustrated in Figure 2.4. We use such regularized functions, instead of classical piecewise affine finite elements, for differentiability and efficiency purposes, by analogy with observations in the 1D case.

2.6.2.2 1D numerical example

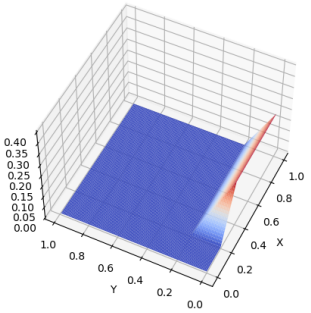
Convergence of the algorithm We tested the algorithm for the marginal laws with densities ρ_μ and ρ_ν defined in Equation (2.77), the quadratic cost function $c(x,y) = |y-x|^2$ and a fixed penalization coefficient $1/\eta$. The exact optimal transport map between μ (abscissa) and ν (ordinate) is represented by the red line on the graphs of Figure 2.7. We present four minimizations:



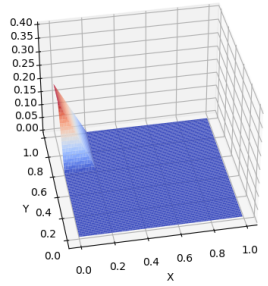
(1) $\varphi^6_{4,2}$



(2) $\varphi^6_{0,0}$



(3) $\varphi^6_{2,0}$



(4) $\varphi^6_{0,4}$

Figure 2.3: Examples of unscaled functions used for the 2D numerical example as defined in (2.83), (2.84), (2.85) and (2.86) for $N = 5$ (out of 36 test functions in total).

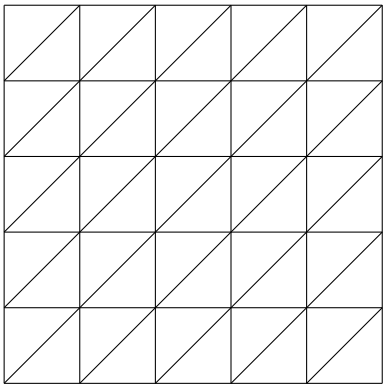
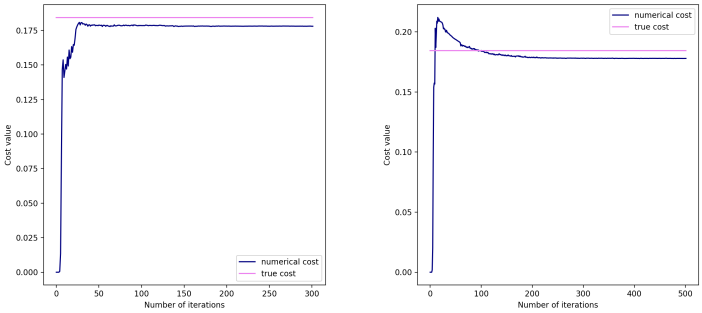


Figure 2.4: Mesh of piecewise affine functions on $[0, 1]^2$.

- Regularized continuous piecewise affine functions (2.80)
  - $N = 10$ and $1/\eta = 100$,
  - $N = 40$ and $1/\eta = 25$.
- Gaussian test functions (2.82)
  - $N = 10$ and $1/\eta = 100$,
  - $N = 20$ and $1/\eta = 100$.

Once each minimization process has converged, in the regularized continuous piecewise affine functions case, the cost for $N = 10$ is 0.17764 and the one for $N = 40$ is 0.17785; in the case with Gaussian test function, the cost for $N = 10$ is 0.17857 and the cost for $N = 20$ is 0.17915. The cost of the optimal transport problem is roughly equal to 0.18444. The convergence of the numerical cost for each case in function of the number of iterations of the gradient algorithm is plotted in Figures 2.5 and 2.6, where the pink line indicates the exact cost of the optimal transport problem that we approximate. The evolution of the configurations through the iterations are represented in each case in the graphs of Figure 2.7. The darker the particle (x_k, y_k), the larger its weight p_k (note that at iteration 0, all the particles have the same weight $1/N$, and we use a darker color to make them visible).

We note on the examples in the regularized continuous piecewise affine functions case (see Figure 2.7) that the particles (x_k, y_k) tend to cluster in some places. This is due to the fact that the cost function is convex and that the test functions are (up to the regularization) locally linear. In contrast, this phenomenon is not observed with Gaussian test function where many particles have a significant weight. Nonetheless, as far as the approximation of the cost is concerned, both choices of test functions lead to a similar accuracy: on our example, the Gaussian test functions lead to a slightly better approximation of the cost.



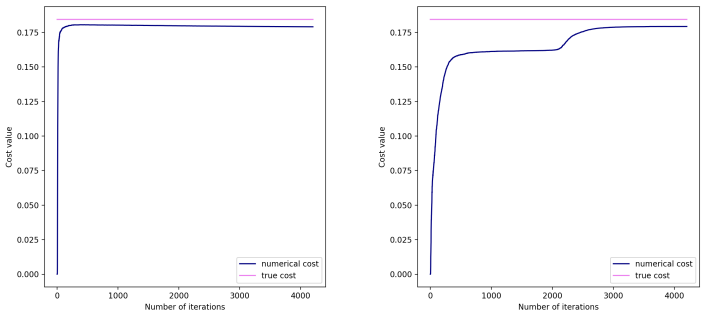
(1) $N = 10$. (2) $N = 40$.

Figure 2.5: Cost in function of the number of iterations in the gradient algorithm for regularized continuous piecewise affine test functions.

2.6.2.3 2D numerical example

We consider two normal marginal laws in $\mathbb{R}^2$: $\mu \sim \mathcal{N}_2(m_\mu, \Sigma_\mu)$ and $\nu \sim \mathcal{N}_2(m_\nu, \Sigma_\nu)$, with

$$m_\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \Sigma_\mu = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad m_\nu = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \Sigma_\nu = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}, \tag{2.87}$$



(1) $N = 10$. (2) $N = 20$.

Figure 2.6: Cost in function of the number of iterations in the gradient algorithm for Gaussian test functions.

and the quadratic cost function. In this case, it is known that the optimal cost is given by $|m_\mu - m_\nu|^2 + \mathrm{Tr}(\Sigma_\mu + \Sigma_\nu - 2(\Sigma_\mu^{1/2}\Sigma_\nu\Sigma_\mu^{1/2})^{1/2})$ and the optimal transport map is given by $x \mapsto m_\nu + \Sigma_\mu^{-1/2}(\Sigma_\mu^{1/2}\Sigma_\nu\Sigma_\mu^{1/2})^{1/2}\Sigma_\nu^{-1/2}$, see e.g. [126]. In Figures 2.8.1 and 2.9, the density of μ (resp. ν) is plotted with different shades of red (resp. blue). We consider regularized piecewise linear test functions on $[-4, 4]^2$ obtained by rescaling the functions (2.83), (2.84), (2.85) and (2.86) on $[0, 1]^2$.

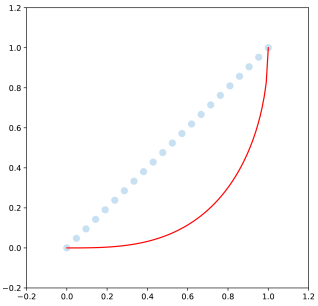
We represent several iterations of the optimization for $N = 36$ and $1/\eta = 2$ in Figure 2.9, where the green arrows represent the transport map computed by the algorithm from μ (red) to ν (blue). The greener the arrow, the more weight it has.

We represent the configuration of particles at convergence on Figure 2.8.1 where each particle consists in a red dot linked to a blue dot. The bigger are the dots, the more mass is transported. The green dots represent the location where the red dot would have been transported if the particle were on the transport plan. Convergence of the cost is represented in Figure 2.8.2 where the pink line represents the cost of the Optimal Transport problem we approximate.

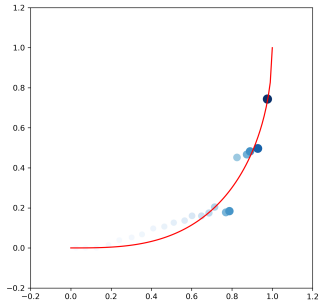
2.6.2.4 Martingale Optimal Transport numerical example

We tested the algorithm for the marginal laws μ and ν being respectively the uniform random variables on $[\frac{1}{4}, \frac{3}{4}]$ and $[0, 1]$, with the cost $c(x, y) = |y - x|^3$. Note that $\int |y-x|^2 \mathrm{d}\pi(x, y) = \int |y|^2 \mathrm{d}\nu(y) - \int |x|^2 \mathrm{d}\mu(x) = 1/16$ for any martingale coupling π. By Jensen's inequality, we have $\int |y-x|^3 \mathrm{d}\pi(x, y) \geq (1/16)^{3/2} = (1/4)^3$ and therefore $\mathrm{d}\pi(x, y) = \mathrm{d}\mu(x)(\frac{1}{2}\mathrm{d}\delta_{x+1/4}(y) + \frac{1}{2}\mathrm{d}\delta_{x-1/4}(y))$ is an optimal martingale coupling and the equality condition in Jensen's inequality shows that this is the unique optimal martingale coupling.

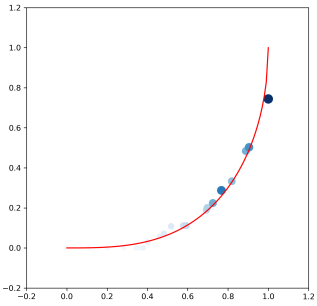
The two lines $y = x + 1/4$ and $y = x - 1/4$ characterizing the optimal martingale coupling are represented by the red lines on Figure 2.11. We have made one minimization with $N = 10$ and $1/\eta = 60$, and $N' = 10$ continuous piecewise affine moment constraints for the martingale constraint, see Problem (2.25). The evolution of the configurations through the iterations are represented in Figure 2.11. The darker the particle (x_k, y_k), the larger the value of its weight p_k. The convergence of the numerical cost is illustrated in Figure 2.10, where the pink line represents the cost of the exact Martingale Optimal Transport problem we approximate.



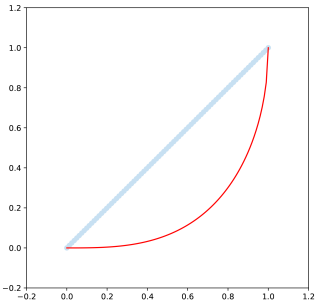
(1) RCPA, $N = 10$, iter. 0.



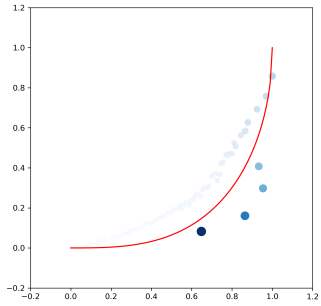
(2) RCPA, $N = 10$, iter. 26.



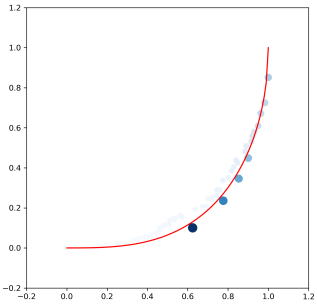
(3) RCPA, $N = 10$, iter. 301.



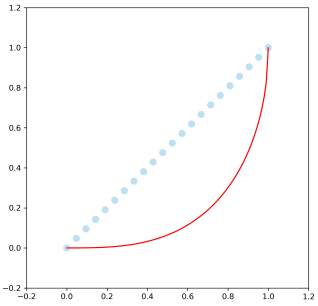
(4) RCPA, $N = 40$, iter. 0.



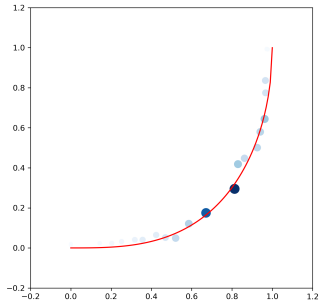
(5) RCPA, $N = 40$, iter. 101.



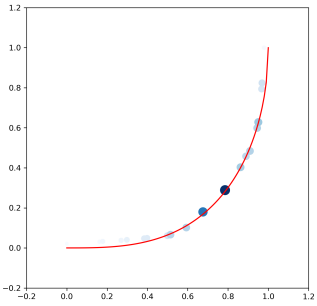
(6) RCPA, $N = 40$, iter. 501.



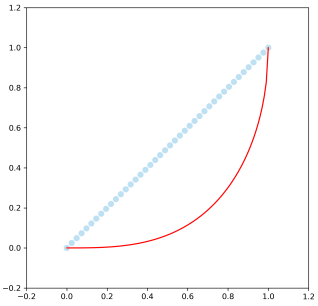
(7) Gauss, $N = 10$, iter. 0.



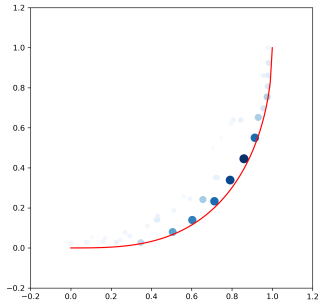
(8) Gauss, $N = 10$, iter. 201.



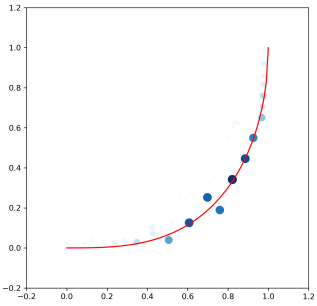
(9) Gauss, $N = 10$, iter. 4201.



(10) Gauss, $N = 20$, iter. 0.

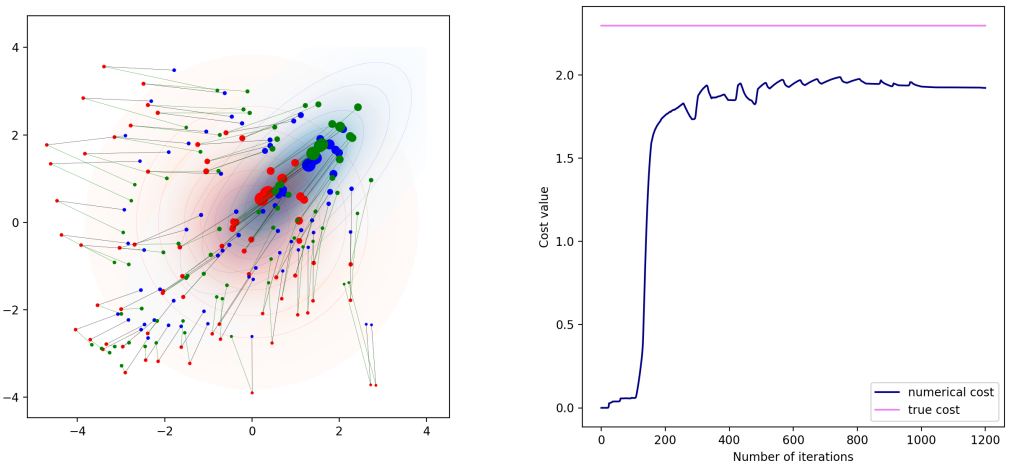


(11) Gauss, $N = 20$, iter. 201.



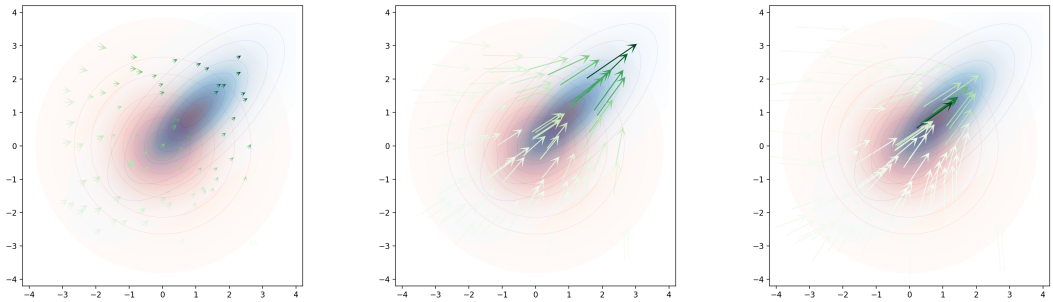
(12) Gauss, $N = 20$, iter. 4201.

Figure 2.7: Particles and weights configurations during the optimization using a gradient-type procedure with regularized continuous piecewise affine (RCPA) test functions and Gaussian test functions (Gauss) for the marginal laws (2.77) and a quadratic cost.



(1) Transport map at convergence. (2) Cost convergence.

Figure 2.8: Cost convergence and approximation of the transport plan at convergence.



(1) iteration 50. (2) iteration 150. (3) iteration 1200.

Figure 2.9: Convergence for two 2D marginal laws with 36 test functions on each set for a quadratic cost.

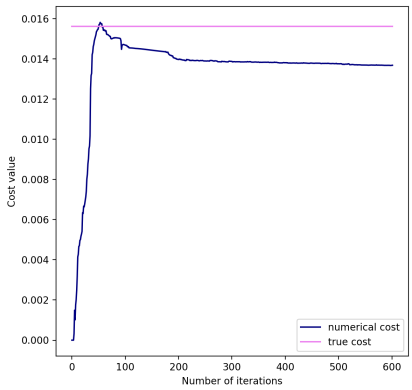
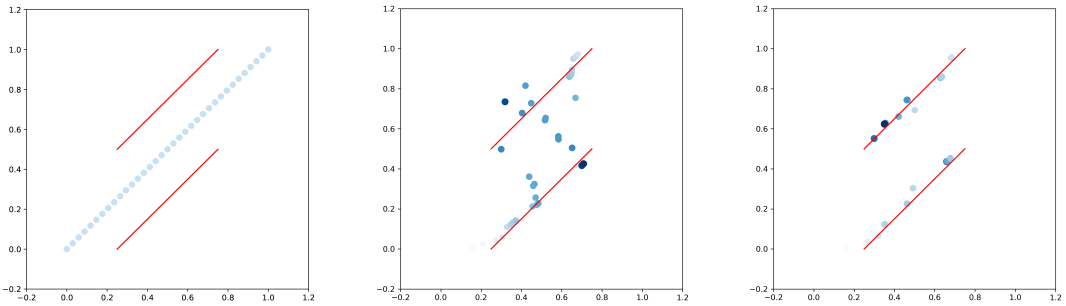


Figure 2.10: Cost in function of the number of iterations in the gradient algorithm.



(1) iteration 0. (2) iteration 61. (3) iteration 601.

Figure 2.11: Convergence with 10 test functions on each set for $c(x, y) = |y - x|^3$

Appendix A

Appendix of Chapter 2

A.1 Technical proofs of Section 2.5

Proof of Proposition 2.14.

Proof. Let us first prove (2.58). Lemma 2.13 implies that

$$\forall 1 \leq m \leq N, \quad \int_{T_m^N} F_1(x)\mathrm{d}x = \int_{T_m^N} F_2(x)\mathrm{d}x \tag{A.1}$$

and

$$\forall 1 \leq k \leq N, \quad F_1\left(\frac{k}{N}\right) = F_2\left(\frac{k}{N}\right). \tag{A.2}$$

Then, using a Taylor expansion, as $F_1, F_2 \in C^2([0,1])$, we get that for all $1 \leq m \leq N$, all $u \in T_m^N$, and all $l = 1, 2$,

$$\left|F_l(u) - F_l\left(\frac{m}{N}\right) - F_l'\left(\frac{m}{N}\right)\left(u - \frac{m}{N}\right)\right| \leq \frac{\|F_l''\|_\infty}{2}\left(u - \frac{m}{N}\right)^2. \tag{A.3}$$

Integrating over T_m^N, one gets

$$\left|\int_{T_m^N} F_l(u)\mathrm{d}u - F_l\left(\frac{m}{N}\right)\frac{1}{N} + F_l'\left(\frac{m}{N}\right)\frac{1}{2N^2}\right| \leq \frac{\|F_l''\|_\infty}{6N^3}.$$

This implies, using (A.1) and (A.2), that

$$\left|F_1'\left(\frac{m}{N}\right) - F_2'\left(\frac{m}{N}\right)\right| \leq \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{3N}. \tag{A.4}$$

Thus, using (A.3), for all $l = 1, 2$ and $u \in T_m^N$,

$$F_l(u) = F_l\left(\frac{m}{N}\right) + \left(u - \frac{m}{N}\right)F_l'\left(\frac{m}{N}\right) + \varphi_l(u),$$

where $|\varphi_l(u)| \leq \frac{\|F_l''\|_\infty}{2}\left(u - \frac{m}{N}\right)^2$. Then, using (A.4), one gets that for all $u \in T_m^N$,

$$|F_1(u) - F_2(u)| \leq \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{3N}\left(\frac{m}{N} - u\right) + \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{2}\left(u - \frac{m}{N}\right)^2. \tag{A.5}$$

Integrating over T_m^N yields that

$$\int_{T_m^N} |F_1(u) - F_2(u)|\,\mathrm{d}u \leq \frac{\|F_1''\|_\infty + \|F_2''\|_\infty}{3N^3}. \tag{A.6}$$

Using the fact that

$$W_1(\mu_1,\mu_2)=\int_{[0,1]}|F_1(u)-F_2(u)|\,\mathrm{d}u=\sum_{m=1}^{N}\int_{T_m^N}|F_1(u)-F_2(u)|\,\mathrm{d}u,$$

we obtain that

$$W_1(\mu_1,\mu_2)\leq\frac{\|F_1''\|_\infty+\|F_2''\|_\infty}{3N^2}.$$

Let us now prove (2.59). The main result needed is the expression of the Wasserstein distance in term of the cumulative distribution functions (cdf) and not their inverse (see [188] Lemma B.3), which holds for $p>1$,

$$W_p^p(F,G)=p(p-1)\int_{\mathbb{R}^2}\mathbf{1}_{\{x<y\}}\left([G(x)-F(y)]^+ + [F(x)-G(y)]^+\right)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y \tag{A.7}$$

because the reasoning of the beginning of this proof introduced a control on the norm between the cdf of the marginal law and the cdf of a marginal law satisfying the same moments.

Then, one can proceed with the following induction. Suppose that we know for $p\in\mathbb{N}^*$ that

$$W_p^p(\mu,\tilde{\mu})\leq\left(\frac{\|F''\|_\infty+\|\tilde{F}''\|_\infty}{3N^2}\right)^p\left(\frac{5}{2}\left(\frac{1}{m_\mu}+\frac{1}{m_{\tilde{\mu}}}\right)\right)^{p-1}p!, \tag{A.8}$$

which holds for $p=1$. Then,

$$\begin{aligned}
&W_{p+1}^{p+1}(\mu,\tilde{\mu})\\
&=(p+1)p\int_{\mathbb{R}^2}\mathbf{1}_{\{x<y\}}\left(\left[F(x)-\tilde{F}(y)\right]^+ + \left[\tilde{F}(x)-F(y)\right]^+\right)(y-x)^{p-1}\mathrm{d}x\mathrm{d}y\\
&=(p+1)p\int_0^1\left(\int_x^1\left(\left[\tilde{F}(x)-F(y)\right]^+ + \left[F(x)-\tilde{F}(y)\right]^+\right)(y-x)^{p-1}\mathrm{d}y\right)\mathrm{d}x\\
&=(p+1)p\int_0^1\left(\int_x^1\left[F(x)-\tilde{F}(y)\right]^+(y-x)^{p-1}\mathrm{d}y\right.\\
&\quad\left.+\int_x^1\left[\tilde{F}(x)-F(y)\right]^+(y-x)^{p-1}\mathrm{d}y\right)\mathrm{d}x.
\end{aligned}$$

Let us treat the first term of the sum, as the second one can be treated symmetrically. If $F(x)\geq\tilde{F}(x)$, we can define $y_x=\tilde{F}^{-1}(F(x))$ and because $\tilde{F}:[0,1]\to[\tilde{F}(0),1]$ is continuous increasing, and we have

$$\begin{aligned}
\int_x^1\left[F(x)-\tilde{F}(y)\right]^+(y-x)^{p-1}\mathrm{d}y&=\int_x^{y_x}\left(F(x)-\tilde{F}(y)\right)(y-x)^{p-1}\mathrm{d}y\\
&\leq\frac{1}{p}\left|F(x)-\tilde{F}(x)\right|(y_x-x)^p.
\end{aligned}$$

Thus, by using (A.3), we get

$$\begin{aligned}
&\int_0^1 \int_x^1 \left[F(x) - \tilde{F}(y)\right]^+ (y-x)^{p-1} \mathrm{d}y\mathrm{d}x \\
&\quad \leq \frac{1}{p} \int_0^1 \mathbf{1}_{\{F(x) \geq \tilde{F}(x)\}} \left|F(x) - \tilde{F}(x)\right| (y_x - x)^p \mathrm{d}x \\
&\quad \leq \frac{1}{p} \sum_{m=1}^N \int_{T_m^N} \mathbf{1}_{\{F(x) \geq \tilde{F}(x)\}} \left|F(x) - \tilde{F}(x)\right| (y_x - x)^p \mathrm{d}x \\
&\quad \leq \frac{1}{p} \sum_{m=1}^N \int_{T_m^N} \left(\frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{3N} \left(\frac{m}{N} - x\right) + \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{2} \left(x - \frac{m}{N}\right)^2 \right) \\
&\qquad \times \mathbf{1}_{\{F(x) \geq \tilde{F}(x)\}} (y_x - x)^p \mathrm{d}x \\
&\quad \leq \frac{5}{6p} \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{N^2} \int_0^1 \mathbf{1}_{\{F(x) \geq \tilde{F}(x)\}} (\tilde{F}^{-1}(F(x)) - F^{-1}(F(x)))^p \mathrm{d}x \\
&\quad \leq \frac{5}{6p} \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{N^2} \int_{F(0)}^1 \mathbf{1}_{\{u \geq \tilde{F}(F^{-1}(u))\}} \left(\tilde{F}^{-1}(u) - F^{-1}(u)\right)^p \frac{\mathrm{d}u}{F'(F^{-1}(u))} \\
&\quad \leq \frac{5}{6p} \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{N^2} \frac{1}{\min_{u \in [0,1]} F'(F^{-1}(u))} \int_0^1 \left|\tilde{F}^{-1}(u) - F^{-1}(u)\right|^p \mathrm{d}u,
\end{aligned}$$

where we used the formula bounding the difference between the cdf (A.5).

Therefore, as $m_\mu > 0$ and $m_{\tilde{\mu}} > 0$, and using the symmetry of the formula (A.7), one gets

$$W_{p+1}^{p+1}(\mu, \tilde{\mu}) \leq \frac{5(p+1)}{2} \frac{\|F''\|_\infty + \|\tilde{F}''\|_\infty}{3N^2} \left(\frac{1}{m_\mu} + \frac{1}{m_{\tilde{\mu}}}\right) W_p^p(\mu, \tilde{\mu}). \tag{A.9}$$

Hence, using the induction hypothesis (A.8), we obtain that (A.8) holds for $p+1$, which gives the claim. □

Lemma A.1. *Let $\mu \in \mathcal{P}([0,1])$ and F_μ its cumulative distribution function. Let $N \in \mathbb{N}^*$. Then, for any $1 \leq m \leq N$, we define $x_m^N \in T_m^N$ by*

$$x_m^N = \begin{cases} \dfrac{m}{N} & \text{if } \; F_\mu\left(\frac{m}{N}\right) = F_\mu\left(\frac{m-1}{N}\right) \\ \dfrac{\int_{T_m^N} F_\mu + \frac{m-1}{N} F_\mu\left(\frac{m-1}{N}\right) - \frac{m}{N} F_\mu\left(\frac{m}{N}\right)}{F_\mu\left(\frac{m}{N}\right) - F_\mu\left(\frac{m-1}{N}\right)} & \text{if } \; F_\mu\left(\frac{m}{N}\right) > F_\mu\left(\frac{m-1}{N}\right), \end{cases}$$

and $\hat{\mu}^N = F_\mu(0)\delta_0 + \sum_{m=1}^N (F_\mu\left(\frac{m}{N}\right) - F_\mu\left(\frac{m-1}{N}\right))\delta_{x_m}$. Then, we have for all $1 \leq m \leq N$,

$$F_{\hat{\mu}^N}\left(\frac{m}{N}\right) = F_\mu\left(\frac{m}{N}\right), \; \int_{T_m^N} F_{\hat{\mu}^N} = \int_{T_m^N} F_\mu, \; \forall x \in T_m^N, \int_{\frac{m-1}{N}}^x F_\mu \geq \int_{\frac{m-1}{N}}^x F_{\hat{\mu}^N}.$$

Besides, if $\mu(\mathrm{d}x) = \rho_\mu(x)\mathrm{d}x$ with a $\rho_\mu \in L^\infty([0,1], \mathrm{d}x; \mathbb{R}_+)$ and $\tilde{\mu} \in \mathcal{P}([0,1])$ is such that $F_{\tilde{\mu}}\left(\frac{m}{N}\right) = F_\mu\left(\frac{m}{N}\right)$ and $\int_{T_m^N} F_{\tilde{\mu}} = \int_{T_m^N} F_\mu$, we have

$$\int_{T_m^N} \left(\int_{\frac{m-1}{N}}^x F_\mu\right) \mathrm{d}x \leq \int_{T_m^N} \left(\int_{\frac{m-1}{N}}^x F_{\tilde{\mu}}\right) \mathrm{d}x + \frac{\|\rho_\mu\|_\infty}{6N^3} \tag{A.10}$$

Proof. If $F_\mu\left(\frac{m}{N}\right) > F_\mu\left(\frac{m-1}{N}\right)$, we have

$$\frac{1}{N} F_\mu\left(\frac{m-1}{N}\right) \leq \int_{T_m^N} F_\mu < \frac{1}{N} F_\mu\left(\frac{m}{N}\right)$$

since F_μ is non-decreasing and right-continuous. Therefore, there is a unique $x_m^N \in T_m^N$ such that

$$\left(x_m^N - \frac{m-1}{N}\right) F_\mu\left(\frac{m-1}{N}\right) + \left(\frac{m}{N} - x_m^N\right) F_\mu\left(\frac{m}{N}\right) = \int_{T_m^N} F_\mu,$$

which is precisely the definition of x_m^N. By construction, we have $F_{\hat{\mu}^N}\left(\frac{m}{N}\right) = F_\mu\left(\frac{m}{N}\right)$ and the previous equation gives

$$\int_{T_m^N} F_{\hat{\mu}^N} = \int_{\frac{m-1}{N}}^{x_m^N} F_\mu\left(\frac{m-1}{N}\right) dx + \int_{x_m^N}^{\frac{m}{N}} F_\mu\left(\frac{m}{N}\right) dx = \int_{T_m^N} F_\mu$$

when $F_\mu\left(\frac{m}{N}\right) > F_\mu\left(\frac{m-1}{N}\right)$ (this identity is obvious if $F_\mu\left(\frac{m}{N}\right) = F_\mu\left(\frac{m-1}{N}\right)$). Last, since for $x \in T_m^N$, $F_\mu\left(\frac{m-1}{N}\right) \leq F_\mu(x) \leq F_\mu\left(\frac{m}{N}\right)$, we get that $x \mapsto \int_{\frac{m-1}{N}}^x (F_\mu - F_{\hat{\mu}^N})$ is non-decreasing on $[\frac{m-1}{N}, x_m^N]$, non-increasing on $[x_m^N, \frac{m}{N}]$ and vanishes for $x \in \{\frac{m-1}{N}, \frac{m}{N}\}$: it is therefore non-negative on T_m^N.

Now, let us assume that μ has a bounded density probability function ρ_μ. We have

$$\begin{aligned}
x \in \left[\frac{m-1}{N}, x_m^N\right], \int_{\frac{m-1}{N}}^x (F_\mu - F_{\hat{\mu}^N}) &= \int_{\frac{m-1}{N}}^x \int_{\frac{m-1}{N}}^z \rho_\mu(u)\mathrm{d}u\mathrm{d}z \leq \frac{\|\rho_\mu\|_\infty}{2}(x - \frac{m-1}{N})^2, \\
x \in \left[x_m^N, \frac{m}{N}\right], \int_{\frac{m-1}{N}}^x (F_\mu - F_{\hat{\mu}^N}) &= -\int_x^{\frac{m}{N}} (F_\mu - F_{\hat{\mu}^N}) \\
&= \int_x^{\frac{m}{N}} \int_z^{\frac{m}{N}} \rho_\mu(u)\mathrm{d}u\mathrm{d}z \leq \frac{\|\rho_\mu\|_\infty}{2}(\frac{m}{N} - x)^2,
\end{aligned}$$

and therefore

$$\int_{T_m^N} \left(\int_{\frac{m-1}{N}}^x (F_\mu - F_{\hat{\mu}^N})\right) \mathrm{d}x \leq \frac{\|\rho_\mu\|_\infty}{6}\left[\left(x_m^N - \frac{m-1}{N}\right)^3 + \left(\frac{m}{N} - x_m^N\right)^3\right] \leq \frac{\|\rho_\mu\|_\infty}{6N^3}. \tag{A.11}$$

Now, we observe that we either have $\int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_\mu\right)\mathrm{d}x \leq \int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_{\tilde{\mu}}\right)\mathrm{d}x$ or $\int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_\mu\right)\mathrm{d}x \geq \int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_{\tilde{\mu}}\right)\mathrm{d}x \geq \int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_{\hat{\mu}^N}\right)\mathrm{d}x$. In the first case, the claim is obvious. In the second one, we then have

$$\int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_\mu\right)\mathrm{d}x \leq \int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x F_{\tilde{\mu}}\right) + \int_{T_m^N}\left(\int_{\frac{m-1}{N}}^x (F_\mu - F_{\hat{\mu}^N})\right),$$

and we get the result using (A.11). □

A.2 Refinements of Theorem 2.3 and Theorem 2.7

We prove in this section some additional results which may be seen as refinements of Theorem 2.3 and Theorem 2.7.

A.2.1 Existence of discrete minimizers for MCOT problems: case of compactly supported test functions

As announced in Remark 2.1 (iv), an alternative statement of Theorem 2.3 which avoids imposing the constraint $\int_{\mathcal{X}\times\mathcal{Y}}(\theta_\mu(|x|) + \theta_\nu(|y|))\mathrm{d}\pi(x,y) \leq A$ can be obtained

under stronger assumptions on the test functions and the cost. In all Subsection A.2.1, we consider the case

$$\mathcal{X} = \mathbb{R}^{d_x} \text{ and } \mathcal{Y} = \mathbb{R}^{d_y},$$

for some $d_x, d_y \in \mathbb{N}^*$, and assume that the cost c is continuous and satisfies:

$$\forall x \in \mathcal{X},\ c(x,y) \underset{|y|\to+\infty}{\longrightarrow} +\infty,\ \forall y \in \mathcal{Y}, c(x,y) \underset{|x|\to+\infty}{\longrightarrow} +\infty, \tag{A.12}$$

$$\exists (x_n) \in \mathcal{X}^{\mathbb{N}}, (y_n) \in \mathcal{Y}^{\mathbb{N}}, |x_n| \to +\infty, |y_n| \to +\infty \text{ and } c(x_n, y_n) = 0. \tag{A.13}$$

This condition is satisfied for example when $d_x = d_y$ and $c(x,y) = H(|x-y|)$, with H continuous satisfying $H(0) = 0$ and $H(r) \underset{r\to+\infty}{\to} +\infty$. We assume also that the test functions ϕ_m, ψ_n, $1 \le m, n \le N$ are continuous with compact support, and define their compact support as follows

$$\mathcal{S}_{\mathcal{X}} := \overline{\{x \in \mathcal{X}, \quad \exists 1 \le m \le N,\ \phi_m(x) \neq 0\}},$$
$$\mathcal{S}_{\mathcal{Y}} := \overline{\{y \in \mathcal{Y}, \quad \exists 1 \le n \le N,\ \psi_n(y) \neq 0\}}.$$

Let $M = \max_{x,y \in \mathcal{S}_{\mathcal{X}} \times \mathcal{S}_{\mathcal{Y}}} c(x,y)$ and let us define

$$\widetilde{\mathcal{S}}_{\mathcal{X}} = \{x \in \mathcal{X} : \exists y \in \mathcal{S}_{\mathcal{Y}},\ c(x,y) \le M+1\} \tag{A.14}$$

$$\widetilde{\mathcal{S}}_{\mathcal{Y}} = \{y \in \mathcal{Y} : \exists x \in \mathcal{S}_{\mathcal{X}},\ c(x,y) \le M+1\} \tag{A.15}$$

together with

$$\mathcal{K} = \left(\mathcal{S}_{\mathcal{X}} \times \tilde{\mathcal{S}}_{\mathcal{Y}}\right) \cup \left(\tilde{\mathcal{S}}_{\mathcal{X}} \times \mathcal{S}_{\mathcal{Y}}\right).$$

It can be easily seen that $\tilde{\mathcal{S}}_{\mathcal{X}}$ (resp. $\tilde{\mathcal{S}}_{\mathcal{Y}}$) is a compact set that contains $\mathcal{S}_{\mathcal{X}}$ (resp. $\mathcal{S}_{\mathcal{Y}}$), and thus the set $\mathcal{K}$ is compact. Then, from (A.13), we take an arbitrary point $(\bar{x}, \bar{y}) \notin \mathcal{K}$ such that $c(\bar{x}, \bar{y}) = 0$, and we define

$$\bar{\mathcal{K}} = \mathcal{K} \cup \{(\bar{x}, \bar{y})\}. \tag{A.16}$$

Lemma A.2. *Let $K \in \mathbb{N}^*$, and for all $1 \le k \le K$, $x_k \in \mathcal{X}$, $y_k \in \mathcal{Y}$, $p_k \ge 0$ such that $\sum_{k=1}^K p_k = 1$. If $\gamma = \sum_{k=1}^K p_k \delta_{x_k, y_k} \in \Pi(\mu, \nu; (\phi_m)_{1\le m\le N}, (\psi_n)_{1\le n\le N})$ then there exist K points $(\tilde{x}_k, \tilde{y}_k) \in \bar{\mathcal{K}}$ for $1 \le k \le K$ such that the discrete probability measure $\tilde{\gamma} = \sum_{k=1}^K p_k \delta_{\tilde{x}_k, \tilde{y}_k} \in \Pi(\mu, \nu; (\phi_m)_{1\le m\le N}, (\psi_n)_{1\le n\le N})$ and*

$$\sum_{k=1}^K p_k c(\tilde{x}_k, \tilde{y}_k) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\tilde{\gamma}(x,y) \le \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) \mathrm{d}\gamma(x,y) = \sum_{k=1}^K p_k c(x_k, y_k).$$

Proof. Consider a measure $\gamma = \sum_{k=1}^K p_k \delta_{x_k, y_k}$ satisfying the assumptions of Lemma A.2. We construct $\tilde{\gamma} = \sum_{k=1}^K p_k \delta_{\tilde{x}_k, \tilde{y}_k}$ using the following procedure.

Case 1. If $(x_k, y_k) \in \mathcal{K}$, then we define $(\tilde{x}_k, \tilde{y}_k) = (x_k, y_k)$.

Case 2. If $x_k \notin \mathcal{S}_{\mathcal{X}}$ and $y_k \notin \mathcal{S}_{\mathcal{Y}}$, then we define $(\tilde{x}_k, \tilde{y}_k) = (\bar{x}, \bar{y})$.

Case 3. Let us suppose $x_k \in \mathcal{S}_{\mathcal{X}}$ and $y_k \notin \tilde{\mathcal{S}}_{\mathcal{Y}}$ (the case $y_k \in \mathcal{S}_{\mathcal{Y}}$ and $x_k \notin \tilde{\mathcal{S}}_{\mathcal{X}}$ is treated symmetrically). By definition of $\tilde{\mathcal{S}}_{\mathcal{Y}}$, we have

$$\forall x \in \mathcal{S}_{\mathcal{X}}, \quad c(x, y_k) > M+1.$$

In particular, we have $c(x_k, y_k) > M + 1$. Let $y^* \in \mathcal{S}_\mathcal{Y}$. Then,

$$c(x_k, y^*) \leq \max_{x,y \in \mathcal{S}_\mathcal{X} \times \mathcal{S}_\mathcal{Y}} c(x, y) = M$$

Let $y_\lambda := (1 - \lambda)y^* + \lambda y_k$ for $\lambda \in [0, 1]$. As c is continuous, there exists λ^* such that $c(x_k, y_{\lambda^*}) = \frac{2M+1}{2}$. Then, $y_{\lambda^*} \notin \mathcal{S}_\mathcal{Y}$ because $\frac{2M+1}{2} > M$, and $y_{\lambda^*} \in \widetilde{\mathcal{S}_\mathcal{Y}}$. Then, we define $(\tilde{x}_k, \tilde{y}_k) = (x_k, y_{\lambda^*})$.

This construction preserves the points in the supports $\mathcal{S}_\mathcal{X}$ and $\mathcal{S}_\mathcal{Y}$, and the points outside the supports are replaced by other points outside the supports. Thus, we have

$$\begin{aligned} \forall 1 \leq m \leq N, \quad \sum_{k=1}^{K} p_k \phi_m(\tilde{x}_k) &= \sum_{k=1}^{N} p_k \phi_m(x_k) \\ \forall 1 \leq n \leq N, \quad \sum_{k=1}^{K} p_k \psi_n(\tilde{y}_k) &= \sum_{k=1}^{N} p_k \psi_n(y_k), \end{aligned}$$

and the moment constraints are satisfied by $\tilde{\gamma}$. In addition, it is clear that the cost does not change in Case 1 and is lowered in Cases 2 and 3. □

Proposition A.3. *Let us assume that $\mathcal{X} = \mathbb{R}^{d_x}$, $\mathcal{Y} = \mathbb{R}^{d_y}$ and $c : \mathbb{R}^{d_x} \times \mathbb{R}^{d_y} \to \mathbb{R}_+$ is continuous and satisfies* (A.12), (A.13).

Let us assume that for all $1 \leq m, n \leq N$, ϕ_m and ψ_n are compactly supported real-valued continuous functions defined on $\mathbb{R}^d$. Then, there exists at least one minimizer to the minimization problem

$$I^N = \inf_{\pi \in \Pi(\mu,\nu;(\phi_m)_{1\leq m\leq N},(\psi_n)_{1\leq n\leq N})} \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\pi(x, y). \tag{A.17}$$

Moreover, there exists $K \in \mathbb{N}$ such that $K \leq 2N + 2$, and for all $1 \leq k \leq K$, $(x_k, y_k) \in \bar{\mathcal{K}}$, $p_k \geq 0$ such that $\sum_{k=1}^{K} p_k = 1$ such that $\widetilde{\pi} := \sum_{k=1}^{K} p_k \delta_{x_k,y_k}$ is a minimum.

Proof. Let us consider a minimizing sequence $(\pi_l)_{l\in\mathbb{N}}$ for Problem (A.17). For all $l \in \mathbb{N}$, we will denote by γ_l a finite discrete measure which has the same cost and same moments than π_l, with at most $2N+2$ points, which exists thanks to Theorem 2.1, and the fact that the test functions are compactly supported. Then, using Lemma A.2, for all $l \in \mathbb{N}$, one can define a measure $\tilde{\gamma}_l$ which satisfies the moment constraints, has a support contained in the set $\bar{\mathcal{K}}$ defined in (A.16), and such that,

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\tilde{\gamma}_l(x, y) \leq \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\gamma_l(x, y).$$

Thus, $(\tilde{\gamma}_l)_{l\in\mathbb{N}}$ is a minimizing sequence. Besides, $(\tilde{\gamma}_l)_{l\in\mathbb{N}}$ is tight since the support of $\widetilde{\gamma}_l$ is included in the compact set $\bar{\mathcal{K}}$ for all $l \in \mathbb{N}$. Then, following the same lines as in the proof of Theorem 2.3, one can extract a weakly converging subsequence, the cost of the limit of which is equal to I^N. The fact that there exists a finite discrete measure charging at most $K \leq 2N + 2$ points can be deduced from Theorem 2.1, following the same lines as in the proof of Theorem 2.3. □

A.2.2 Convergence of the MCOT problem towards the OT problem: bounded test functions on compact sets

We now assume that $\mathcal{X}$ and $\mathcal{Y}$ are compact subsets of $\mathbb{R}^{d_x}$ and $\mathbb{R}^{d_y}$. As announced in Remark 2.4 (ii), we state a result analogous to Theorem 2.7 which holds without the additional moment constraint and for possibly discontinuous test functions. We consider two sequences of bounded measurable real-valued test functions $(\phi_m)_{m\in\mathbb{N}^*} \subset L^\infty(\mathcal{X})$ and $(\psi_n)_{n\in\mathbb{N}^*} \subset L^\infty(\mathcal{Y})$ that satisfy

$$\forall f \in C^0(\mathcal{X}), \inf_{v_N \in \mathrm{Span}\{\phi_m,\, 1\le m\le N\}} \|f - v_N\|_\infty \underset{N\to+\infty}{\longrightarrow} 0 \tag{A.18}$$

and

$$\forall f \in C^0(\mathcal{Y}), \inf_{v_N \in \mathrm{Span}\{\psi_n,\, 1\le n\le N\}} \|f - v_N\|_\infty \underset{N\to+\infty}{\longrightarrow} 0. \tag{A.19}$$

It is easy then to see that the properties (2.33) and (2.34) are satisfied for any $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. For any $N \ge 1$, we consider the following MCOT problem:

$$I^N = \inf_{\pi\in\Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y) \right\}. \tag{A.20}$$

Proposition A.4. *Let us assume that $\mathcal{X}$ and $\mathcal{Y}$ are compact sets and let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Let $(\phi_m)_{m\in\mathbb{N}^*} \subset L^\infty(\mathcal{X})$ and $(\psi_n)_{n\in\mathbb{N}^*} \subset L^\infty(\mathcal{Y})$ satisfying* (A.18) *and* (A.19)*. Let us assume that $I < +\infty$. Then, it holds that $I^N \le I$ and*

$$I^N \xrightarrow[N\to\infty]{} I.$$

Moreover, from every sequence $(\pi^N)_{N\in\mathbb{N}}$ such that for all $N \in \mathbb{N}^*$, the measure $\pi^N \in \Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})$ satisfies

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi^N(x,y) \le I_N + \epsilon_N, \tag{A.21}$$

with $\epsilon_N \underset{n\to+\infty}{\longrightarrow} 0$, one can extract a subsequence which converges towards a measure $\pi^\infty \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ which is a minimizer to Problem (2.35)*.*

Remark A.1. *From Theorem 2.1, there exists $0 \le K_N \le 2N+2$, $x_1, \cdots, x_{K_N} \in \mathcal{X}$, $y_1, \cdots, y_{K_N} \in \mathcal{Y}$ and $w_1, \cdots, w_{K_N} \ge 0$ such that $\gamma^N := \sum_{k=1}^{K_N} w_k \delta_{(x_k,y_k)} \in \Pi(\mu,\nu;(\phi_m)_{1\le m\le N},(\psi_n)_{1\le n\le N})$ and*

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\gamma^N(x,y) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi^N(x,y) \le I_N + \epsilon_N. \tag{A.22}$$

In other words, any sequence $(\pi^N)_{N\in\mathbb{N}^*}$ satisfying the assumptions of Proposition A.4 can be chosen as a discrete measure charging at most $2N+2$ points.

Proof of Proposition A.4. Since $\mathcal{X}$ and $\mathcal{Y}$ are compact, the sequence (π^N) is tight, and we can assume, up to the extraction of a subsequence, that it weakly converges to π^∞. For $N \in \mathbb{N}^* \cup \{\infty\}$, we denote the marginal laws of π^N respectively by $\mathrm{d}\mu^N(x) := \int_{\mathcal{Y}} \mathrm{d}\pi^N(x,y)$ and $\mathrm{d}\nu^N(y) := \int_{\mathcal{X}} \mathrm{d}\pi^N(x,y)$. For $f \in C^0(\mathcal{X})$, it holds that

$$\int_{\mathcal{X}} f\mathrm{d}\mu^N \xrightarrow[N\to\infty]{} \int_{\mathcal{X}} f\mathrm{d}\mu^\infty.$$

Let $\epsilon > 0$. Using the density condition (A.18), one can find $M \in \mathbb{N}^*$ and $\lambda_1, ..., \lambda_M \in \mathbb{R}$ such that $\sup_{x\in\mathcal{X}} \left| f(x) - \sum_{i=1}^{M} \lambda_i \phi_i(x) \right| \leq \epsilon$. Thus,

$$\left| \int_{\mathcal{X}} f \mathrm{d}\mu - \sum_{i=1}^{M} \lambda_i \mu_i \right| \leq \epsilon \tag{A.23}$$

and for $K > M$, $\left| \int_{\mathcal{X}} f \mathrm{d}\mu^K - \sum_{i=1}^{M} \lambda_i \int_{\mathcal{X}} \phi_i \mathrm{d}\mu^K \right| \leq \epsilon$, i.e.

$$\left| \int_{\mathcal{X}} f \mathrm{d}\mu^K - \sum_{i=1}^{M} \lambda_i \mu_i \right| \leq \epsilon. \tag{A.24}$$

Then, (A.23) and (A.24) imply that $\left| \int_{\mathcal{X}} f \mathrm{d}\mu^K - \int_{\mathcal{X}} f \mathrm{d}\mu \right| \leq 2\epsilon$, and taking $K \to \infty$ leads to

$$\left| \int_{\mathcal{X}} f \mathrm{d}\mu^\infty - \int_{\mathcal{X}} f \mathrm{d}\mu \right| \leq 2\epsilon. \tag{A.25}$$

As (A.25) holds for any $\epsilon > 0$, one gets that for any $f \in C^0(\mathcal{X})$,

$$\int_{\mathcal{X}} f \mathrm{d}\mu^\infty = \int_{\mathcal{X}} f \mathrm{d}\mu,$$

which yields that $\mu^\infty = \mu$. Similarly, we have $\nu^\infty = \nu$. Therefore, $\pi^\infty \in \Pi(\mu, \nu)$ and

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\pi^\infty(x, y) \geq I. \tag{A.26}$$

Now, we use the same arguments as in the proof of Theorem 2.7 to deduce that $\int_{\mathcal{X}\times\mathcal{Y}} c(x, y) \mathrm{d}\pi^\infty(x, y) \leq I$, which gives the result. □

Chapter 3

Constrained overdamped Langevin dynamics for symmetric multimarginal optimal transportation

This chapter is an article written with Aurélien Alonsi and Virginie Ehrlacher and submitted to *Mathematical Models and Methods in Applied Sciences* [7].

Abstract

The Strictly Correlated Electrons (SCE) limit of the Levy-Lieb functional in Density Functional Theory (DFT) gives rise to a symmetric multi-marginal optimal transport problem with Coulomb cost, where the number of marginal laws is equal to the number of electrons in the system, which can be very large in relevant applications. In this work, we design a numerical method, built upon constrained overdamped Langevin processes to solve Moment Constrained Optimal Transport (MCOT) relaxations (introduced in Chapter 2 and in A. Alfonsi, R. Coyaud, V. Ehrlacher and D. Lombardi, Math. Comp. 90, 2021, 689–737) of symmetric multi-marginal optimal transport problems with Coulomb cost. Some minimizers of such relaxations can be written as discrete measures charging a low number of points belonging to a space whose dimension, in the symmetrical case, scales linearly with the number of marginal laws. We leverage the sparsity of those minimizers in the design of the numerical method and prove that any local minimizer to the resulting problem is actually a *global* one. We illustrate the performance of the proposed method by numerical examples which solves MCOT relaxations of 3D systems with up to 100 electrons.

3.1 Introduction

Optimal transport (OT) problems [291, 319] appear in numerous application fields such as data science [277], finance [27], economics [82, 146, 147] or physics [318]. Hence an increasing interest in developing efficient numerical methods for this types of problems among the applied mathematics community.

In this article, we specifically focus on multi-marginal symmetric optimal transportation problems arising from quantum chemistry. Density Functional Theory (DFT) [268] is one of the most popular theories in quantum chemistry in order to compute the ground state of electrons within a molecule. It is exact in principle,

due to the Hohenberg-Kohn theorem, up to the knowledge of the Levy-Lieb functional, which is unfortunately not computable in practice. Hence, a wide zoology of electronic structure models have been developped in the chemistry community where approximations of this Levy-Lieb functional are computed [228]. Actually, it has been recently proved [50, 66, 67, 106, 107, 142, 226] that the semi-classical limit of this Levy-Lieb functional is the solution of a symmetric multi-marginal optimal transport problem which we state now.

For all $p \in \mathbb{N}^*$ (where $\mathbb{N}^*$ denotes the set of positive integers $\{1, 2, 3, \ldots\}$), we denote by $\mathcal{P}(\mathbb{R}^p)$ the set of probability measures on $\mathbb{R}^p$. For $d \in \mathbb{N}^*$, for all $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $M \in \mathbb{N}^*$ a fixed number of marginal laws (the number of electrons in DFT), we will denote the set of M-couplings for μ by

$$\Pi(\mu; M) := \left\{ \pi \in \mathcal{P}\left((\mathbb{R}^d)^M\right) : \forall 1 \leq m \leq M, \int_{(\mathbb{R}^d)^{M-1}} d\pi(x_1, \ldots, x_M) = d\mu(x_m) \right\}. \tag{3.1}$$

Let $c : (\mathbb{R}^d)^M \to \mathbb{R}_+ \cup \{+\infty\}$ be a M-symmetric (i.e. such that for all $(x_1, \cdots, x_M) \in \left(\mathbb{R}^d\right)^M$, $c(x_1, \ldots, x_M) = c(x_{\sigma(1)}, \ldots, x_{\sigma(M)})$ for $\sigma \in \mathcal{S}_M$ a M-permutation) non-negative lower semi-continuous (l.s.c.) function. The function c is called hereafter the *cost function*. Then, the multimarginal symmetric optimal transport problem associated to μ, M and c is defined as

$$I(\mu) = \inf_{\pi \in \Pi(\mu; M)} \int_{(\mathbb{R}^d)^M} c(x_1, \ldots, x_M) d\pi(x_1, \ldots, x_M). \tag{3.2}$$

In DFT applications, the cost c is defined as the Coulomb cost $c(x_1, \ldots, x_M) = \sum_{m_1 < m_2} \frac{1}{|x_{m_1} - x_{m_2}|}$. Then, this multimarginal symmetric optimal transport problem allows to compute the interaction energy between electrons, given an electronic density (equal to $M\mu$), in the Strictly Correlated Electrons (SCE) limit – bringing interest in numerical methods for large multimarginal systems.

A straightforward discretization of problem (3.2) (using a discretization of the state space $\mathbb{R}^d$ with a discrete d dimensional grid for instance) leads to a linear programming problem, whose size scales exponentially with M. Hence, for large values of M, specific numerical methods have to be used in order to circumvent the curse of dimensionality. Hence, new application or efficiency oriented approaches have been developed for such problems, using entropic relaxation and the Sinkhorn algorithm [41, 42], dual formulations of the problem [251] or sparsity structure of the minimizers of the discrete problems [143, 320], which can be combined with a semidefinite relaxation [192, 193].

In Chapter 2, the authors considered a relaxation of the optimal transport problem (Moment Constrained Optimal Transport – MCOT) which boils down to considering a particular instance of Generalized Moment Problem [179, 210, 211]. The idea of the proposed approach is to change the discretization approach in the sense the state space $\mathbb{R}^d$ is not discretized anymore, but the marginal constraints in (3.2) are relaxed into a finite number of moment constraints. Taking advantage of the M-symmetry of the problem, it was proved in Proposition 2.5 that some minimizers of the obtained relaxed problems could be written as discrete measures charging a low number of points which scales independently of M.

Thus, a natural idea inspired from this result is to restrict the minimization set considered in the MCOT problem to the set of probability measures of $(\mathbb{R}^d)^M$ which can be written as discrete measures charging a low number of points and satisfying

the associated moment constraints. The resulting problem, called hereafter the *particle problem*, amounts to optimize the positions of the points and the weights charging the associated Dirac measures[i]. In principle, the low number of points needed to obtain a representation of a minimizer to the MCOT problem should help in tackling the curse of dimensionality. However, the non-convexity of the particle problem remains a numerical challenge.

One of the first contribution of this paper is to prove that, despite the non-convexity of the obtained particle problem, any of its local minimizers are actually **global minimizers**. Besides, we prove that the set of local minimizers, which is hence identical to the set of global minimizers, is polygonally connected. This first result is stated in Section 3.2 of the article.

The second contribution of the paper is to propose a numerical scheme in order to find an optimum solution to the particle problem. The numerical method builds on the use of a constrained overdamped Langevin process projected on a submanifold defined by the constraints of the problem, in the spirit of [99, 217, 218, 219, 220, 324]. Such processes are actually already used in the context of molecular dynamics (for which the constraint is defined through the use of a so-called *reaction coordinate* function). We give in this paper some elements of theoretical analysis justifying the interest of such processes for the resolution of multi-marginal optimal transportation problems and outline the link between such constrained overdamped Langevin processes and entropic regularization of optimal transport problems. This is the object of Section 3.3. Finally, we present the numerical scheme we consider in this article in Section 3.4 and the numerical results obtained with this approach in Section 3.5. Proofs of our main theoretical results are postponed until Section 3.6.

We want here to stress on the fact that this numerical scheme enabled us to obtain approximations of solutions to (3.2) for very high-dimensional problems, for instance in cases where $d = 3$ and $M = 100$. Such a method thus appears to be a very promising approach in order to solve large-scale problems in DFT for systems involving a large number of electrons.

Let us point out here that algorithms based on constrained overdamped Langevin dynamics can also be used in principle for the resolution of general multimarginal optimal transport problems and multimarginal martingale optimal transport problems, as there exist an MCOT approximation for both types of problems (see Section 2.3.2). In these cases, the number of marginal constraints to be imposed scales linearly in M[ii], hence the practical implementation of the numerical method proposed in this paper is more intricate than in the symmetric case studied here, where the number of constraints is independant of M.

[i]Note that we use, in this article, the term *particle* to designate a *Dirac measure* (seen in the minimization problem as a vector in $\mathbb{R}_+ \times (\mathbb{R}^d)^M$ accounting for a nonnegative weight and the coordinates of a point in $(\mathbb{R}^d)^M$), and not with the physics meaning that encompasses electrons – the electronic density of which, in the DFT application, would correspond in this article to M times the marginal law μ.

[ii]In the case of multimarginal martingale optimal transport, if there is no assumption of Markovian relationship between the marginal laws, the scaling in the number of constraints for the approximation of the martingale constraints may be exponential in M.

3.2 Mathematical properties of MCOT particle problems

We recall in this section the MCOT problem which was introduced in Chapter 2, together with the associated particle problem. We also state here our first theoretical results which describe the set of minimizers associated to the particle problem.

3.2.1 MCOT and particle problems

As introduced in Chapter 2, the Moment Constrained Optimal Transport (MCOT) problem is a particular case of Generalized Moment Problem [211] which may be seen as a relaxation of optimal transport where the marginal constraints are alleviated and replaced by a finite number of moment constraints. In the following, we restrain our analysis to symmetrical multimarginal optimal transport for the sake of clarity but let us mention here that the results presented here can be extended to general multimarginal optimal transport, as well as martingale optimal transport.

Let $d \in \mathbb{N}^*$, $\mu \in \mathcal{P}(\mathbb{R}^d)$, $M \in \mathbb{N}^*$ and $c : (\mathbb{R}^d)^M \to \mathbb{R}_+ \cup \{+\infty\}$ be a lower semi-continuous symmetric function. The MCOT problem is a relaxation of the optimal transport problem (3.2) which we present now. Let $N \in \mathbb{N}^*$ and let us consider a set $(\phi_n)_{1\leq n\leq N} \subset L^1(\mathbb{R}^d, \mu; \mathbb{R})$ of N continuous real-valued functions, integrable with respect to μ and called hereafter *test functions*. For all $1 \leq n \leq N$, let us denote by

$$\mu_n = \int_{\mathbb{R}^d} \phi_n(x) \mathrm{d}\mu(x), \tag{3.3}$$

the moments of μ, by

$$\begin{aligned} \Pi(\mu; (\phi_n)_{1\leq n\leq N}; M) := \Bigg\{ & \pi \in \mathcal{P}\left((\mathbb{R}^d)^M\right) : \\ & \forall 1 \leq n \leq N, \int_{(\mathbb{R}^d)^M} \sum_{m=1}^{M} |\phi_n(x_m)| \mathrm{d}\pi(x_1, \dots, x_M) < \infty, \\ & \int_{(\mathbb{R}^d)^M} \left(\frac{1}{M} \sum_{m=1}^{M} \phi_n(x_m) \right) \mathrm{d}\pi(x_1, \dots, x_M) = \mu_n \Bigg\}, \end{aligned} \tag{3.4}$$

the set of probability measures on $(\mathbb{R}^d)^M$ for which the mean of the moments against the test functions of the marginal laws are equal to the one of μ, and by

$$\begin{aligned} \Pi^S(\mu; (\phi_n)_{1\leq n\leq N}; M) := \Bigg\{ & \pi \in \mathcal{P}\left((\mathbb{R}^d)^M\right) : \\ & \forall 1 \leq n \leq N, \int_{(\mathbb{R}^d)^M} \sum_{m=1}^{M} |\phi_n(x_m)| \mathrm{d}\pi(x_1, \dots, x_M) < \infty, \\ & \forall 1 \leq m \leq M, \int_{(\mathbb{R}^d)^M} \phi_n(x_m) \mathrm{d}\pi(x_1, \dots, x_M) = \mu_n \Bigg\} \end{aligned} \tag{3.5}$$

the set of probability measures on $(\mathbb{R}^d)^M$ that have, for each marginal law, the same moments as μ against the test functions.

For technical reasons linked to the fact that the optimal problem is defined on the unbounded state space $\mathbb{R}^d$, we assume in addition that there exists a non-decreasing

non-negative continuous function $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\theta(r) \xrightarrow[r\to+\infty]{} +\infty$ and for which there exists $C > 0$ and $0 < s < 1$ such that for all $1 \leq n \leq N$ and all $x \in \mathbb{R}^d$,

$$|\phi_n(x)| \leq C(1 + \theta(|x|))^s. \tag{3.6}$$

We finally choose a positive real number $A > 0$ satisfying $A \geq A_0 := \int_{\mathbb{R}^d} \theta(|x|) d\mu(x)$.

Then, the *MCOT problem* is defined by

$$I^N := \inf_{\substack{\pi \in \Pi^S(\mu;(\phi_n)_{1\leq n\leq N};M) \\ \frac{1}{M}\int_{(\mathbb{R}^d)^M} \sum_{m=1}^M \theta(|x_m|) d\pi(x_1,\dots,x_M) \leq A}} \int_{(\mathbb{R}^d)^M} c(x_1,\dots,x_M) d\pi(x_1,\dots,x_M). \tag{MCOT^S}$$

Under appropriate assumptions on the family of test functions $(\phi_n)_{1\leq n\leq N}$, it is proved in Chapter 2 that the value of I^N can be made arbitrarily close to I as N, the number of test functions, goes to infinity. Besides, converging subsequences of minimizers to (MCOT^S) necessarily converge to some minimizer of (3.2). This is the reason why (MCOT^S) can be seen as a particular discretization approach for the numerical approximation of Problem (3.2).

Remark 3.1. *It is proved in Chapter 2 that the value of I^N does not depend on the value of A provided that A satisfies $A \geq A_0$.*

Using the symmetry of the cost c and the marginal constraints, it can be easily checked that I^N is also equal to

$$I^N = \inf_{\substack{\pi \in \Pi(\mu;(\phi_n)_{1\leq n\leq N};M) \\ \frac{1}{M}\int_{(\mathbb{R}^d)^M} \sum_{m=1}^M \theta(|x_m|) d\pi(x_1,\dots,x_M) \leq A}} \int_{(\mathbb{R}^d)^M} c(x_1,\dots,x_M) d\pi(x_1,\dots,x_M). \tag{MCOT}$$

Then, from Proposition 2.5 there exists at least one minimizer to problem (MCOT), which can be written as

$$\pi^N = \sum_{k=1}^K w_k \delta_{\left(x_1^k,\dots,x_M^k\right)}, \tag{3.7}$$

for some $0 < K \leq N + 2$, with $w_k \geq 0$ and $x_m^k \in \mathbb{R}^d$ for all $1 \leq m \leq M$ and $1 \leq k \leq K$. Besides, the symmetrized measure associated to π^N, which is defined by

$$\pi_S^N := \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} \sum_{k=1}^K w_k \delta_{\left(x_{\sigma(1)}^k,\dots,x_{\sigma(M)}^k\right)} \tag{3.8}$$

where $\mathcal{S}_M$ is the set of permutations of $\{1, \cdots, M\}$, is a minimizer to (MCOT^S).

The proof of this result makes use of Tchakaloff's theorem [24, Corollary 2], which is recalled in Theorem 3.4 in Section 3.6.1. Note that since $\Pi(\mu;(\phi_n)_{1\leq n\leq N};M) \subset \Pi(\mu;M)$, when I is finite, it naturally holds that $I^N \leq I < \infty$.

These theoretical results naturally lead us to consider an optimization problem similar to (MCOT) but where the optimization set is reduced to the set of measures of $\Pi(\mu;(\phi_n)_{1\leq n\leq N};M)$ which can be written as discrete measures under the form (3.7) for some $K \in \mathbb{N}^*$. This naturally leads to the following optimization problem, which we call hereafter the *MCOT particle problem* with K particles:

$$I_K^N := \inf_{(W,Y)\in\mathcal{U}_K^N} \quad \sum_{k=1}^K w_k c\left(X^k\right), \tag{MCOT^K}$$

where

$$\mathcal{U}_K^N := \Bigg\{ (W,Y) \in \mathbb{R}_+^K \times \left((\mathbb{R}^d)^M\right)^K, \quad W=(w_k)_{1\le k\le K},\ Y=(X^k)_{1\le k\le K}, \tag{3.9}$$
$$\sum_{k=1}^K w_k = 1, \quad \sum_{k=1}^K w_k \vartheta(X^k) \le A, \quad \forall 1\le n\le N,\ \sum_{k=1}^K w_k \varphi_n(X^k) = \mu_n \Bigg\},$$

with, for all $X=(x_1,\cdots,x_M)\in(\mathbb{R}^d)^M$ and all $1\le n\le N$,

$$\vartheta(X) := \frac{1}{M}\sum_{m=1}^M \theta\left(|x_m|\right)) \quad \text{and} \quad \varphi_n(X) := \frac{1}{M}\sum_{m=1}^M \phi_n(x_m). \tag{3.10}$$

In view of Proposition 2.5 we have $I_K^N = I^N$ as soon as $K \ge N+2$.

A few remarks are in order at this point.

Remark 3.2. *(i) Considering problem MCOT^K as a starting point for a numerical scheme seems very appealing, especially in contexts when M is large. Indeed, in principle, the resolution of (MCOT^K) only requires the optimization of at most $K(M+1)$ scalars, thus would require the resolution of an optimization problem defined on a continuous optimization set involving a number of parameters which only scales **linearly** with respect to the number of marginal laws. Thus, gradient-based algorithms are natural to consider for the numerical resolution of (MCOT^K), at least for differentiable test functions.*

(ii) Problem MCOT^K is highly non-convex, whereas the original MCOT problem (MCOT) reads as a (high-dimensional) linear problem[iii]*. This definitely makes the numerical resolution of (MCOT^K) a challenging task. This is the reason why we consider in this article **randomized versions of gradient-based** algorithms for the resolution of (MCOT^K). Nevertheless, strikingly, we prove in this article that, despite the lack of convexity, any local minimizers to the MCOT particle problem (MCOT) are actually **global** minimizers, provided that $K \ge 2N+6$. This is the object of Section 3.2.2 to state this result and further mathematical properties of the set of minimizers to (MCOT^K).*

The main focus of this article is to propose numerical schemes relying on stochastic versions of gradient-based algorithms in order to find minimizers to the MCOT particle problem. Such numerical schemes actually make use of *constrained overdamped Langevin processes*, which are usually encountered in the context of molecular dynamics simulations [218, 219]. In Section 3.3, we relate such stochastic processes with MCOT problems and entropic regularizations of the latter.

In numerical tests, and especially in the 3D case, the schemes proposed in this article perform better when using a large number of particles K, with weights w_k assumed to be fixed and equal to $\frac{1}{K}$ which are not optimized upon. That is why we introduce here the resulting optimization, called the *MCOT fixed-weight particle problem* with K particles, which reads as follows:

$$J_K^N := \inf_{\substack{Y:=\left(X^k\right)_{1\le k\le K}\in((\mathbb{R}^d)^M)^K,\\ \forall 1\le n\le N,\ \frac{1}{K}\sum_{k=1}^K \varphi_n(X^k)=\mu_n,\\ \frac{1}{K}\sum_{k=1}^K \vartheta(X^k)\le A}} \sum_{k=1}^K \frac{1}{K} c\left(X^k\right). \tag{MCOT^K -fixed weight}$$

[iii] More generally, any non-linear minimization problem can be reframed as a linear minimization problem in a much larger space (the measure space), as $\min_{x\in\mathbb{R}^d} c(x) = \min_{\mathbb{P}} \int_{\mathbb{R}^d} c(y)\mathrm{d}\mathbb{P}(y)$.

Remark 3.3. *(i) Let us stress on the fact that the existence of a solution to (MCOT^K -fixed weight) is not guaranteed in general. This stems from the fact that there may not exist a set of points $Y = \left(X^k\right)_{1\leq k\leq K}$ satisfying the constraints of problem (MCOT^K -fixed weight). However, for all $N, K \in \mathbb{N}^*$, it always holds that $J_K^N \geq I_K^N$.*

Let however consider $(W, Y) \in \mathcal{U}_{N+2}^N$ a minimizer of (MCOT^K) *and assume that the cost c and the test functions ϕ_n are bounded. Then, by rounding the weights w_k to a multiple of $1/K$, and then by using ℓ copies of particles with weight ℓ/K, we can construct $\tilde{Y} = \left(\tilde{X}^k\right)_{1\leq k\leq K}$ such that*

$$\frac{1}{K}\sum_{k=1}^{K}\varphi_n(\tilde{X}^k) \approx \mu_n + \mathcal{O}\left(\frac{1}{K}\right).$$

Thus, $\tilde{Y}$ satisfies the moment constraints of problem (MCOT^K -fixed weight) up to an error of order $\mathcal{O}\left(\frac{1}{K}\right)$ and achieves a cost that is also $\mathcal{O}\left(\frac{1}{K}\right)$ away from the optimal cost achieved by (W, Y).

Furthermore, in the limit $K \to \infty$ optima of problems (MCOT^K -fixed weight) *(with an accepted error $\mathcal{O}\left(\frac{1}{K}\right)$ on the constraints) converge to the optimum of the problem* (MCOT).

(ii) Yet, in the numerical experiments in the fixed weight case in 3D, the convergence in K appears to be faster than $\mathcal{O}\left(\frac{1}{K}\right)$ and even low values of K can give sharp approximations of the optimum of (MCOT).

3.2.2 Properties of the set of minimizers of the particle problem

The aim of this section is to present the first main theoretical result of this paper, which states some mathematical properties on the set of minimizers of the particle problem MCOT^K.

For any $(W, Y) \in \mathbb{R}_+^K \times ((\mathbb{R}^d)^M)^K$, we define by

$$\mathcal{I}(W, Y) := \sum_{k=1}^{K} w_k c(X^k),$$

where $W := (w_k)_{1\leq k\leq K}$ and $Y := (X^k)_{1\leq k\leq K}$. Problem (MCOT^K) can then be equivalently rewritten as

$$I_K^N = \inf_{(W,Y)\in\mathcal{U}_K^N} \mathcal{I}(W, Y). \tag{3.11}$$

We begin this section by Theorem 3.1, which states that for any two elements of $\mathcal{U}_K^N$, there exists a continuous path with values in $\mathcal{U}_K^N$ which connects these two elements, and such that $\mathcal{I}$ monotonically varies along this path.

Theorem 3.1. *Let us assume that $K \geq 2N + 6$. Let $(W_0, Y_0), (W_1, Y_1) \in \mathcal{U}_K^N$. Then, there exists a continuous application $\psi : [0, 1] \to \mathcal{U}_K^N$ made of a polygonal chain such that $\psi(0) = (W_0, Y_0)$, $\psi(1) = (W_1, Y_1)$ and such that the application $[0, 1] \ni t \mapsto \mathcal{I}(\psi(t))$ is monotone.*

In order to explain the main ideas of the proof of Theorem 3.1, let us remark that, using Tchakaloff's theorem (recalled in Section 3.6.1), for any measure $\pi \in \Pi(\mu; (\phi_n)_{1\leq n\leq N}; M)$, satisfying,

$$\int_{(\mathbb{R}^d)^M} \vartheta d\pi \leq A, \tag{3.12}$$

and charging $K \geq 2N + 6$ points, one can find a measure $\tilde{\pi} \in \Pi(\mu; (\phi_n)_{1\leq n\leq N}; M)$ charging $N + 3$ points, whose support is included in the one of π, and having the same cost and the same moment against ϑ. Then, the segment $((1-t)\pi + t\tilde{\pi})_{t\in[0,1]}$ is included in $\Pi(\mu; (\phi_n)_{1\leq n\leq N}; M)$, charges at most $2N+6$ points and keeps the cost and the moment against ϑ constant. Besides, let $\tilde{\pi}_0, \tilde{\pi}_1 \in \Pi(\mu; (\phi_n)_{1\leq n\leq N}; M)$ be two measures with support on at most $N+3$ points, and such that for $i = 0, 1$, $\tilde{\pi}_i$ satisfies (3.12). Then, the segment $((1-t)\tilde{\pi}_0 + t\tilde{\pi}_1)_{t\in[0,1]}$ is included in $\Pi(\mu; (\phi_n)_{1\leq n\leq N}; M)$, satisfies the inequality constraint (3.12) for all $t \in [0, 1]$, charges at most $2N + 6$ points, and the cost varies linearly along it. By identifying (W_0, Y_0) with π_0 (resp. (W_1, Y_1) with π_1), one can join π_0 to π_1 by segments (with appropriately defined intermediate measures $\tilde{\pi}_0$ and $\tilde{\pi}_1$) satisfying the constraints, and along which the cost varies linearly. The adaptation of these ideas to vectors $(W_0, Y_0), (W_1, Y_1) \in \mathcal{U}_K^N$, which requires to take into account the displacement of the positions between Y_0 and Y_1 as well as the ordering of the coordinates, is the object of Section 3.6.2.

A direct consequence of Theorem 3.1 is then Corollary 3.2 which states that any local minimizer to problem (3.11) (or equivalently problem MCOT^K) is actually a *global minimizer* as soon as $K \geq 2N + 6$. In addition, the set of minimizers forms an polygonally connected (and thus arc-connected) set.

Corollary 3.2. *Let us assume that $K \geq 2N + 6$. Then, any local minimizer of the MCOT particle problem* (MCOT^K) *is actually a global minimizer. Besides, the set of (local or global) minimizers of the MCOT particle problem* (MCOT^K) *is an polygonally connected subset of $\mathbb{R}_+^K \times ((\mathbb{R}^d)^M)^K$.*

3.3 Overdamped Langevin processes for MCOT particle problems

The motivation of this section is twofold: first, the numerical method used in this article for the resolution of the particle problems (MCOT^K) and (MCOT^K -fixed weight) can be seen as a time discretization of constrained overdamped Langevin dynamics, which are usually encountered in molecular dynamics simulation; second, we draw here a link, on the formal level, between the long-time and large number of particles limit of these processes and a regularized version of the MCOT problem (MCOT) using the so-called Kullback-Leibler entropy regularization, very similar to the regularization which is at the core of the Sinkhorn algorithm for the resolution of optimal transportation problem [277].

The objective of Section 3.3.1 is to recall some fundamental properties of general constrained overdamped Langevin processes. Then, in Section 3.3.2, we consider specific processes which are related to the MCOT problem presented in Section 3.2.

3.3.1 Properties of general constrained overdamped Langevin processes

3.3.1.1 Definition

Let $p \in \mathbb{N}^*$. Let us first introduce *unconstrained* overdamped Langevin processes in the state space $\mathbb{R}^p$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An overdamped Langevin stochastic process is a stochastic process $(Y_t)_{t\geq 0}$ solution to the following stochastic differential equation

$$dY_t = -\nabla V(Y_t)\, dt + \beta\, dW_t,$$

where $V : \mathbb{R}^p \to \mathbb{R}$ is a smooth function, called hereafter the *potential function* of the overdamped Langevin process, $\beta > 0$ is a positive coefficient which is proportional to the square root of the temperature of the system in molecular dynamics ($\beta = \sqrt{2\bar{T}}$ with $\bar{T}$ the temperature), and $(W_t)_{t\geq 0}$ is a p-dimensional Brownian motion.

Constrained overdamped Langevin processes are overdamped Langevin processes whose trajectory is enforced to be included into a given submanifold. In the sequel, we assume that the submanifold is characterized as the zero isovalued set of a given smooth function $\Gamma : \mathbb{R}^p \to \mathbb{R}^q$ for some $q \in \mathbb{N}^*$, so that the corresponding submanifold is defined by

$$\mathcal{M} = \{Y \in \mathbb{R}^p, \Gamma(Y) = 0\}.$$

We assume in the sequel that the submanifold $\mathcal{M}$ is arc connected. In addition, let us assume that there exists a neighborhood $\mathcal{W}$ of $\mathcal{M}$ such that, for all $Y \in \mathcal{W}$,

$$G(Y) := \nabla\Gamma(Y)^T \nabla\Gamma(Y) \in \mathbb{R}^{q\times q} \tag{3.13}$$

is an invertible matrix, where $\nabla\Gamma(Y)_{i,j} = \partial_i \Gamma_j$ for $1 \leq i \leq p$ and $1 \leq j \leq q$. These two assumptions on the function Γ, together with the implicit function theorem, imply that $\mathcal{M}$ is a regular $(p-q)$-dimensional submanifold.

A constrained overdamped Langevin process [218, Section 3.2.3] is a $\mathbb{R}^p$-valued stochastic process $(Y_t)_{t\geq 0}$ that solves the stochastic differential equation

$$\begin{cases} dY_t = -\nabla V(Y_t)\, dt + \beta dW_t + \nabla\Gamma(Y_t) d\Lambda_t, \\ \Gamma(Y_t) = 0, \end{cases} \tag{3.14}$$

where $\beta > 0$, $(W_t)_{t\geq 0}$ is a p-dimensional Brownian process and $(\Lambda_t)_{t\geq 0}$ is a q-dimensional stochastic adapted stochastic process, which ensures that Y_t belongs to the submanifold $\mathcal{M}$ almost surely for all $t \in \mathbb{R}^+$. More precisely, Λ_t is the Lagrange multiplier associated to the constraint $\Gamma(Y_t) = 0$ and is defined by

$$d\Lambda_t = G^{-1}(Y_t) \left[\left(\nabla\Gamma(Y_t)^T \nabla V(Y_t) - \frac{\beta^2}{2} \begin{pmatrix} \sum_{i=1}^p \partial_i^2 \Gamma_1(Y_t) \\ \vdots \\ \sum_{i=1}^p \partial_i^2 \Gamma_q(Y_t) \end{pmatrix} \right) dt - \beta \nabla\Gamma(Y_t)^T dW_t \right]. \tag{3.15}$$

Thus, if we define $P(y) = \mathrm{Id} - \nabla\Gamma(y)^T G^{-1}(u) \nabla\Gamma(y)$ the projection operator, we get

$$dY_t = P(Y_t)[-\nabla V(Y_t) + \beta dW_t] - \frac{\beta^2}{2} \nabla\Gamma(Y_t)^T G^{-1}(Y_t) \begin{pmatrix} \sum_{i=1}^p \partial_i^2 \Gamma_1(Y_t) \\ \vdots \\ \sum_{i=1}^p \partial_i^2 \Gamma_q(Y_t) \end{pmatrix} dt.$$

Let us assume in addition that

$$Z := \int_{\mathbb{R}^p} e^{-\frac{2V(Y)}{\beta^2}}\, d\sigma_{\mathcal{M}}(Y) < +\infty, \tag{3.16}$$

where $d\sigma_{\mathcal{M}}$ is the surface measure (induced by the Lebesgue measure in $\mathbb{R}^p$, see [218, Remark 3.4] for a precise definition) on the submanifold $\mathcal{M}$. Let us introduce the probability measure $\eta \in \mathcal{P}(\mathbb{R}^p)$ defined by

$$d\eta(Y) := \frac{1}{Z} e^{-\frac{2V(Y)}{\beta^2}} |\det G(Y)|^{-1/2} d\sigma_{\mathcal{M}}(Y). \tag{3.17}$$

Under suitable assumptions, [218, Proposition 3.20] states that η is the unique equilibrium distribution of the stochastic process Y_t solution to the constrained overdamped Langevin dynamics (3.14) and that

$$Y_t \text{ weakly converges to } \eta \text{ as } t \to +\infty. \tag{3.18}$$

3.3.1.2 Long-time and large number of particles limit

We recall here some results proved in [290, Section 2.3 and Proposition 5.1], where the authors consider the so-called large-particle limit of constrained overdamped Langevin dynamics subject to average moment constraints. The objective of the work [290] was to study the properties of the constrained overdamped Langevin process in a large number of particles limit and to show the convergence towards η of the invariant distribution of the approximating particle system when the number of particles $K \to \infty$.[iv]More precisely, from now on, let us consider $p' = Kp$ for some $K \in \mathbb{N}^*$. We define for any $K \in \mathbb{N}^*$ the potential function V^K and the constraint function Γ^K by:

$$\forall Y = (X^k)_{1\leq k\leq K} \in (\mathbb{R}^p)^K,\ V^K(Y) := \frac{1}{K}\sum_{k=1}^K V\left(X^k\right) \text{ and } \Gamma^K(Y) := \frac{1}{K}\sum_{k=1}^K \Gamma\left(X^k\right).$$

We then consider the following constrained overdamped Langevin process $(Y_t^K)_{t\geq 0}$ that is assumed to be solution to the stochastic differential equation

$$\begin{cases} dY_t^K = -\nabla V^K(Y_t^K)\, dt + \beta dW_t^K + \nabla\Gamma^K(Y_t^K) d\Lambda_t^K, \\ \Gamma^K(Y_t^K) = 0, \end{cases} \tag{3.19}$$

where $(W_t^K)_{t\geq 0}$ is a Kp-dimensional Brownian process and $(\Lambda_t^K)_{t\geq 0}$ is a q-dimensional stochastic adapted stochastic process, which ensures that Y_t^K satisfies the constraint $\Gamma^K(Y_t^K) = 0$ almost surely. The process Y^K is usually called a particle system: each coordinate X^k for $1 \leq k \leq K$ is seen as a particle. The large number of particles limit consists in considering the limit as K goes to infinity of the stochastic process $(Y_t^K)_{t\geq 0}$.

It follows from (3.18) that, under suitable assumptions, as t goes to ∞, the law of the process Y_t^K converges to the probability measure $\eta^K \in \mathcal{P}\left((\mathbb{R}^p)^K\right)$ defined for all $Y^K = (X^1, \cdots, X^K) \in (\mathbb{R}^p)^K$ by

$$d\eta^K(Y^K) = \frac{1}{Z^K}\left(\Pi_{k=1}^K e^{-\frac{2V(X^k)}{\beta^2}}\right) d\sigma_{\mathcal{M}^K}(Y^K), \tag{3.20}$$

where

$$\mathcal{M}^K := \left\{Y^K \in (\mathbb{R}^p)^K,\ \Gamma^K(Y^K) = 0\right\},$$

[iv]We use here the notation K for the number of particles in view of the use of the Langevin dynamics to solve (MCOT^K) problems, from Section 3.3.2, for which $K \geq 2N+6$. Yet, the results recalled in Section 3.3.1.2 are general and unrelated to MCOT applications.

$$Z^K := \int_{(\mathbb{R}^p)^K} e^{-\frac{2V^K(Y^K)}{\beta^2}} d\sigma_{\mathcal{M}^K}(Y^K),$$

and

$$G^K(Y^K) := \nabla\Gamma^K(Y^K)^T\nabla\Gamma^K(Y^K) \in \mathbb{R}^{q\times q}.$$

For $1 \leq k \leq K$, $\left(X_t^k\right)_{t\geq 0}$ is a p-dimensional stochastic process. Let us denote by $\zeta_t^K \in \mathcal{P}(\mathbb{R}^p)$ the law of the first particle X_t^1. Then, the symmetry of the functions V^K and Γ^K implies that ζ_t^K weakly converges in law when $t \to \infty$ to the probability measure ζ_∞^K defined for all $X \in \mathbb{R}^p$ by

$$d\zeta_\infty^K(X) = \int_{(\mathbb{R}^p)^{K-1}} d\eta^K(X, X^2, \cdots, X^K). \tag{3.21}$$

Under appropriate assumptions on V and Γ which we do not detail here [290, Proposition 5.1], the sequence $\left(\zeta_\infty^K\right)_{K\in\mathbb{N}^*}$ weakly converges in $\mathcal{P}(\mathbb{R}^p)$ as K goes to infinity to a probability measure $\pi_\beta^* \in \mathcal{P}(\mathbb{R}^p)$ which is the unique solution to

$$\pi_\beta^* := \underset{\substack{\pi\in\mathcal{P}(\mathbb{R}^p)\\ \int_{\mathbb{R}^p}\Gamma\, d\pi=0}}{\arg\min} \int_{\mathbb{R}^p} \ln\left(\frac{d\pi(X)}{(Z^\infty)^{-1}e^{-\frac{2V(X)}{\beta^2}}\, dX}\right) d\pi(X), \tag{3.22}$$

where $Z^\infty := \int_{\mathbb{R}^p} e^{-\frac{2v(X)}{\beta^2}}\, dX$. In other words, π_β^* is thus a probability measure on $\mathbb{R}^p$, which is absolutely continuous with respect to the Lebesgue measure and which is solution to

$$\pi_\beta^* := \underset{\substack{\pi\in\mathcal{P}(\mathbb{R}^p)\\ \int_{\mathbb{R}^p}\Gamma\, d\pi=0}}{\arg\min} \int_{\mathbb{R}^p} V(X)\, d\pi(X) + \frac{\beta^2}{2}\int_{\mathbb{R}^p} \ln\left(\frac{d\pi(X)}{dX}\right) d\pi(X). \tag{3.23}$$

3.3.2 Application to MCOT problems

The aim of this section is to illustrate the link between the MCOT problems presented in Section 3.2 and the constrained overdamped Langevin processes introduced in Section 3.3.1. We start by considering the *fixed weight MCOT particle problem* (MCOT^K -fixed weight), before considering the MCOT particle problem with adaptive weights (MCOT).

3.3.2.1 Fixed-weight MCOT particle problem

We first draw the link between constrained Langevin overdamped dynamics and the fixed weight MCOT particle problem (MCOT^K -fixed weight). Then, for all $K \in \mathbb{N}^*$, let us consider $(Y_t^K)_{t\geq 0}$ a constrained overdamped Langevin process solution to the stochastic differential equation (3.19) with $p = dM$, $q = N$, $V = c$ and $\Gamma = (\varphi_1 - \mu_1, \cdots, \varphi_N - \mu_N)$ where for all $1 \leq n \leq N$, φ_n is defined by (3.10).

Then, the stochastic dynamics (3.19) can be viewed as a randomized version of a constrained gradient numerical method for the resolution of problem (MCOT^K -fixed weight), where for all $t \geq 0$, $Y_t^K = (X_t^1, \cdots, X_t^K) \in ((\mathbb{R}^d)^M)^K$ and where for all $1 \leq k \leq K$,

$$X_t^k = \left(x_{1,t}^k, \cdots, x_{M,t}^k\right) \in \left(\mathbb{R}^d\right)^M.$$

Note that it is not clear in general that V and Γ satisfy the regularity assumptions which ensure the convergence results stated in Section 3.3.1.2 to hold true. But, formally, assuming that the long-time limit and large number of particles convergence

holds nevertheless, the associated measure π^*_β solution (3.23) can be equivalently rewritten as

$$\pi^*_\beta := \underset{\substack{\pi\in\mathcal{P}\left((\mathbb{R}^d)^M\right)\\ \forall 1\le n\le N,\, \int_{(\mathbb{R}^d)^M}\varphi_n\, d\pi=\mu_n}}{\arg\min} \mathcal{J}(\pi), \tag{3.24}$$

where

$$\mathcal{J}(\pi) := \int_{(\mathbb{R}^d)^M} c\,(X)\; d\pi(X) + \frac{\beta^2}{2}\int_{(\mathbb{R}^d)^M} \ln\left(\frac{d\pi(X)}{dX}\right) d\pi(X).$$

Recall that π^*_β is the large number of particles limit of the long-time limit of the law of one particle associated to the constrained overdamped Langevin process. Notice that π^*_β can be equivalently seen as the solution of an entropic regularization of the MCOT problem (MCOT), where the term $\int_{(\mathbb{R}^d)^M} \ln\left(\frac{d\pi(X)}{dX}\right) d\pi(X)$ can be identified as the Kullback-Leibler entropy of the measure π with respect to the Lebesgue measure. Thus, Problem 3.24 is close to the entropic regularization of optimal transport problems used in several works [41, 118, 262, 277], in particular for the so-called Sinkhorn algorithm [41].

Let us point out here that, at least on the formal level, we expect the family $(\pi^*_\beta)_{\beta>0}$ to weakly converge to a minimizer of (MCOT) as β goes to 0 (a similar result is proven in [83, Theorem 2.7]).

3.3.2.2 Adaptive-weight MCOT particle problem

A similar link can be drawn between constrained Langevin overdamped dynamics and the MCOT particle problem (MCOT^K) with adaptive weights.

In order to fit in the framework of the constrained Langevin overdamped dynamics, without any positivity constraint, let us introduce a continuous surjective function $f : \mathbb{R} \to \mathbb{R}_+$, which we call hereafter a *weight function*. We assume that f satisfies the following assumption: there exists an interval $I \subset \mathbb{R}$ such that the Lebesgue measure of I is equal to 1 and such that $\int_I f = 1$. A simple choice of admissible weight function can be given by $f(a) = a^2$ for all $a \in \mathbb{R}$ with $I = (\frac{1}{2}, \frac{25^{1/3}}{2})$.

Then, for all $K \in \mathbb{N}^*$, let us consider $\left(\overline{Y}^K_t\right)_{t\ge 0}$ a constrained overdamped Langevin process solution to the stochastic differential equation (3.19) with $p = dM + 1$, $q = N + 1$, and where for all $\overline{X} = (a, X) \in \mathbb{R} \times (\mathbb{R}^d)^M$, $\overline{V}(\overline{X}) = f(a)c(X)$ and $\overline{\Gamma}(\overline{X}) = (f(a) - 1, f(a)\varphi_1(X) - \mu_1, \cdots, f(a)\varphi_N(X) - \mu_N)$. Then, the stochastic dynamics (3.19) can be viewed as a randomized version of a constrained gradient numerical method for the resolution of the optimization problem

$$\inf_{(A,Y)\in\mathcal{V}^N_K} \quad \sum_{k=1}^K f(a_k)c\left(X^k\right), \tag{3.25}$$

where

$$\mathcal{V}^N_K := \Bigg\{ (A, Y) \in \mathbb{R}^K \times \left((\mathbb{R}^d)^M\right)^K, \quad A = (a_k)_{1\le k\le K},\; Y = (X^k)_{1\le k\le K}, \tag{3.26}$$
$$\sum_{k=1}^K f(a_k) = 1, \quad \sum_{k=1}^K f(a_k)\vartheta(X^k) \le A, \quad \forall 1 \le n \le N,\; \sum_{k=1}^K f(a_k)\varphi_n(X^k) = \mu_n \Bigg\},$$

which is equivalent to problem (MCOT^K) using the surjectivity of f.

Note that the choice of the function f can influence the dynamics as it regulates both the way the brownian motion W affects the weights, and the balance, in the minimization of $\overline{V}$ and in the enforcement of the constraint $\overline{\Gamma}(\overline{X}) = 0_N$, between a displacement of particles and a change in weights.

Here again, it is not clear in general that $\overline{V}$ and $\overline{\Gamma}$ satisfies the regularity assumptions which ensures the convergence results stated in Section 3.3.1.2 to hold true. But, using formal computations, we can consider the associated measure $\overline{\pi}^{\mathrm{a}}_{\beta} \in \mathcal{P}\left(\mathbb{R} \times (\mathbb{R}^d)^M\right)$ solution to

$$\overline{\pi}^{\mathrm{a}}_{\beta} := \underset{\substack{\overline{\pi} \in \mathcal{P}\left(\mathbb{R}\times(\mathbb{R}^d)^M\right) \\ \int_{a\in\mathbb{R}} \int_{X\in(\mathbb{R}^d)^M} f(a)\, d\overline{\pi}(a,X)=1 \\ \forall 1\leq n\leq N,\, \int_{a\in\mathbb{R}} \int_{X\in(\mathbb{R}^d)^M} f(a)\varphi_n(X)\, d\overline{\pi}(a,X)=\mu_n}}{\arg\min} \overline{\mathcal{J}}(\overline{\pi}), \tag{3.27}$$

where

$$\overline{\mathcal{J}}(\overline{\pi}) := \int_{a\in\mathbb{R}} \int_{X\in(\mathbb{R}^d)^M} f(a)c\left(X\right)\, d\overline{\pi}(a,X) + \frac{\beta^2}{2} \int_{a\in\mathbb{R}} \int_{X\in(\mathbb{R}^d)^M} \ln\left(\frac{d\overline{\pi}(a,X)}{dadX}\right) d\overline{\pi}(a,X).$$

Let us introduce now $\pi^{\mathrm{a}}_{\beta} \in \mathcal{P}\left((\mathbb{R}^d)^M\right)$ defined by

$$d\pi^{\mathrm{a}}_{\beta}(X) = \int_{a\in\mathbb{R}} f(a)\, d\overline{\pi}^{\mathrm{a}}_{\beta}(a,X).$$

Then π^{a}_{β} satisfies the constraints of problem (3.24) and

$$\overline{\mathcal{J}}(\overline{\pi}^{\mathrm{a}}_{\beta}) = \int_{X\in(\mathbb{R}^d)^M} c(X)\, d\pi^{\mathrm{a}}_{\beta}(X) + \frac{\beta^2}{2} \int_{a\in\mathbb{R}} \int_{X\in(\mathbb{R}^d)^M} \ln\left(\frac{d\overline{\pi}^{\mathrm{a}}_{\beta}(a,X)}{dadX}\right) d\overline{\pi}^{\mathrm{a}}_{\beta}(a,X).$$

Notice that, as a consequence, problem (3.27) may be seen as a second kind of entropic regularization of (MCOT) and that π^{a}_{β} is expected to be an approximation of some minimizer to (MCOT) as β goes to 0.

Let us notice here that the assumption made on f ensures that, for all $\pi \in \mathcal{P}\left((\mathbb{R}^d)^M\right)$, there exists a probability measure $\overline{\pi} \in \mathcal{P}\left(\mathbb{R} \times (\mathbb{R}^d)^M\right)$ such that

$$d\pi(X) = \int_{a\in\mathbb{R}} f(a)\, d\overline{\pi}(a,X).$$

Indeed, defining $d\overline{\pi}(a,X) := \mathbb{1}_I(a)\, da \otimes d\pi(X)$ yields the desired result. Besides, we easily check that $\overline{\mathcal{J}}(\overline{\pi}) = \mathcal{J}(\pi)$, which leads immediately to $\overline{\mathcal{J}}(\overline{\pi}^{\mathrm{a}}_{\beta}) \leq \mathcal{J}(\pi^*_{\beta})$ from the optimality of $\overline{\pi}^{\mathrm{a}}_{\beta}$.

3.4 Numerical optimization method

We present in this section the numerical procedure we use in our numerical tests to compute approximate solutions to the particle problems with fixed weights (MCOT^K -fixed weight) or adaptive weights (MCOT^K) for a fixed given $K \in \mathbb{N}^*$. Note that (MCOT^K -fixed weight) can be equivalently rewritten as

$$J^N_K := \underset{\substack{Y^K\in((\mathbb{R}^d)^M)^K, \\ \Gamma^K(Y^K)=0, \\ \Theta^K(Y^K)\leq A}}{\inf} V^K(Y^K), \tag{3.28}$$

where V^K and Γ^K are defined in Section 3.3.2.1, and where

$$\Theta^K : \left\{ \begin{array}{ccc} ((\mathbb{R}^d)^M)^K & \mapsto & \mathbb{R} \\ Y := (X^1, \cdots, X^K) & \to & \frac{1}{K}\sum_{k=1}^K \vartheta(X^k). \end{array} \right.$$

Besides, problem (MCOT^K) can be rewritten equivalently as

$$I_K^N := \inf_{\substack{\overline{Y}^K \in \left(\mathbb{R}\times((\mathbb{R}^d)^M)\right)^K \\ \overline{\Gamma}^K(\overline{Y}^K)=0, \\ \overline{\Theta}^K(\overline{Y}^K)\leq A}} \overline{V}^K(\overline{Y}^K), \tag{3.29}$$

where $\overline{V}^K$ and $\overline{\Gamma}^K$ are defined in Section 3.3.2.2, and where

$$\overline{\Theta}^K : \left\{ \begin{array}{ccc} (\mathbb{R}\times(\mathbb{R}^d)^M)^K & \mapsto & \mathbb{R} \\ \overline{Y} := ((a^1, X^1), \cdots, (a^K, X^K)) & \to & \frac{1}{K}\sum_{k=1}^K f(a^k)\vartheta(X^k). \end{array} \right.$$

For the sake of simplicity, we restrict the presentation here to the method used for the resolution of (3.28), since the method used for the resolution of (3.29) follows exactly the same lines.

3.4.1 Time-discretization of constrained overdamped Langevin dynamics

The numerical procedure considered in this paper consists in a time discretization of the dynamics (3.14) with an adaptive time step and noise level. The main idea of the algorithm is the following: let $(W_n)_{n\in\mathbb{N}}$ be a sequence of iid normal vectors of dimension dMK. At each iteration $n \in \mathbb{N}^*$ of the procedure, starting from an initial guess $Y_0^K \in \mathcal{M}^K$ for $n = 0$, a new approximation $Y_{n+1}^K \in \mathcal{M}^K$ is computed as the projection in some sense of $Y_{n+1/2}^K := Y_n^K - \nabla V^K(Y_n^K)\Delta t_n + \beta_n\sqrt{\Delta t_n}W_n$ onto $\mathcal{M}^K$, where $\Delta t_n > 0$ is the time step and $\beta_n > 0$ is the noise level at iteration n. Precisely, the next iterate Y_{n+1}^K is computed as $Y_{n+1/2}^K + \nabla\Gamma^K(Y_n^K)\cdot\Lambda_{n+1}^K$ where $\Lambda_{n+1}^K \in \mathbb{R}^N$ is a Lagrange multiplier which ensures that the constraint $\Gamma^K(Y_{n+1}^K) = 0$ is satisfied.

The complete resulting procedure is summarized in Algorithm 2.

We discuss here three main difficulties about the algorithm we propose:

- the initialization step which consists in finding an element $Y_0^K \in \mathcal{M}^K$;
- the choice of the values of the time step Δt_n and noise level β_n at each iteration of the algorithm;
- the practical method used in order to compute a projection of $Y_{n+1/2}^K$ onto the submanifold $\mathcal{M}^K$, and in particular the value of the Lagrange multiplier Λ_{n+1}^K.

The procedure chosen to adapt the time step and noise level is discussed in Section 3.4.2. The algorithm used to compute a projection of $Y_{n+1/2}^K$ onto the submanifold $\mathcal{M}^K$ and the value of the Lagrange multiplier Λ_{n+1}^K is detailed in Section 3.4.3. Finally, the initialization procedure used to compute a starting guess $Y_0^K \in \mathcal{M}^K$ is exaplined in Section 3.4.4.

Algorithm 2 Constrained Overdamped Langevin Algorithm

Input $Y_0^K \in \mathcal{M}^K$, $\Delta t_0 > 0$, $\beta_0 > 0$, $\tau_0 > 0$, $i_{\text{const}} \in \mathbb{N}^*$, $i_{\max} \in \mathbb{N}^*$, NoiseDecrease : $\mathbb{R}^+ \times \mathbb{N} \to \mathbb{R}^+$, $n_{\max} \in \mathbb{N}^*$
Fix $n = 0$, $\Lambda_0^K = 0$.
Define $(W_n)_{n\in\mathbb{N}}$ a sequence of i.i.d. normal vectors of the same dimension as Y_0^K.

while $n \leq n_{\max}$ **do**
  AdaptTimeStep$(Y_n^K, \Lambda_n^K, \Delta t_n, \beta_n, \tau_n)$
  $Y_{n+1/2}^K := Y_n^K - \nabla V^K(Y_n^K)\Delta t_n + \beta_n\sqrt{\Delta t_n}W_n$
  if Projection$(Y_{n+1/2}^K, \nabla\Gamma^K(Y_n^K), \Lambda_n^K, i_{\max})$ succeeds **then**
    $Y_{n+1}^K, \Lambda_{n+1}^K, i_n \leftarrow$ Projection$(Y_{n+1/2}^K, \nabla\Gamma^K(Y_n^K), \Lambda_n^K, i_{\max})$
    if $i_n \leq i_{\text{const}}$ **then**
      $\tau_{n+1} \leftarrow 2\tau_n$
    end if
    $\beta_{n+1} \leftarrow$ NoiseDecrease(β_n, n)
    $\Delta t_{n+1} \leftarrow \Delta t_n$; $\tau_{n+1} \leftarrow \tau_n$
    $n \leftarrow n + 1$
  else
    $\tau_n \leftarrow \tau_n/2$
  end if
end while
return $\min(V^K(Y_n^K), 0 \leq n \leq n_{\max})$

3.4.2 Time step and noise level adaptation procedure

Two remarks are in order to motivate the procedure we propose here:

(i) the computation of the Lagrange multiplier Λ_{n+1}^K at each iteration n of the algorithm and of the resulting value of Y_{n+1}^K must be fast (as it is executed at each step).

(ii) the time-step Δt_n must be:

  (a) small enough for the procedure that computes the Lagrange multiplier to be well-defined,

  (b) large enough for the total number of iterations needed to observe convergence to be reasonable. In practice, $n_{\max}$ was chosen to be of the order of 20000 in the numerical experiments presented in Section 3.5.

To address item (i), we use a Newton method similar to the one proposed in [219, 220] to enforce the constraints and compute the Lagrange multiplier Λ_{n+1}^K which is summarized in Algorithm 4 and detailed in Section 3.4.3. This method is observed to converge fast if the value $Y_{n+1/2}^K$ is close enough to the submanifold $\mathcal{M}^K$. The tolerance threshold allowed at each step on the satisfiability of the constraints is given by $\tau_n > 0$, the value of which is also adapted at each step. Its precise value is inferred as follows: if the Newton method converges fast enough (i.e. if the number of iterations needed to ensure convergence i_n is lower than some fixed value i_{const}), then the value of τ_n is multiplied by 2. On the other hand, if the Newton method does not converge in a maximum number of iterations (given by $i_{\max}$), then τ_n is divided by 2. This step may involve a new time-step computation for iteration n, which we detail below.

The time-step Δt_n is adapted (in order to answer item (ii)) through Algorithm 3 (AdaptTimeStep subprocedure). It is increased at each step n if the constraints are satified up to a tolerance threshold *lower* than τ_n (in order to answer (iib)). Otherwise, the time-step is divided by 2 as many times as needed for $Y^K_{n+1/2}$ to satisfy the constraints defining the submanifold up to a tolerance lower than τ_n (in order to satisfy item (iia)).

Moreover, the noise-level β_n is decreased at each iteration n at a rate inspired from Robbins-Siegmund Lemma [266, Theorem 6.1] for non-constrained stochastic gradient optimization, using the NoiseDecrease function in Algorithm 2. This is managed through the NoiseDecrease function. In the numerical experiments presented in Section 3.5, we used two possible choices of NoiseDecrease function defined respectively by $(\beta, n) \mapsto \beta$ (noise level unchanged) and $(\beta, n) \mapsto \sqrt{\frac{n}{n+1}}\beta$ (slow decrease of the noise level: note that this is the relative decrease, so that after n steps, the noise is $\beta_0/\sqrt{1+n}$).

Algorithm 3 AdaptTimeStep subprocedure

Input: Y^K, Λ, Δt, β, τ, n
if $\|\Gamma^K(Y^K - \nabla V^K(Y^K)2\Delta t + W_n\sqrt{2\Delta t}\beta)\| \leq \tau$ **then**
  $\Delta t \leftarrow 2\Delta t$;
else
  while $\|\Gamma^K(Y^K - \nabla V^K(Y^K)\Delta t + W_n\sqrt{2\Delta t}\beta)\| \geq \tau$ **do**
    $\Delta t \leftarrow \Delta t/2$; $\Lambda \leftarrow \Lambda/2$
  end while
end if

3.4.3 Projection method

As mentioned earlier, to compute $Y^K_{n+1} \in \mathcal{M}^K$ and Λ^K_{n+1} from $Y^K_{n+1/2}$, we use a Newton method similar to the one proposed in [219, 220]. We refer the reader to [220, Section 2.2.2] for theoretical considerations on such projections.

More precisely, the procedure reads as follows: given $Y^K_n, Y^K_{n+1/2} \in ((\mathbb{R}^d)^M)^K$, the aim of the Newton procedure is to find a solution $\Lambda^K_{n+1} \in \mathbb{R}^N$ to the equation

$$\Gamma^K\left(Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda^K_{n+1}\right) = 0.$$

We numerically observe that this Newton procedure only converges in cases when $Y^K_{n+1/2}$ and Y^K_n are close enough to the manifold $\mathcal{M}^K$. Provided that Y^K_n belongs to $\mathcal{M}^K$, $Y^K_{n+1/2}$ can be made arbitrarily close to the submanifold provided that the value of the time step Δt_n is chosen small enough. We also refer the reader to [279, Theorem 1.4.1] for theoretical conditions which guarantee the convergence of this Newton procedure.

This projection procedure, together with the routine for the adaptation of the error tolerance τ_n on the satisfiability of the constraints, is summarized in Algorithm 4. Note that this Newton algorithm requires the inversion of matrices of the form

$$\nabla\Gamma^K(Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda)^T\cdot\nabla\Gamma(Y^K_n)$$

for $\Lambda \in \mathbb{R}^N$ and that we cannot theoretically guarantee the invertibility of this matrix in general. In practice, it naturally depends significantly on the choice of test functions $(\phi_n)_{1\leq n\leq N}$.

Algorithm 4 Projection subprocedure (Newton method)

Input: $Y^K_{n+1/2}$, $\nabla\Gamma^K(Y^K_n)$, Λ^K_n, $i_{\max}$
$i = 0$, $\Lambda'_0 \leftarrow \Lambda_n$
while $\|\Gamma^K(Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda'_i)\| > 10^{-16}$ and $i \leq i_{\max}$ **do**
  $\Lambda'_{i+1} \leftarrow \Lambda'_i - \left(\nabla\Gamma^K(Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda'_i)^T\cdot\nabla\Gamma(Y^K_n)\right)^{-1}\cdot\Gamma^K((Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda'_i)$
  $i \leftarrow i+1$
end while
if $\|\Gamma^K(Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda'_i)\| \leq 10^{-16}$ **then**
  return $Y^K_{n+1/2} + \nabla\Gamma^K(Y^K_n)\cdot\Lambda'_i$, Λ'_i, i
else
  return Projection failure.
end if

3.4.4 Initialization procedure

Algorithm 2 is initialized with an initial guess Y^K_0 which is assumed to belong to the constraints submanifold $\mathcal{M}^K$. In practice, finding an element which belongs to this submanifold is a delicate task, especially when the number of test functions is large. Indeed, as mentioned in the preceding section, the Newton procedure described in Section 3.4.3 only converges if the starting point of the algorithm is sufficiently close to the manifold $\mathcal{M}^K$. This is the reason why this initialization step is rather performed using a method inspired from [324, Section 5 example 3]. A Runge-Kutta 3 (Bogacki-Shampine) numerical scheme [285, (5.8-42)] is used in order to discretize the dynamics

$$\frac{d}{dt}Y^K(t) = F(Y^K(t))$$

starting from a random initial state $Y^K(t=0) = Y^{K,0} \in \left((\mathbb{R}^d)^M\right)^K$, where F is defined as

$$\forall Y^K \in \left((\mathbb{R}^d)^M\right)^K, \quad F(Y^K) = -\left\|\Gamma^K\left(Y^K\right)\right\|_2^2 \frac{\nabla\Gamma^K(Y^K)\cdot\Gamma^K(Y^K)}{\|\nabla\Gamma^K(Y^K)\cdot\Gamma^K(Y^K)\|_2^2}. \tag{3.30}$$

We observe that such a numerical procedure is more robust than a Newton algorithm, even if it can converge very slowly.

Let us mention here that, in the case of the particle problem (3.29) with adaptive weights, an additional step may be used prior to such a Runge-Kutta method, which consists in using a Carathéodory-Tchakaloff subsampling procedure. Carathéodory-Tchakaloff subsampling [278, 313] has been introduced to compute low nodes cardinality cubatures.

In our context, this method can be adapted to find a low nodes cardinality starting point, as close as possible to the constraints submanifold $\overline{\mathcal{M}}^K$. More precisely, the method works as follows: we fix a value $K_\infty \gg K$ and compute $(X^1, \cdots, X^{K_\infty})$ iid samples of random vectors according to the probability law μ. A Non-Negative Least Squares (NNLS) is then used to find a sparse solution to the optimization problem

$$w^* \in \underset{w\in\mathbb{R}_+^{K_\infty}}{\arg\min} \|\Phi w - \bar{\mu}\|^2, \tag{3.31}$$

where $\Phi := (\Phi_{n,k})_{1\leq n\leq N+1, 1\leq k\leq K_\infty} \in \mathbb{R}^{N\times K_\infty}$, $\overline{\mu} = (\mu_1, \cdots, \mu_N, 1) \in \mathbb{R}^{N+1}$ and

$$\forall 1 \leq k \leq K_\infty, \quad \forall 1 \leq n \leq N,\ \Phi_{n,k} = \varphi_n(X^k) \text{ and } \Phi_{N+1,k} = 1.$$

By Kuhn-Tucker conditions for the NNLS problem [214, Theorem (23.4)], there exists a solution $w^* := (w_k^*)_{1\leq k\leq K_\infty} \in \mathbb{R}_+^{K_\infty}$ to (3.31) such that $\#J \leq N+1$ with $J := \{1 \leq k \leq K_\infty, w_k^* > 0\}$. Common algorithms such as the Lawson-Hanson method [214, Theorem (23.10)] enable to compute such a sparse solution. Let us point out that any solution w^* to (3.31) then satisfies

$$\sum_{n=1}^{N} \left| \sum_{k\in J} w_k^* \varphi_n(X^k) - \bar{\mu}_n \right|^2 \leq \sum_{n=1}^{N} \left| \frac{1}{K_\infty} \sum_{k=1}^{K_\infty} \varphi_n(X^k) - \bar{\mu}_n \right|^2. \tag{3.32}$$

In practice, in the case when $\#J \leq K$, the positions and weights returned by the Carathéodory-Tchakaloff Subsampling procedure are subdivided and randomly perturbed with a small amount of noise.

3.4.5 Test functions scaling

From a numerical perspective, the MCOT approximation of an OT problem is never exactly computed. In particular, constraints are never exactly satisfied, but rather up to a machine precision ϵ : $\|\overline{\Gamma}^K(\overline{Y}^K)\|_\infty \leq \epsilon$ for $\overline{Y}^K \in (\mathbb{R}\times((\mathbb{R}^d)^M)^K$ numerically satisfying the constraints. Hence, replacing $\overline{\Gamma}^K$ by $D\cdot\overline{\Gamma}^K$ in the optimization procedure, for D a non-singular diagonal matrix, can change the numerical solution. We discuss hereafter of a way of choosing an appropriate *scaling* D.

Let $\overline{Y}^{K,*}$ be a minimizer of (3.29)

$$\overline{Y}^{K,*} \in \underset{\substack{\overline{Y}^K \in \left(\mathbb{R}\times((\mathbb{R}^d)^M)\right)^K \\ \overline{\Gamma}^K(\overline{Y}^K)=0, \\ \overline{\Theta}^K(\overline{Y}^K)\leq A}}{\arg\min} \overline{V}^K(\overline{Y}^K),$$

let $\overline{Y}^K \in (\mathbb{R}\times((\mathbb{R}^d)^M)^K$, and let us assume that f and c are $\mathcal{C}^2$. Then,

$$\begin{aligned}
\overline{V}^K(\overline{Y}^{K,*}) &= \overline{V}^K(\overline{Y}^K) + \nabla\overline{V}^K(\overline{Y}^{K,*})^T \cdot (\overline{Y}^{K,*} - \overline{Y}^K) + \mathcal{O}\left(\alpha^2\right) \\
\overline{\Gamma}^K(\overline{Y}^{K,*}) &= \overline{\Gamma}^K(\overline{Y}^K) + \nabla\overline{\Gamma}^K(\overline{Y}^{K,*})^T \cdot (\overline{Y}^{K,*} - \overline{Y}^K) + \mathcal{O}\left(\alpha^2\right),
\end{aligned}$$

where $\alpha = \|\overline{Y}^{K,*} - \overline{Y}^K\|_2$.

As $\overline{Y}^{K,*}$ is a minimizer, $\overline{\Gamma}^K(\overline{Y}^{K,*}) = 0_N$, and there exists $\lambda^* \in \mathbb{R}^N$ such that

$$\nabla\overline{V}^K(\overline{Y}^{K,*}) = \nabla\overline{\Gamma}^K(\overline{Y}^{K,*}) \cdot \lambda^*.$$

Thus,

$$\begin{aligned}
\overline{V}^K(\overline{Y}^{K,*}) - \overline{V}^K(\overline{Y}^K) &= (\overline{Y}^{K,*} - \overline{Y}^K)^T \cdot \nabla\overline{\Gamma}^K(\overline{Y}^{K,*}) \cdot \lambda^* + \mathcal{O}\left(\alpha^2\right) \\
&= -\overline{\Gamma}^K(\overline{Y}^K) \cdot \lambda^* + \mathcal{O}\left(\alpha^2\right).
\end{aligned} \tag{3.33}$$

Therefore, if $\overline{Y}^K$ is a numerical solution, under the hypothesis that for all $1 \leq n \leq N$, $\overline{\Gamma}^K(\overline{Y}^K)_n = \pm\epsilon$, the numerical error on the optimal cost is minimized if there exists $a \in \mathbb{R}$ such that for all $1 \leq n \leq N$, $\lambda_n^* = \pm a$ – which is the condition

for choosing an appropriate scaling D. In practice, although the value of such a λ^* might not be known exactly, it might be of the same magnitude as a solution to

$$\arg\min_{\lambda \in \mathbb{R}^N} \left\| \nabla \overline{\Gamma}^K(\overline{Y}^K) \cdot \lambda - \nabla \overline{V}^K(\overline{Y}^K) \right\|_2^2,$$

if $\overline{\Gamma}^K(\overline{Y}^K)$ is well-conditioned.

3.5 Numerical tests

The aim of this section is to illustrate the results obtained via the numerical procedure described in Section 3.4 for the resolution of the particle problems (3.29) and (3.28) in different test cases.

Section 3.5.1 is devoted to results obtained in cases where $d = 1$ and Section 3.5.2 contains numerical results obtained in examples where $d = 3$. The experiments presented in this section have been implemented in python 3 using scipy and numpy modules, and tested on a server with an Intel Xeon processor with 32 cores (hyper-threaded) and 192 Go RAM.

3.5.1 One-dimensional test cases ($d = 1$)

3.5.1.1 Theoretical elements

In the case where $d = 1$, the solution to the optimal transport problem (3.2) is analytically known in the case when c is a symmetric repulsive cost from [102, Theorem 1.1]. For sake of completeness, we recall their result for the cost function that we consider in our numerical experiments.

Theorem 3.3 (Colombo, De Pascale, Di Marino, 2015)**.** *Let $\epsilon \geq 0$ and $c : \mathbb{R}^M \to [0, +\infty]$ be the cost defined by*

$$\forall x_1, \cdots, x_M \in \mathbb{R}, \quad c(x_1, \ldots, x_M) = \sum_{1 \leq i,j \leq M, i \neq j} \frac{1}{\epsilon + |x_i - x_j|}. \tag{3.34}$$

Let μ be an non atomic probability measure on $\mathbb{R}$ such that

$$\min_{\pi \in \Pi(\mu; M)} \int_{\mathbb{R}^M} c(x_1, \ldots, x_M) d\pi(x_1, \ldots, x_M) < +\infty. \tag{3.35}$$

Let $-\infty = d_0 < d_1 < \cdots < d_M = +\infty$ be such that

$$\mu([d_i, d_{i+1}]) = \frac{1}{M}, \quad i = 0, \ldots, M-1. \tag{3.36}$$

Let $T : \mathbb{R} \to \mathbb{R}$ be the unique (up to μ-null sets) function increasing on each interval $[d_i, d_{i+1}]$, $i = 0, \ldots, M-1$ and such that

$$\begin{aligned} T\#(\mathbf{1}_{[d_i, d_{i+1}]}\mu) &= \mathbf{1}_{[d_{i+1}, d_{i+2}]}\mu, \quad i = 0, \ldots, M-2 \\ T\#(\mathbf{1}_{[d_{M-1}, d_M]}\mu) &= \mathbf{1}_{[d_0, d_1]}\mu. \end{aligned} \tag{3.37}$$

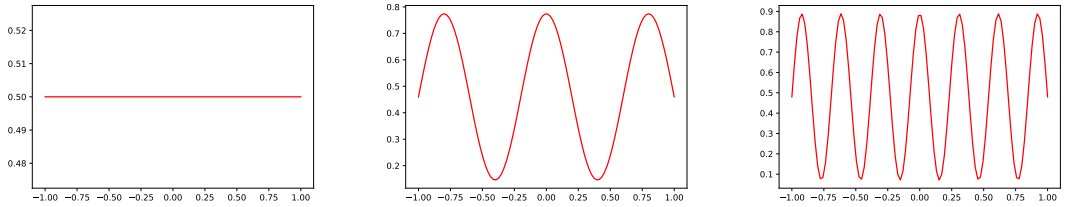
Then T is an admissible map for

$$\inf_{T : \mathbb{R} \to \mathbb{R} \text{ Borel}, T\#\mu = \mu, T^{(M)} = \mathrm{Id}} \int_{\mathbb{R}} c(x, T(x), \ldots, T^{(M-1)}(x)) d\mu(x), \tag{3.38}$$

where $T^{(i)} = \overbrace{T \circ \cdots \circ T}^{i \text{ times}}$.

Moreover, the only symmetric optimal transport plan is the symmetrization of the plan induced by the map T.

We make use of Theorem 3.3 to compare the exact solution of problem (3.2) together with the approximation given by the numerical procedure described in Section 3.4 to solve the MCOT particle problems with fixed or adaptive weights.



(1) μ_1 (2) μ_2 (3) μ_3

Figure 3.1: Densities of the marginal laws tested for 1D numerical tests.

3.5.1.2 Marginals, test functions, cost and weight functions

Marginals. The numerical experiments in this section were realized with three different marginal laws, which are respectively denoted by μ_1, μ_2 and μ_3 and defined by

$$d\mu_1(x) := \frac{1}{2}\mathbf{1}_{[-1,1]}(x)\,dx, \tag{3.39}$$

$$d\mu_2(x) := \left[\frac{\pi}{10}\cos\left(\frac{5\pi}{2}x\right) + 0.46\right]\mathbf{1}_{[-1,1]}(x)\,dx, \tag{3.40}$$

$$d\mu_3(x) := \left[0.13\pi\cos\left(\frac{13\pi}{2}x\right) + 0.48\right]\mathbf{1}_{[-1,1]}(x)\,dx. \tag{3.41}$$

The densities of μ_1, μ_2, μ_3 are plotted in Figure 3.1.

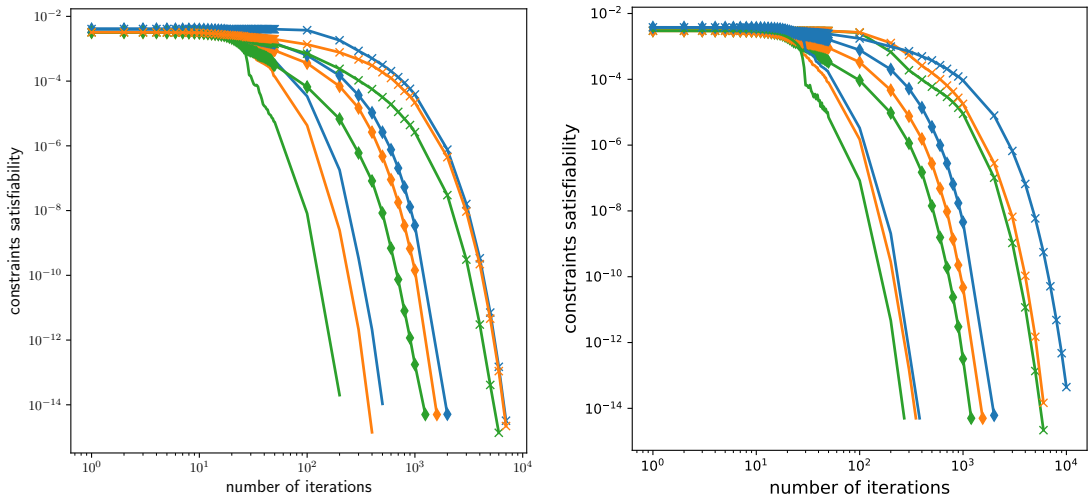
Test functions. The test functions $(\phi_n)_{1\leq n\leq N}$ used are Legendre Polynomials with the following scaling

$$\phi_n = \frac{\sqrt{2n+\frac{1}{2}}}{n+1}P_n, \tag{3.42}$$

where P_n is the Legendre Polynomial of degree n. As the marginal laws considered have their support in $[-1,1]$, we chose the Legendre polynomials for their orthogonality property. Besides, by using polynomials, the matrix $\nabla\Gamma(X)$ is related to a Vandermonde matrix, the invertibility of which (crucial to enforce the constraints by Algorithm 4 or the Runge-Kutta method) is ensured as long as particles are spread on more than N locations. In view of Section 3.4.5, the scaling between the polynomials comes from the assumption of an L^2 convergence of the function $\sum_{n=1}^{N}\lambda_n\phi_n$ as N goes to $+\infty$.

Cost. We use in all experiments the regularized Coulomb cost function (3.34) with $\epsilon = 10^{-1}$.

Weight functions. Two different choices of weight functions f are studied in the numerical experiments presented below: the squared weight function $f : \mathbb{R} \ni a \mapsto a^2$ and the exponential weight function $f : \mathbb{R} \ni a \mapsto e^{-a}$. Although we do not have strong criteria to chose a weight function, the intuition behind the squared weight function is that it can behave well regarding the enforcement of the constraints by a Newton method, given that $\overline{\Gamma}$ is then a polynomial. The intuition behind the exponential weight function is that it could slow down the cancellation of the weights of the particles, keeping alive more degrees of freedom for the optimization process.



(1) μ_2 (2) μ_3

Figure 3.2: Evolution of $\|\Gamma^K(Y_m^K)\|_\infty$ for different weight functions as a function of the number of iterations m of the Runge-Kutta 3 procedure. Tests were performed with $M = 5$, $K = 10000$. Blue curves uses fixed weights, orange curves uses an exponential weight function and green curves a squared weight function. No marker is for $N = 10$, a diamond marker for $N = 20$ and a "+" marker for $N = 40$. Caratheodory-Tchakaloff subsampling gave initial values of 1.11×10^{-16} (3.83×10^{-16}, 3.02×10^{-16}) for μ_2, $N = 10$ (resp. $N = 20$, $N = 40$) and 3.33×10^{-16} (1.28×10^{-16}, 4.66×10^{-16}) for μ_3, $N = 10$ (resp. $N = 20$, $N = 40$).

3.5.1.3 Initialization step – Figure 3.2

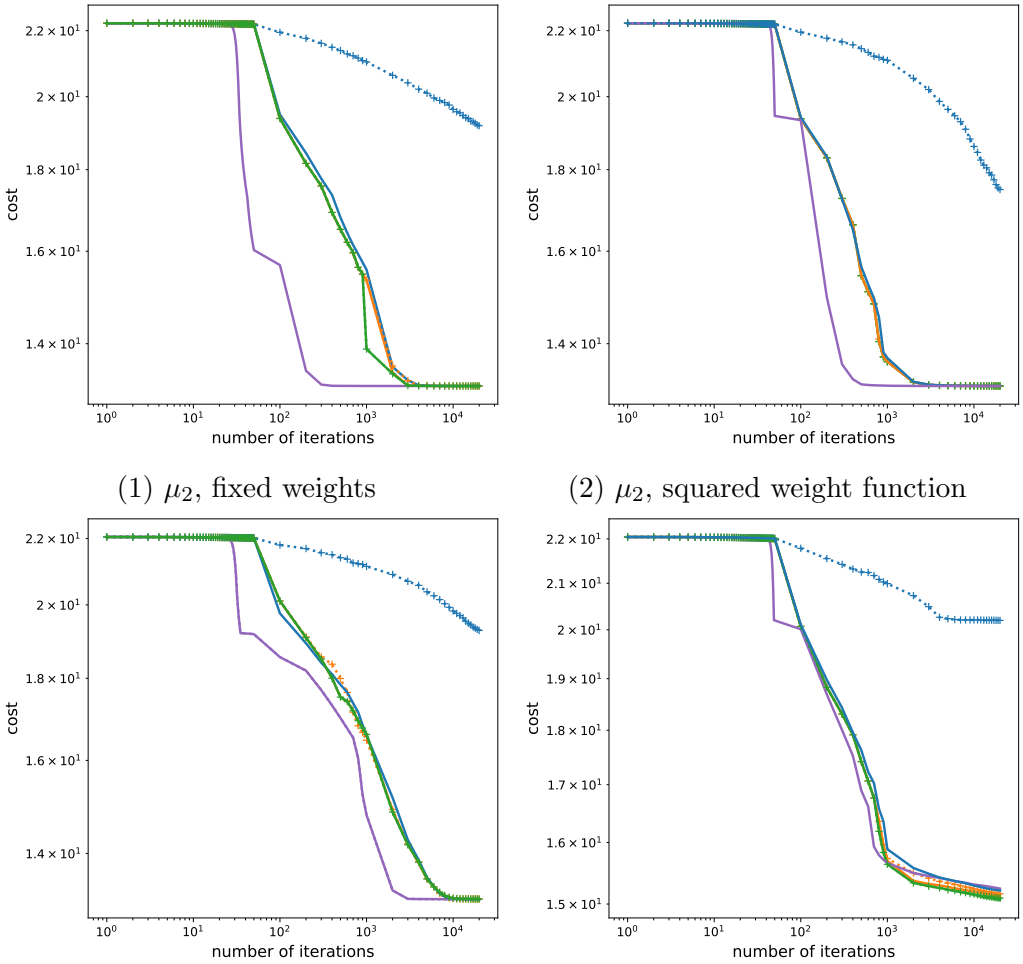
The aim of Figure 3.2 is to plot the decrease of $\|\Gamma^K(Y_m^K)\|_\infty$ as a function of the number of iterations of the Runge-Kutta 3 method presented in Section 3.4.4, in a test case where $M = 5$. We numerically observe here that, as expected, as N increases, the number of iterations needed by the Runge-Kutta procedure to reach convergence increases.[v]Besides, we observe that the additional degrees of freedom of the cases using weight functions allow a faster initial optimization – yet not heavily pronounced, as well as an initialization slightly faster for the squared weight function compared to the exponential one.

3.5.1.4 Decrease of the cost function – Figures 3.3, 3.4 and 3.5

The aim of Figures 3.3, 3.4 and 3.5 is to plot the evolution of $V^K(Y_n^K)$ (or $\overline{V}^K(\overline{Y}_n^K)$) as a function of n the number of iterations of the constrained overdamped Langevin algorithm presented in Section 3.4 for various values of N, various weight functions, values of β_0 and NoiseDecrease functions, and using or not a subsampling at initialization. We observe in Figure 3.3 that decreasing the noise as the squareroot of the number of iterations n converges faster than keeping it constant, and that keeping $\beta_0 = 0$ is the fastest. In Figure 3.4 we remark that the higher N the slower the optimization (with the particular case of μ_3, $N = 20$ with the squared weight function which does not converge in 20000 iterations), and that cases initialized by Caratheodory-Tchakaloff subsampling tend to start with a higher cost. In Figure 3.5, we observe that with $K = 10000$ particles, considering fixed or variable weights does not strongly change the speed of convergence (but for the case μ_3, $N = 20$ with

[v]The problem of finding common roots of polynomials is linked with Bezout's theorem. The complexity of such problems have been studied in [307, 308].

the squared weight function mentioned above). However, using variable weights with $K = 100$ particles seems to be the fastest set of parameters.



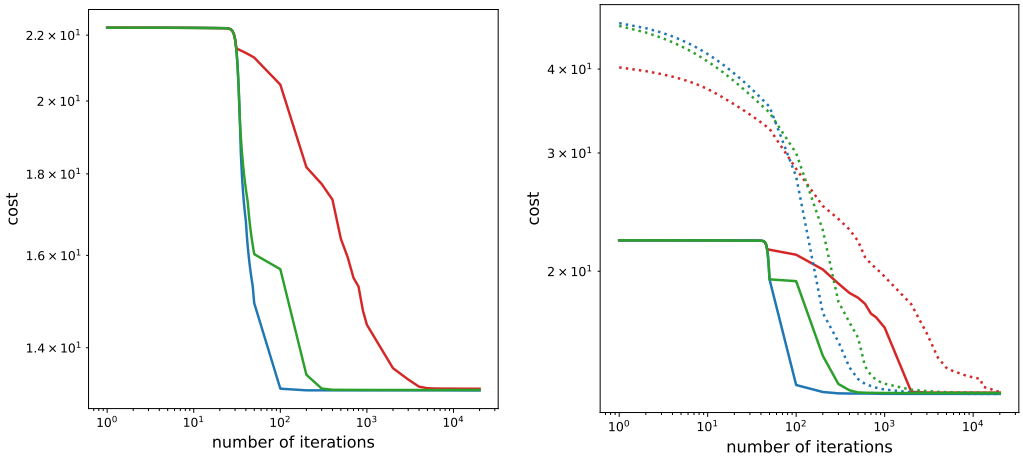
(1) μ_2, fixed weights

(2) μ_2, squared weight function

(3) μ_3, fixed weights

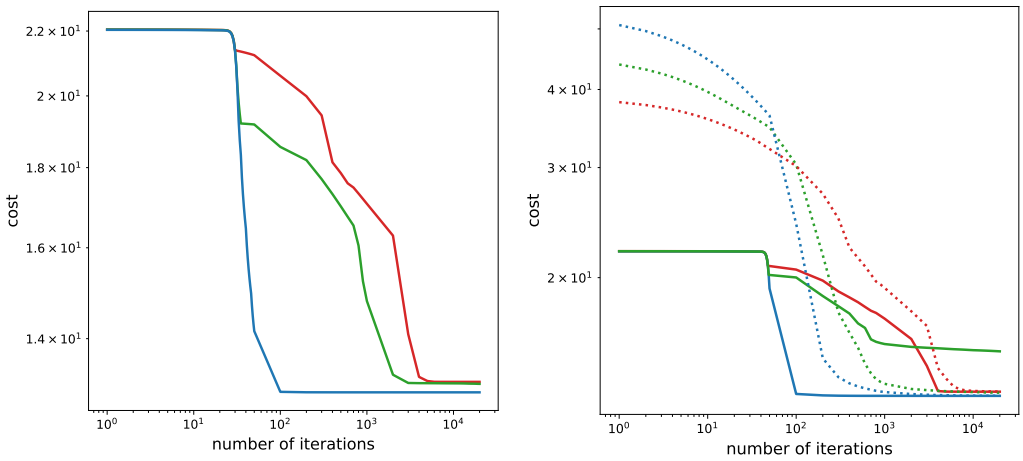
(4) μ_3, squared weight function

Figure 3.3: Evolution of the cost as a function of the number of iterations n for various weight functions and values of β_0, for μ_2 and μ_3. Tests were performed with $M = 5$, $N = 20$, $K = 10000$ and $\Delta t_0 = 10^{-3}$. Blue curves are for $\beta_0 = 10^{-1.5}$, orange curves for $10^{-3.5}$, green curves for $10^{-5.5}$ and purple curves for $\beta_0 = 0$. Solid lines have a decrease of the noise in the squareroot of time whereas dotted lines with a "+" marker have no decrease of the noise.



(1) μ_2, fixed weights

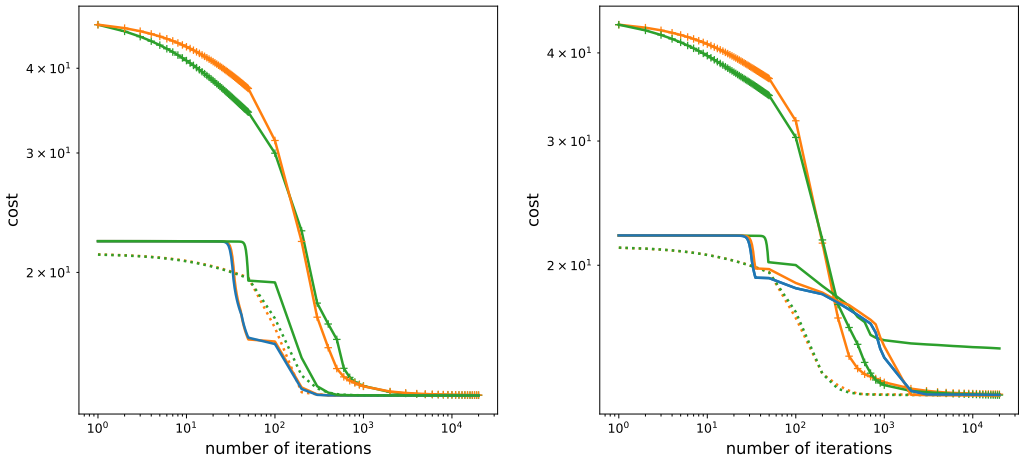
(2) μ_2, squared weight function



(3) μ_3, fixed weights

(4) μ_3, squared weight function

Figure 3.4: Evolution of the cost as a function of the number of iterations n for various weight functions and values of N, for μ_2 and μ_3. Tests were performed with $M = 5$, $\beta_0 = 0$, $K = 10000$ and $\Delta t_0 = 10^{-3}$. Blue curves for $N = 10$, green curves for $N = 20$ and red curves for $N = 40$. Dotted lines correspond to tests initialized by Caratheodory-Tchakaloff subsampling whereas tests solid lines correspond to tests initialized by Runge-Kutta 3 method.



(1) μ_2, (2) μ_3,

Figure 3.5: Evolution of the cost as a function of the number of iterations n for various weight functions, for μ_2 and μ_3. Tests were performed with $M = 5$, $N = 20$, $\beta_0 = 0$ and $\Delta t_0 = 10^{-3}$. Blue curves uses fixed weights, orange curves uses an exponential weight function and green curves a squared weight function. $K = 10000$ particles for solid lines and $K = 100$ particles for dotted lines. Optimization following a Caratheodory-Tchakaloff subsampling at initialization uses "+" markers.

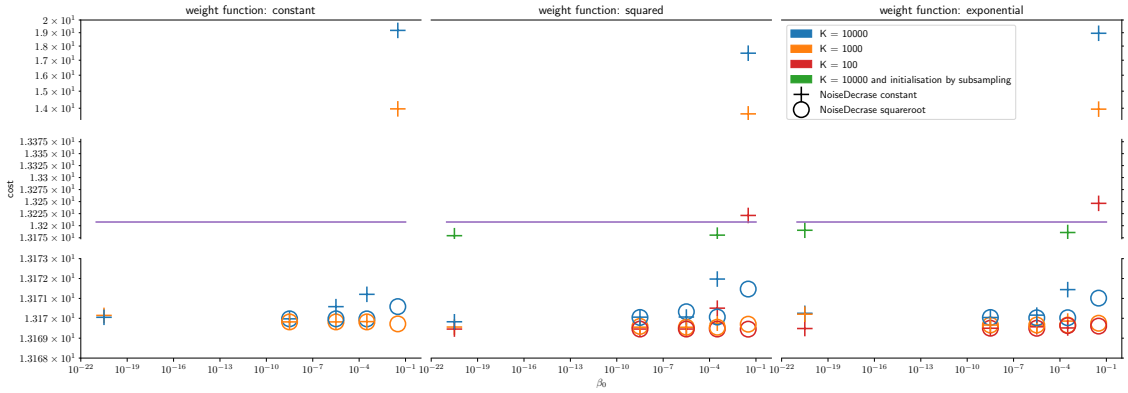


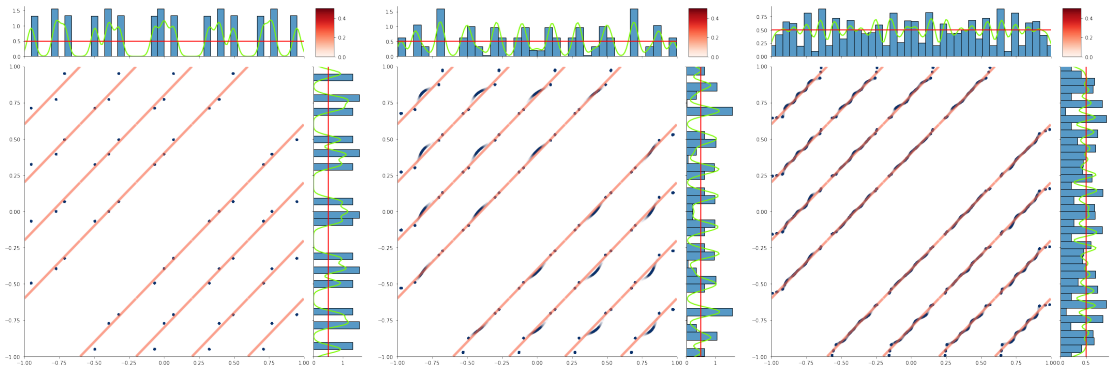
Figure 3.6: Lowest cost value reached during optimization by the constrained overdamped Langevin algorithm in function of the β_0, for various weight functions, values of K and choices of NoiseDecrease functions. The purple line corresponds to the optimal transport cost. The marginal law is μ_2, $N = 20$, $M = 5$, $\Delta t_0 = 10^{-3}$.

3.5.1.5 Minimal values of cost – Figure 3.6

The goal of Figure 3.6 is to compare the minimal values of the cost obtained by the algorithm for different parameters together with its analytic value. We observe that considering adaptive weights enables to reach lower optimal costs than with fixed weights, but the relative difference between the approximate minimal cost values is lower than 0.1%. When the noise level decreases in the square root of the number of iterations a lower optimal cost can be reached compared to a constant noise level. In the variable weights cases, the lower K the lower the optimal cost, but when the optimization starts with a Tchakaloff subsampling solution, for which the lowest cost reached is 0.3% higher than with the Runge-Kutta 3 method.

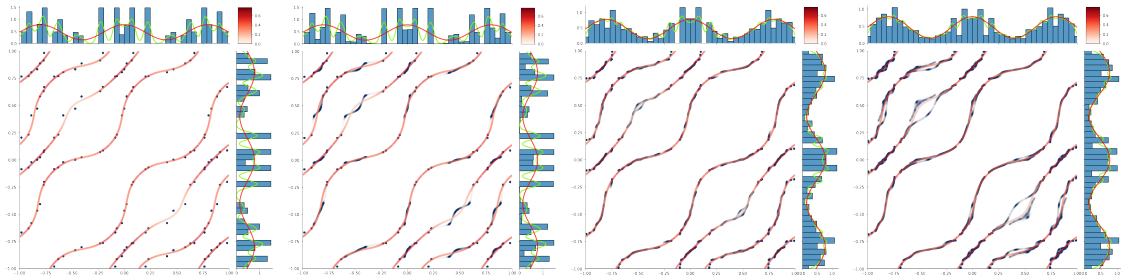
3.5.1.6 Optimal position of particles – Figures 3.7, 3.8 and 3.9

The aim of Figures 3.7, 3.8 and 3.9 is to plot the positions of the particles obtained by the numerical procedure presented in Section 3.4 for respectively μ_1, μ_2 and μ_3 and different values of K, N, β_0, initialization methods, and in fixed and variable weights cases. We numerically observe that the obtained particles are located close to the support of the exact optimal transport plan, and that the higher the value of N the more precise the approximation of this transport map is (see Theorem 2.7) Also, when $K = 10000$ and even more when $\beta_0 = 10^{3.5}$, particles are more spreaded around the transport map.



(1) $N = 10$ (2) $N = 20$ (3) $N = 40$

Figure 3.7: Optimal transport with μ_1 and $M = 5$, $\Delta t_0 = 10^{-3}$. In each plot, on the main graph $\frac{1}{M(M-1)} \sum_{k=1}^{K} \sum_{m \neq m'=1}^{M} w_k \delta_{x_m^k, x_{m'}^k}$ is represented by blue particles. The darker the heavier the particle. Particles have some transparency which allows to see more clearly areas of high concentration. Red curves represent the functions T^i for $i \in \{1, \dots, M-1\}$ defined in Theorem 3.3. The higher the density the darker. On side graphs are represented in blue a weighted histogram of the particles, in red the marginal law and in green a normal kernel density estimate based on the weighted particles (with a bandwidth rule based on Scott's rule with $d = 0$).

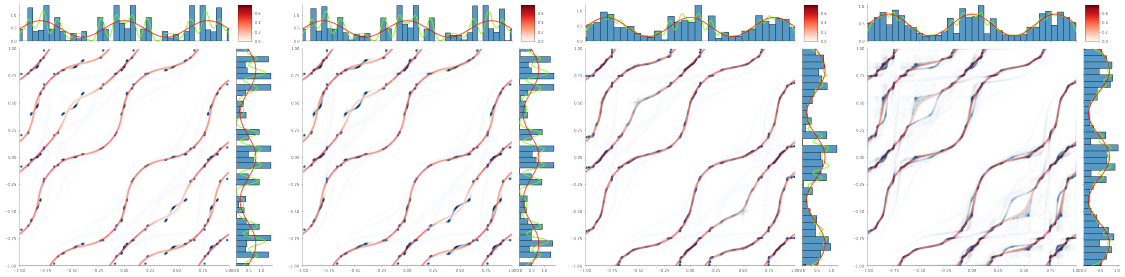


(1) $N = 20$, $\beta_0 = 0$, $K = 10000$, fixed weights

(2) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, fixed weights

(3) $N = 40$, $\beta_0 = 0$, $K = 10000$, fixed weights

(4) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, fixed weights

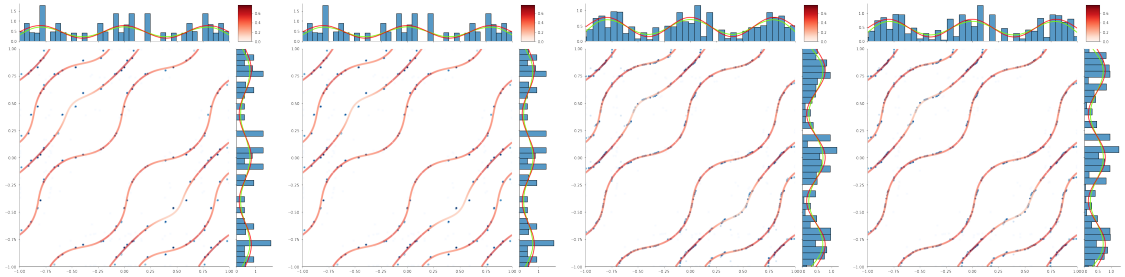


(5) $N = 20$, $\beta_0 = 0$, $K = 10000$, squared weights

(6) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared weights

(7) $N = 40$, $\beta_0 = 0$, $K = 10000$, squared weights

(8) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared weights

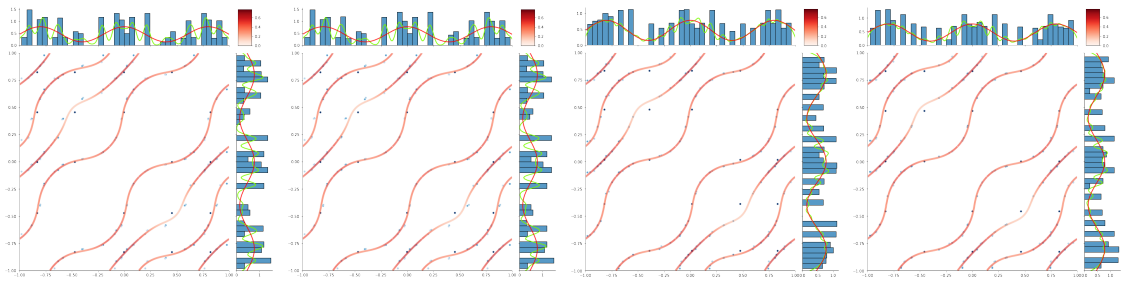


(9) $N = 20$, $\beta_0 = 0$, $K = 100$, squared weights

(10) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 100$, squared weights

(11) $N = 40$, $\beta_0 = 0$, $K = 100$, squared weights

(12) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 100$, squared weights



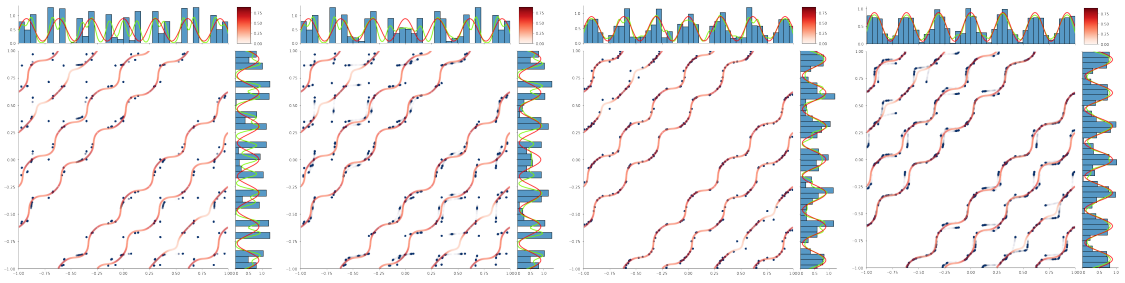
(13) $N = 20$, $\beta_0 = 0$, $K = 10000$, squared weights, with initial subsampling

(14) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared weights, with initial subsampling

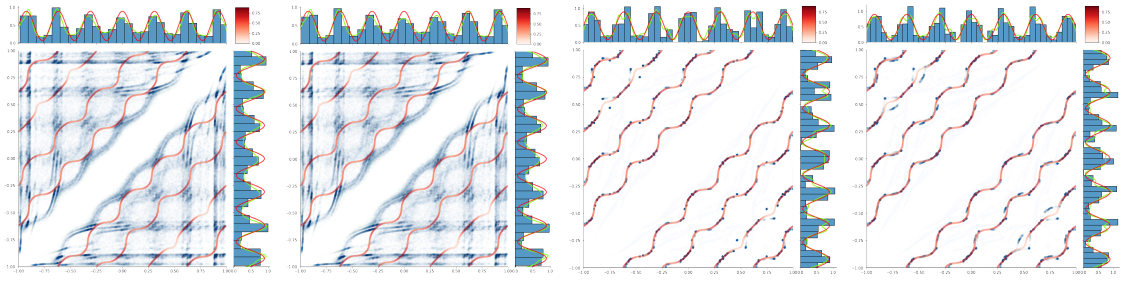
(15) $N = 40$, $\beta_0 = 0$, $K = 10000$, squared weights, with initial subsampling

(16) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared weights, with initial subsampling

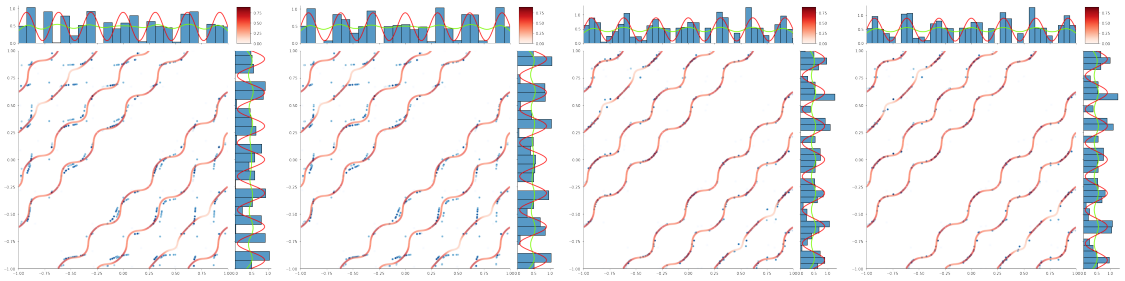
Figure 3.8: Optimal transport with μ_2 and $M = 5$, $\Delta t_0 = 10^{-3}$. In each plot, on the main graph $\frac{1}{M(M-1)} \sum_{k=1}^{K} \sum_{m \neq m'=1}^{M} w_k \delta_{x_m^k, x_{m'}^k}$ is represented by blue particles. The darker the heavier the particle. Particles have some transparency which allows to see more clearly areas of high concentration. Red curves represent the functions T^i for $i \in \{1, \dots, M-1\}$ defined in Theorem 3.3. The higher the density the darker. On side graphs are represented in blue a weighted histogram of the particles, in red the marginal law and in green a normal kernel density estimate based on the weighted particles (with a bandwidth rule based on Scott's rule with $d = 0$).



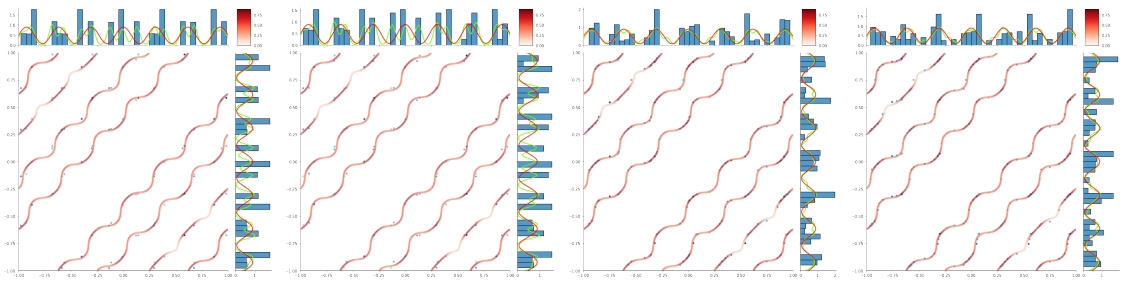
(1) $N = 20$, $\beta_0 = 0$, $K = 10000$, fixed (2) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, fixed (3) $N = 40$, $\beta_0 = 0$, $K = 10000$, fixed (4) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, fixed



(5) $N = 20$, $\beta_0 = 0$, $K = 10000$, squared (6) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared (7) $N = 40$, $\beta_0 = 0$, $K = 10000$, squared (8) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared



(9) $N = 20$, $\beta_0 = 0$, $K = 100$, squared (10) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 100$, squared (11) $N = 40$, $\beta_0 = 0$, $K = 100$, squared (12) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 100$, squared



(13) $N = 20$, $\beta_0 = 0$, $K = 10000$, squared, with initial subsampling (14) $N = 20$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared, with initial subsampling (15) $N = 40$, $\beta_0 = 0$, $K = 10000$, squared, with initial subsampling (16) $N = 40$, $\beta_0 = 10^{-3.5}$, $K = 10000$, squared, with initial subsampling

Figure 3.9: Optimal transport with μ_3 and $M = 5$, $\Delta t_0 = 10^{-3}$. In each plot, on the main graph $\frac{1}{M(M-1)} \sum_{k=1}^{K} \sum_{m \neq m'=1}^{M} w_k \delta_{x_m^k, x_{m'}^k}$ is represented by blue particles. The darker the heavier the particle. Particles have some transparency which allows to see more clearly areas of high concentration. Red curves represent the functions T^i for $i \in \{1, \dots, M-1\}$ defined in Theorem 3.3. The higher the density the darker. On side graphs are represented in blue a weighted histogram of the particles, in red the marginal law and in green a normal kernel density estimate based on the weighted particles (with a bandwidth rule based on Scott's rule with $d = 0$).

3.5.2 Three-dimensional test cases ($d = 3$)

3.5.2.1 Tests design

The numerical experiments were realized with four different marginal laws that are named afterwards as follows:

$$\mu_1 \sim \mathcal{N}(0_3, \mathrm{Id}_3), \tag{3.43}$$

$$\mu_2 \sim \frac{2}{3}\mathcal{N}(0_3, \begin{pmatrix} 1 & 0.5 & 0.75 \\ 0.5 & 2 & 1.5 \\ 0.75 & 1.5 & 3 \end{pmatrix}) + \frac{1}{3}\mathcal{N}(\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0.8 & 0.22 \\ 0.8 & 2 & 1.8 \\ 0.22 & 1.8 & 3 \end{pmatrix}), \tag{3.44}$$

$$\mu_3 \sim \frac{1}{10}\mathcal{N}(0_3, C) + \frac{1}{5}\mathcal{N}(\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}, C) + \frac{1}{5}\mathcal{N}(\begin{pmatrix} 8 \\ 0 \\ 0 \end{pmatrix}, C) + \frac{1}{5}\mathcal{N}(\begin{pmatrix} 12 \\ 0 \\ 0 \end{pmatrix}, C)$$
$$+ \frac{1}{5}\mathcal{N}(\begin{pmatrix} 16 \\ 0 \\ 0 \end{pmatrix}, C) + \frac{1}{10}\mathcal{N}(\begin{pmatrix} 20 \\ 0 \\ 0 \end{pmatrix}, C), \qquad \text{with} \quad C = \begin{pmatrix} 1 & 0.5 & 0.75 \\ 0.5 & 2 & 1.5 \\ 0.75 & 1.5 & 3 \end{pmatrix}, \tag{3.45}$$

$$\mu_4 \sim \mathcal{U}(\mathcal{B}(0,1)). \tag{3.46}$$

And, for $i = 1, 2, 3, 4$, using as test functions[vi] tensor products of 1D orthonormal polynomials $(P_l^{\mu_i,j})_{\substack{1\le j\le 3, \\ l\in\mathbb{N}}}$, defined as, for $j = 1, 2, 3$, $l \in \mathbb{N}$,

$$\text{degree}\left(P_l^{\mu_i,j}\right) = l, \quad \forall l' < l, \int_{\mathbb{R}^3} P_l^{\mu_i,j}(x_j) P_{l'}^{\mu_i,j}(x_j) d\mu_i(x_1, x_2, x_3) = \frac{1}{(l+1)^2}\delta_{l,l'}. \tag{3.47}$$

As for a finite number of multivariate polynomials (and under a suitable control of mixed derivatives), the hyperbolic cross [113] seems to behave better than using all polynomials up to a given degree, we used, for a number of test functions N appropriately chosen, the polynomials $P_{l_1}^{\mu_i,1} \otimes P_{l_2}^{\mu_i,2} \otimes P_{l_3}^{\mu_i,3}$, where

$$(l_1 + 1)(l_2 + 1)(l_3 + 1) \le L_N, \tag{3.48}$$

where L_N is defined such that $\#\{(l_1, l_2, l_3) | (l_1 + 1)(l_2 + 1)(l_3 + 1) \le L_N\} = N$. The map between maximum degree of the polynomials $(L_N - 1)$ and N is shown in Table 3.1.

$L_N - 1$	6	7	8	9	10	11
N	28	38	44	53	56	74

Table 3.1: Map between the maximum degree of 1D polynomials and the number of test functions using hyperbolic cross in 3D.

In the numerical examples presented afterwards, as all weights are fixed to $\frac{1}{K}$, there is no need to use the polynomial of degrees $(l_1, l_2, l_3) = (0, 0, 0)$, hence values of N decreased by 1 compared to the values of Table 3.1.

[vi] These polynomials were chosen after a few numerical tests on some optimization procedures for their better convergence properties than the polynomials they were compared to. Their tensorised form both eases the computation of the moments and allows some parallelisation. Note also that the matrix $\nabla\Gamma^K(Y^K)$ is a multivariate Vandermonde matrix. We checked numerically its invertibility throughout the optimization process.

Remark 3.4. *One of the main advantages of using sums of Normal functions (or a uniform measure on a ball) as marginal laws and polynomials as test functions is that their exists in that case close formulas for the computation of the moments (see Appendix B.1). From our experiments in dimension 1, the precision of the computation of the moments is important both for the solution of the MCOT problem to be well-defined (and thus for the algorithm to converge) – numerically computed moments, though not exact, must allow the existence of* $Y^K \in ((\mathbb{R}^d)^M)^K$ *such that* $\|\Gamma^K(Y^K)\|_\infty \leq \epsilon$ *for* ϵ *the machine-precision; and for the convergence as* N *increases of the MCOT cost towards the OT cost – numerically computed moments not precise enough might hide this convergence. Numerical quadratures in 3D could be implemented for dealing with more general marginal laws and test functions, however, their computation and convergence speed put it beyond the scope of this article.*

Mean-Covariance. Tests were also performed using as test functions the mean and covariance matrix for μ_1 and μ_2, in order to notice on examples how much those test functions do constrain an optimal transport problem. Note that this problem of optimal transport when the mean and covariance structure are given may be interesting per se, when only partial information on the distribution is known. We have indicated in Table 3.2 the optimal costs obtained with our algorithm for μ_1 and μ_2 with mean-covariance constraints ($N = 9$) and with many moment constraints ($N = 52$). We observe on our examples a relative difference around 15-20%.

	μ_1, $M = 10$	μ_1, $M = 100$	μ_2, $M = 10$	μ_2, $M = 100$
$N = 9$	10.65	1395	8.007	1074
$N = 52$	12.50	1599	9.107	1201

Table 3.2: Optimal value of the cost obtained for μ_1 and μ_2 with mean-covariance constraints ($N = 9$) and with many moment constraints ($N = 52$).

Cost. In order to avoid too high values of the cost function, we used in all experiments a regularized Coulomb cost $c(x_1, \dots, x_M) = \sum_{m \neq m'=1}^{M} \frac{1}{\epsilon + |x_m - x_{m'}|}$, with $\epsilon = 10^{-3}$ and $\forall i = 1, \dots, M, x_m \in \mathbb{R}^3$.

Fixed weights. After several tests comparing fixed and variable weights (with various weight functions), we observe that in dimension 3, for the marginal laws considered, both initialization and optimization using variable weights were much slower than using fixed weights. Therefore, all following tests have been performed using fixed weights. Heuristically, when using variable weights, some particles tend to have large weights and are strongly constrained while other ones become lightweight and do not move much since the gradient on positions is proportional to weights.

3.5.2.2 Initialization and constraints enforcement – Figure 3.10

Initialization was performed by a sampling K particles according to the marginal law, and then using the Runge-Kutta method showed in Section 3.4.4 to bring the particles on the submanifold of the constraints $\mathcal{M}^K$. This method has been tested for various values of N and K, presented respectively in Figures 3.10.

As N increases (Figure 3.10), the submanifold of the constraints becomes harder to reach using the Runge-Kutta method (similarly to the 1D case)[vii], and large

values of N ($L_N \geq 11$) could not be attained in the time of the numerical experiment (remind that the number of computations involved at each iteration grows linearly with N). In the case of each marginal laws (μ_1 and μ_2) for which the tests have been performed, despite the assymmetry of μ_2, the dependence on N of the convergence speed appears to be similar.

Note also that as we use symmetrised test functions (regarding the marginal laws) with fixed weights, the number of independant coordinates involved in the Runge-Kutta method to satisfy the constraints is linear in KM (M being the number of marginal laws). Thus, solving the problem of finding a starting point on the submanifold with 100 marginal laws and 10^3 particles is numerically the same as the one with 10 marginal laws and 10^4 particles. Although in the case where weights are variable this remark can not be applied, as coordinates on different marginal laws of the same particle share the same weight, increasing the number of marginal laws relaxes the problem of finding a starting point on the submanifold.

3.5.2.3 Optimization procedure – Figures 3.11, 3.12, 3.13 and 3.14

The aim of Figures 3.11 and 3.12 is to plot the evolution of $V^K(Y_n^K)$ as a function of n the number of iterations of the constrained overdamped Langevin algorithm presented in Section 3.4 for various values of N and values of β_0. As we observed (Figure 3.11), and similarly to the tests in dimension 1, that tests with $\beta_0 = 0$ converges faster than $\beta_0 > 0$, we kept $\beta_0 = 0$ for all the other tests. The convergence of the cost for various values of N and K, various number of marginal laws and for μ_1 and μ_2 is presented in Figure 3.12. And a presentation of how particles move during the optimization procedure can be seen in Figures 3.13 and 3.14.

On all subgraphs of Figure 3.12, one can observe that the optimization procedure reaches a cost close to the optimal one for the MCOT problem in 50-200 iterations, when K is large enough for a given N (e.g. $K = 1000$ is sufficient when $N = 27$ but not when $N = 43$). As N increases the value of the optimal costs does as well, which is expected, as MCOT problems get more and more constrained. As K increases, the value of the cost computed converges towards the MCOT cost. Indeed, the slight decrease of the computed MCOT cost at the 20000^{th} iteration as K increases that can be observed in Table 3.3 from $K = 320$ to $K = 10000$ suggests that their exists $K_0 \in \mathbb{N}$ such that for $K \geq K_0$, the gain in an increase in K reflects weakly on the MCOT cost computed.

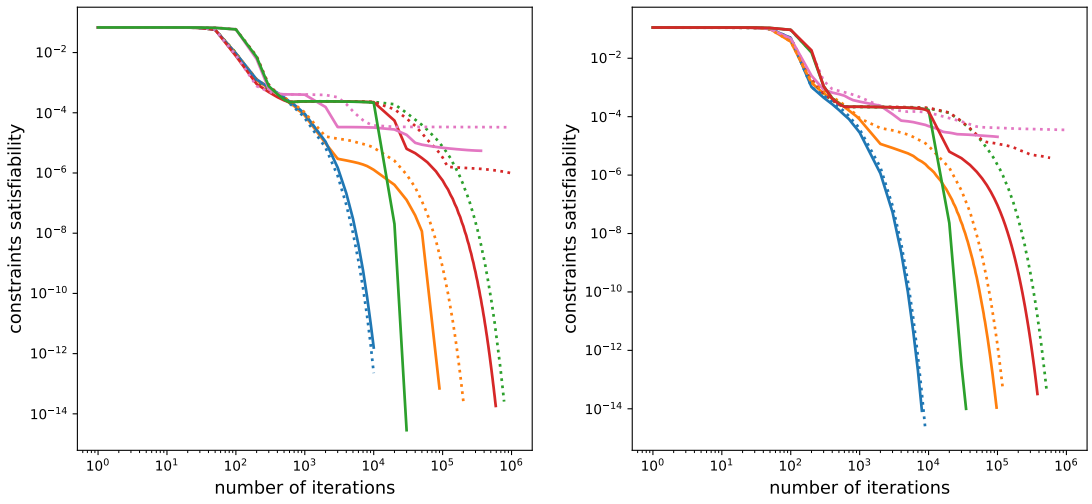
On Figures 3.13 and 3.14 is plotted the evolution of some symmetrized visualizations of the process during the optimization for an MCOT problem on μ_1. Although at each iteration it satisfies the moment constraints, it deviates from a Normal sample rapidly and tends to concentrate on some points (a bit like in Tchakaloff's theorem and Theorem 2.3).

3.5.2.4 Minimas – Figures 3.15, 3.16, 3.17 and 3.18

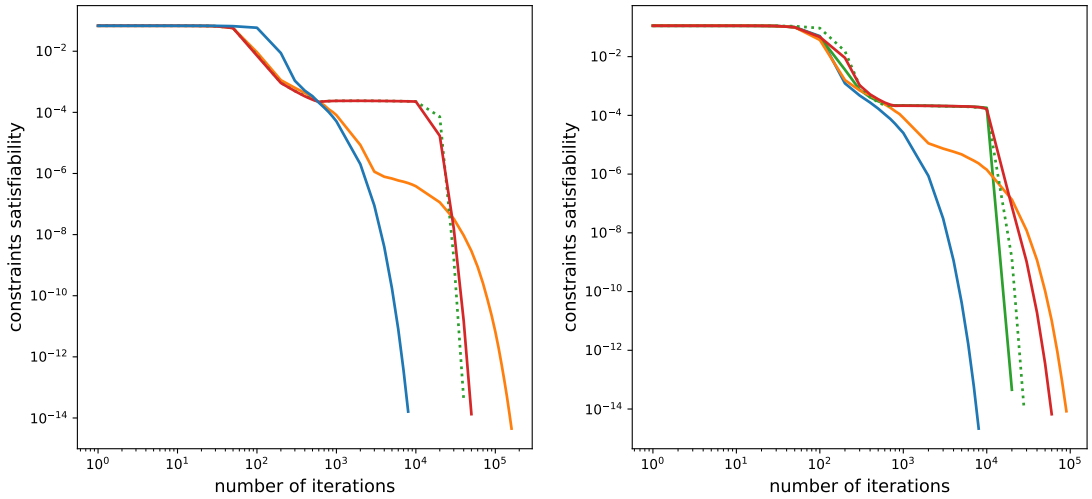
As K increases, the symmetrized minimizers of Figures 3.15 and 3.16 tends to be visually more and more concentrated on some particular points. According to Table 3.3, higher values of K tends to have lower costs.

Some symmetrized visualizations of minimizers for MCOT problems for the non-symmetrical measures μ_2 and μ_3 are presented in Figures 3.17 and 3.18. In those cases, the 1D couplings obtained on each axis (X, Y or Z) are not the same (Figure

[vii] Similarly to Footnote v, the problem of finding particles satisfying the constraints is linked with the multivariate Bezout's theorem [307, 308].



(1) μ_1, $M = 10$ (2) μ_2, $M = 10$



(3) μ_1, $M = 100$ (4) μ_2, $M = 100$

Figure 3.10: Evolution of $\|\Gamma^K(Y_m^K)\|_\infty$ for values of N ranging from 27 to 52 and K between 1000 (dotted lines) and 10000 (solid lines) as a function of the number of iterations m of the Runge-Kutta 3 procedure. $\Delta t_0 = 10^{-4}$. Blue curves are for $N = 27$, orange ones for $N = 37$, green ones for $N = 43$, red ones for $N = 52$ and pink ones for $N = 73$.

3.17). A higher number of marginal laws M seems to spread more the particules, although their 1D coupling still shows particles highly concentrated around a few values in the considered examples. Higher values of N increases the concentration of the particles around fewer values in the μ_3 examples. The planar representation of the minimizers for large M (Figure 3.18), shows that particules are not distributed spatially as a Normal function and tend to concentrate on some 1D curves (for the considered 2D projections) with a higher spreading than for lower values of M.

3.5.2.5 Optimization for μ_4 - Figure 3.19

Optimal transport for μ_4 with a large number of electrons is of theoretical interest as it might provide approximations for a uniform electronic density in a large space [227]. Numerical results for its MCOT relaxation with $M = 100$ and $N = 52$ are presented in Figure 3.19. Although the cost has been optimized (Figure 3.19.1), it is only 3% lower than the initial uniform sampling (after a Runge-Kutta 3 initial-

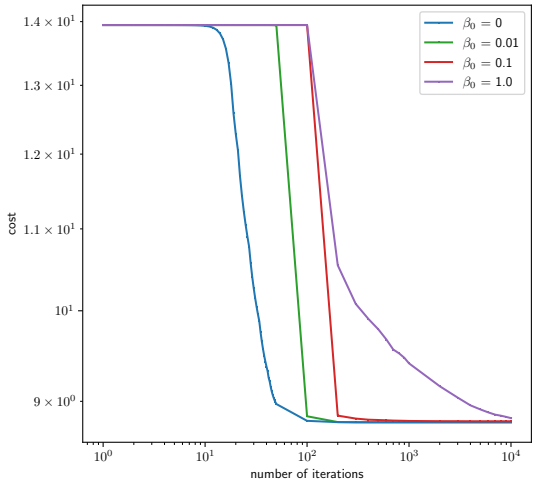
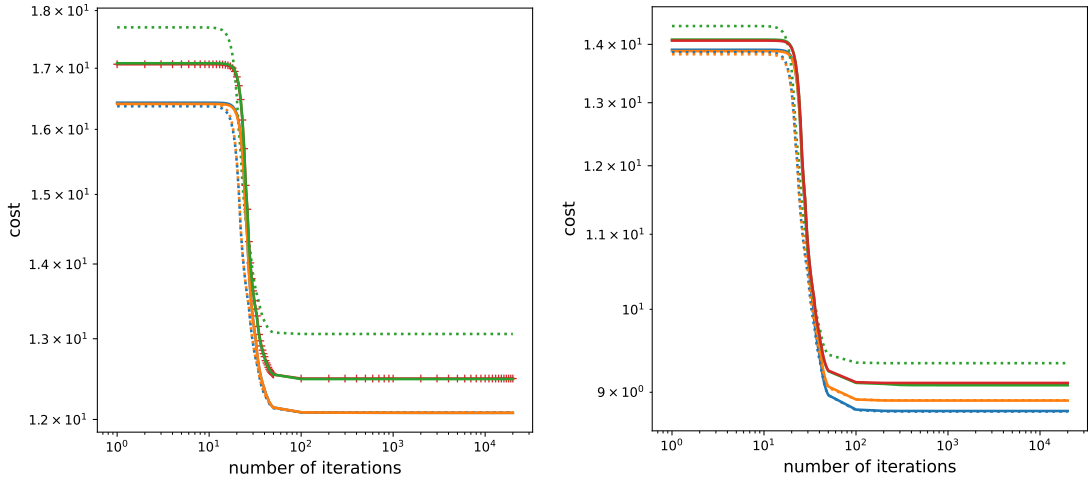


Figure 3.11: Evolution of the cost as a function of the number of iterations n for various values of β_0. The marginal law is μ_2, β_0 varies from 0 to 1 and the other parameters are $\Delta t_0 = 10^{-4}$, noise level decreases as the squareroot of the number of iterations, $N = 27$, $M = 10$, $K = 160$.

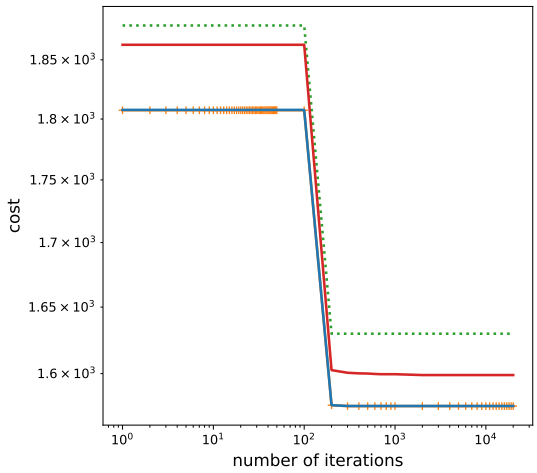
K	40	80	160
cost	12.2558198	12.1747815	12.1457150
lower cost	12.1981977	12.0864398	12.0862042
K	320	1000	10000
cost	12.0916662	12.0821615	12.0785749
lower cost	12.0855486	12.0821615	12.0785745

Table 3.3: Values of the regularized Coulomb cost (see here-named paragraph in Section 3.5.2.1) for the MCOT problem with μ_1, $M = 10$, $N = 27$, $\Delta t_0 = 10^{-4}$, $\beta_0 = 0$ and K ranging from 40 to 10000. The *cost* line corresponds to the value of the regularized cost associated to the minimizing process at iteration 20000 (which also corresponds to the minimizers represented in the graphs of Figures 3.15 and 3.16). The *lower cost* line corresponds to the lower value of the regularized cost encountered by the minmizing process before or at iteration 20000 for each value of K.

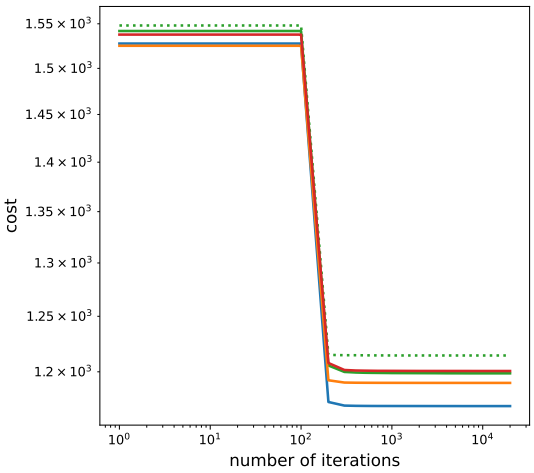


(1) μ_1, $M = 10$

(2) μ_2, $M = 10$

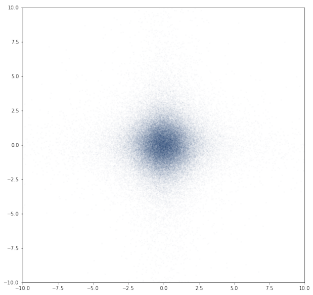


(3) μ_1, $M = 100$

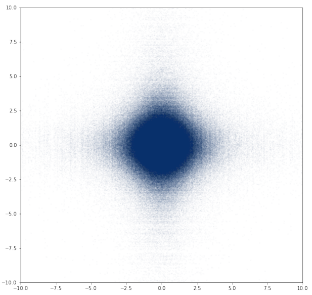


(4) μ_2, $M = 100$

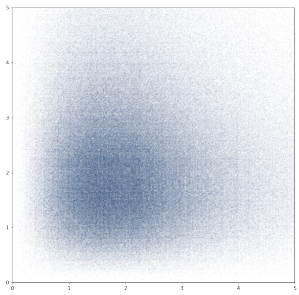
Figure 3.12: Evolution of the cost as a function of the number of iterations n for various values of N and K from 1000 (dotted lines) to 10000 (solid lines). $\Delta t_0 = 10^{-4}$, $\beta_0 = 0$. Blue curves are for $N = 27$, orange ones for $N = 37$, green ones for $N = 43$, red ones for $N = 52$ and pink ones for $N = 73$. On Figures 3.12.1 and 3.12.3, "+" signs are added to better distinguish overlaid curves.



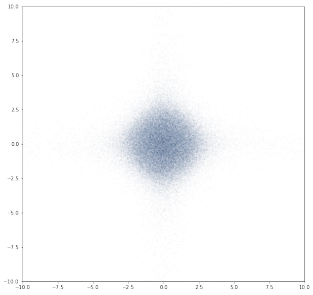
(a1) plane XY, iteration 1



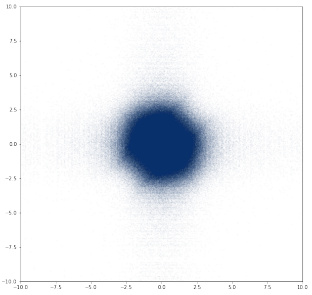
(b1) X axis, iteration 1



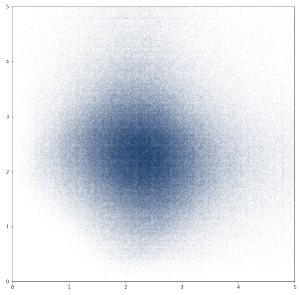
(c1) radial, iteration 1



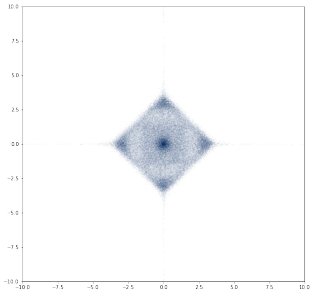
(a2) plane XY, iteration 30



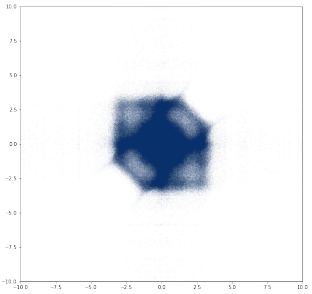
(b2) X axis, iteration 30



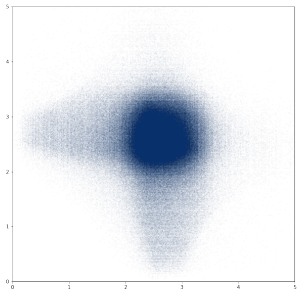
(c2) radial, iteration 30



(a3) plane XY, iteration 50

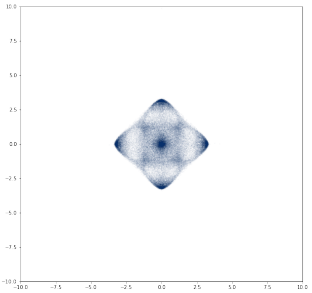


(b3) X axis, iteration 50

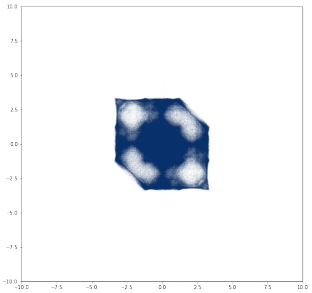


(c3) radial, iteration 50

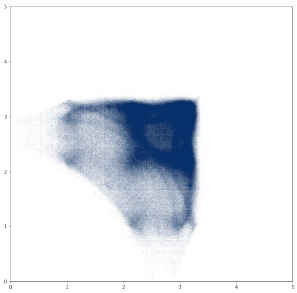
Figure 3.13: Transport along optimization for μ_1, $M = 10$, $K = 10000$, $N = 27$, $\beta_0 = 0$, $\Delta t_0 = 10^{-4}$. In figures of column (a) is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In figures of column (b) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In figures of column (c) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. The evolution of the corresponding cost can be seen in Figure 3.12.1.



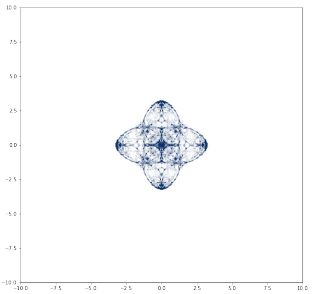
(a4) plane XY, iteration 100



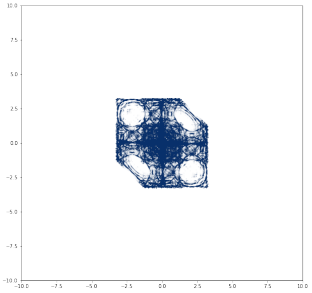
(b4) X axis, iteration 100



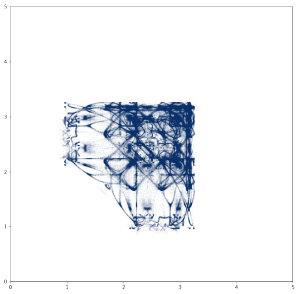
(c4) radial, iteration 100



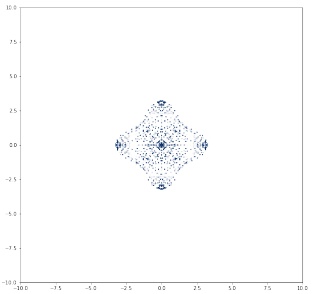
(a5) plane XY, iter. 1000



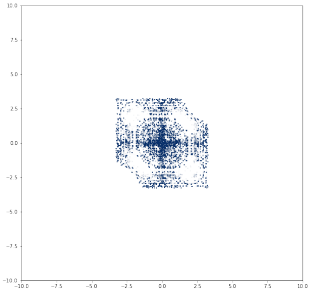
(b5) X axis, iteration 1000



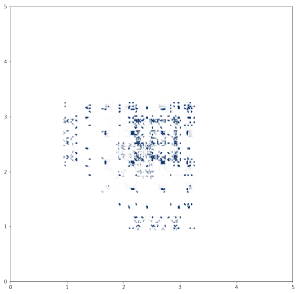
(c5) radial, iteration 1000



(a6) plane XY, iter. 20000

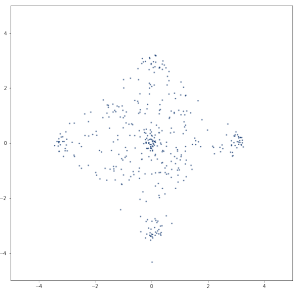


(b6) X axis, iteration 20000

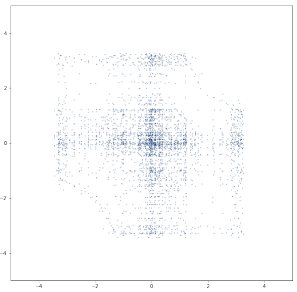


(c6) radial, iteration 20000

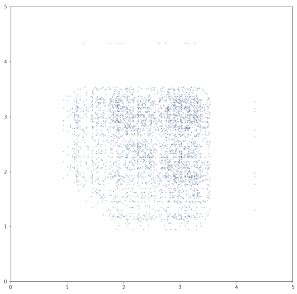
Figure 3.14: Transport along optimization for μ_1, $M = 10$, $K = 10000$, $N = 27$, $\beta_0 = 0$, $\Delta t_0 = 10^{-4}$. In figures of column (a) is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In figures of column (b) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In figures of column (c) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. The evolution of the corresponding cost can be seen in Figure 3.12.1.



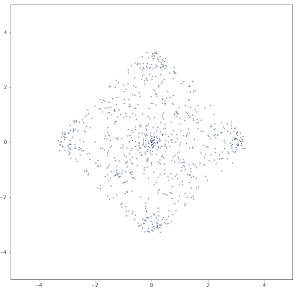
(a1) plane XY, K=40



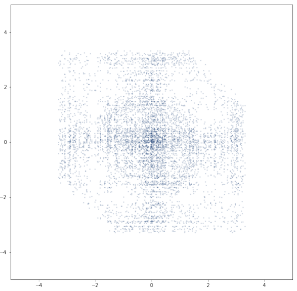
(b1) X axis, K=40



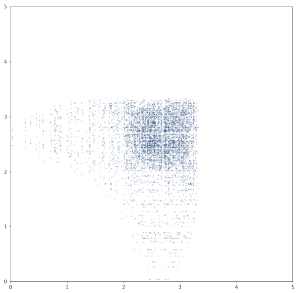
(c1) radial, K=40



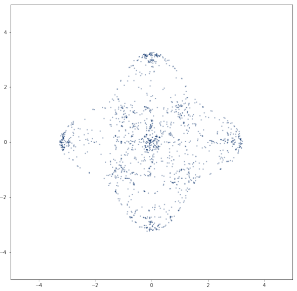
(a2) plane XY, K=80



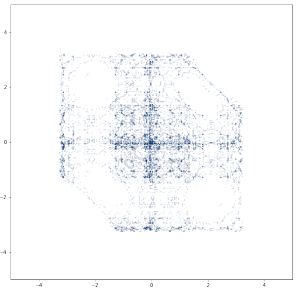
(b2) X axis, K=80



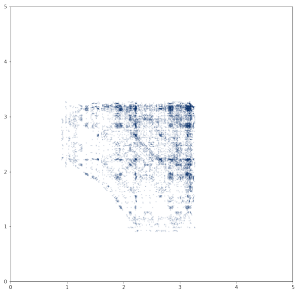
(c2) radial, K=80



(a3) plane XY, K=160

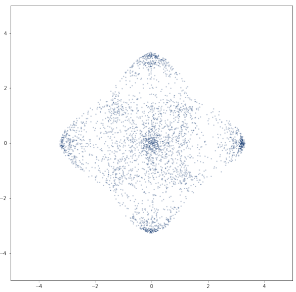


(b3) X axis, K=160

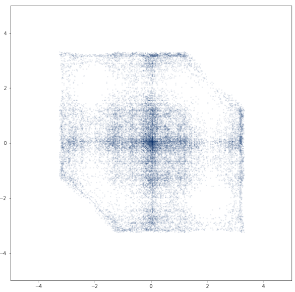


(c3) radial, K=160

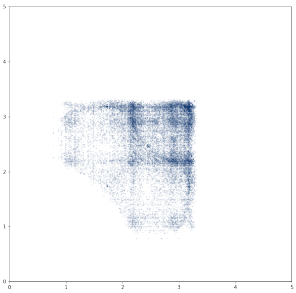
Figure 3.15: Optimal transport with μ_1, $M = 10$, $N = 27$, $\beta_0 = 0$ and $\Delta t_0 = 10^{-4}$, for $K = 40, 80, 160$. In figures of column (a) is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In figures of column (b) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In figures of column (c) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. Corresponding costs can be found in Table 3.3.



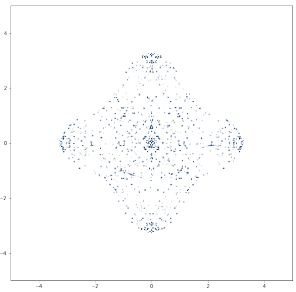
(a4) plane XY, K=320



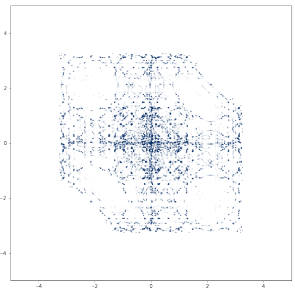
(b4) X axis, K=320



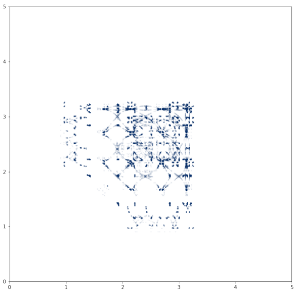
(c4) radial, K=320



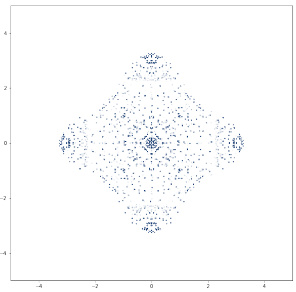
(a5) plane XY, K=1000



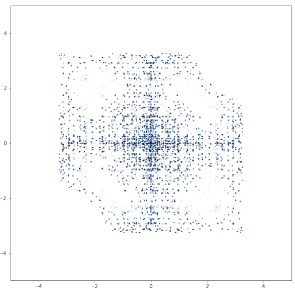
(b5) X axis, K=1000



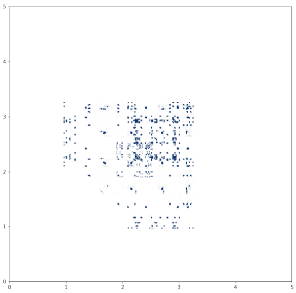
(c5) radial, K=1000



(a6) plane XY, K=10000

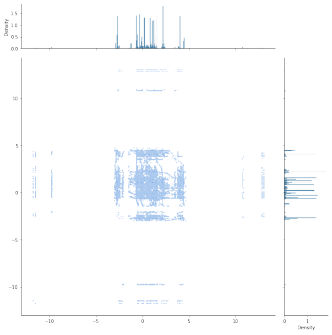


(b6) X axis, K=10000

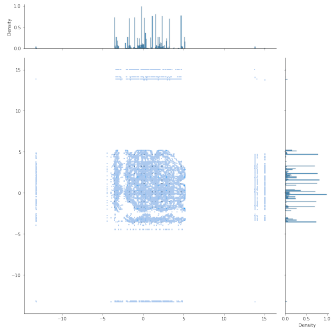


(c6) radial, K=10000

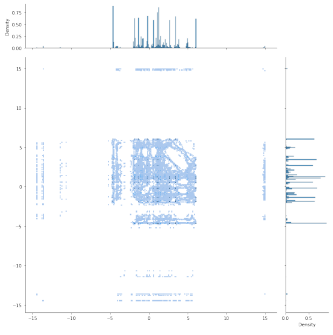
Figure 3.16: Optimal transport with μ_1, $M = 10$, $N = 27$, $\beta_0 = 0$ and $\Delta t_0 = 10^{-4}$, for $K = 320, 1000, 10000$. In figures of column (a) is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In figures of column (b) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In figures of column (c) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. Corresponding costs can be found in Table 3.3.



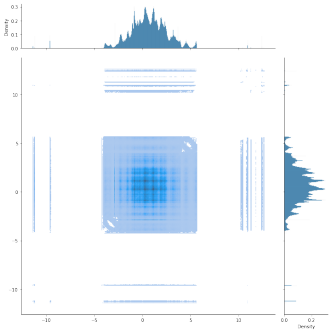
(a1) μ_2, X axis, $M = 10$, $N = 27$



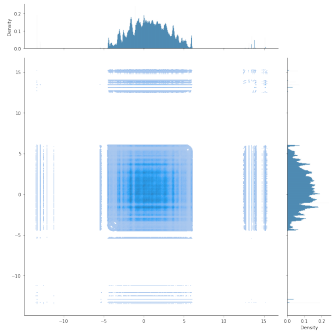
(b1) μ_2, Y axis, $M = 10$, $N = 27$



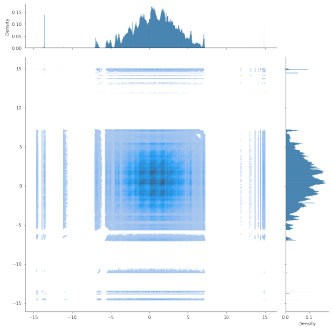
(c1) μ_2, Z axis, $M = 10$, $N = 27$



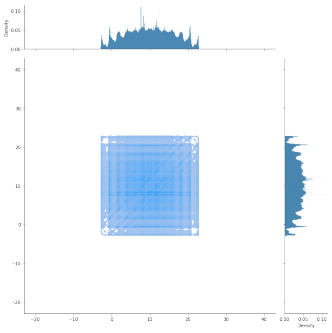
(a2) μ_2, X axis, $M = 100$, $N = 27$



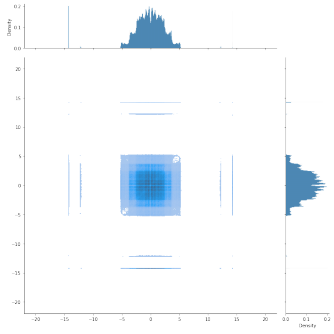
(b2) μ_2, Y axis, $M = 100$, $N = 27$



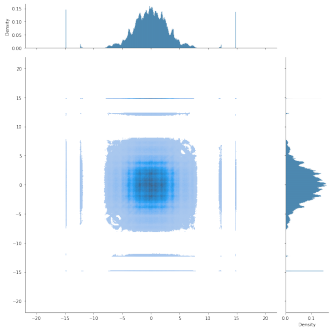
(c2) μ_2, Z axis, $M = 100$, $N = 27$



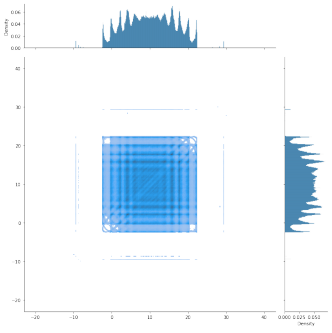
(a3) μ_3, X axis, $M = 100$, $N = 27$



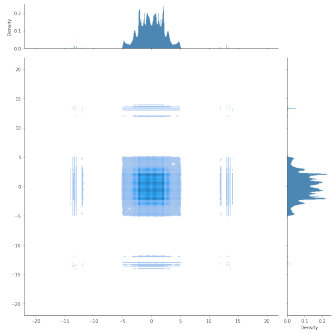
(b3) μ_3, Y axis, $M = 100$, $N = 27$



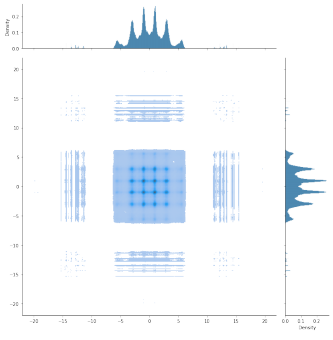
(c3) μ_3, Z axis, $M = 100$, $N = 27$



(a4) μ_3, X axis, $M = 100$, $N = 52$

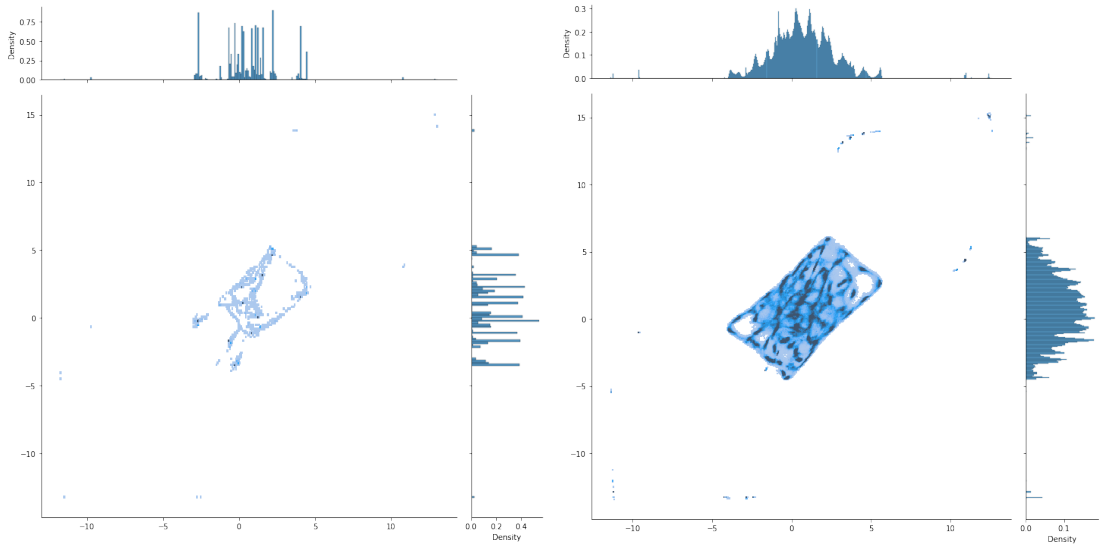


(b4) μ_3, Y axis, $M = 100$, $N = 52$

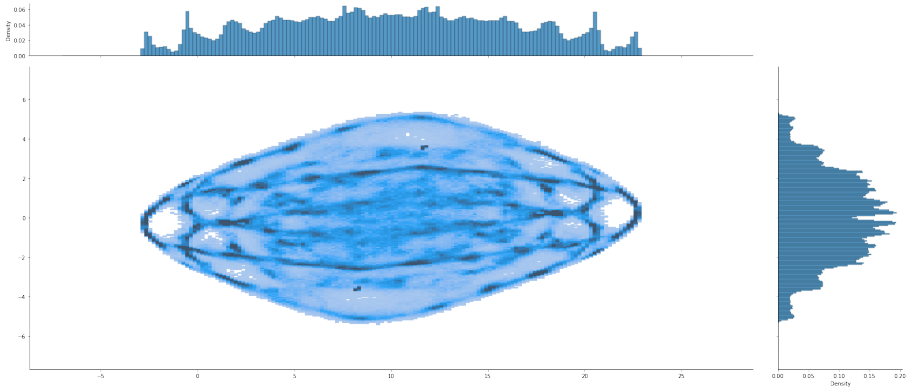


(c4) μ_3, Z axis, $M = 100$, $N = 52$

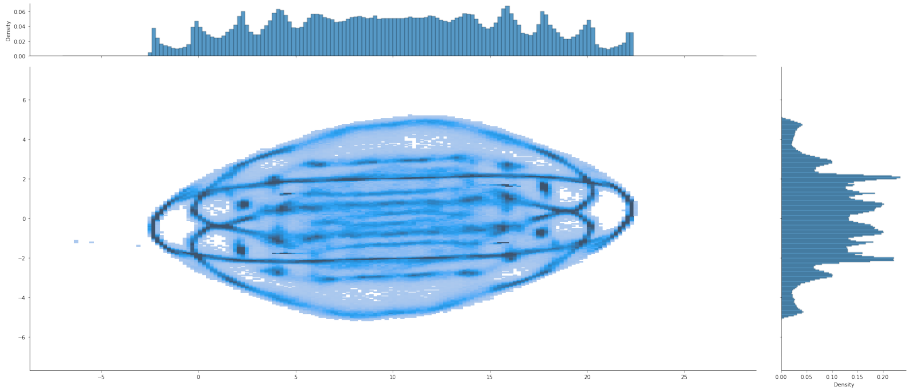
Figure 3.17: Optimal transport for μ_2 and μ_3, $M = 10, 100$, $N = 27, 52$, $\beta_0 = 0$, $K = 10000$ and $\Delta t_0 = 10^{-4}$ In figures of column (a) is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In figures of column (b) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In figures of column (c) is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. In order to better distinguish between areas of low and high particles density, plots are represented as 2D histograms.



(1) μ_2, plane XY, $M = 10$, $N = 27$ (2) μ_2, plane XY, $M = 100$, $N = 27$



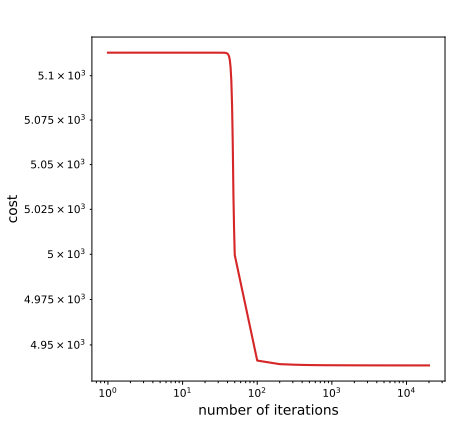
(3) μ_3, plane XY, $M = 100$, $N = 27$



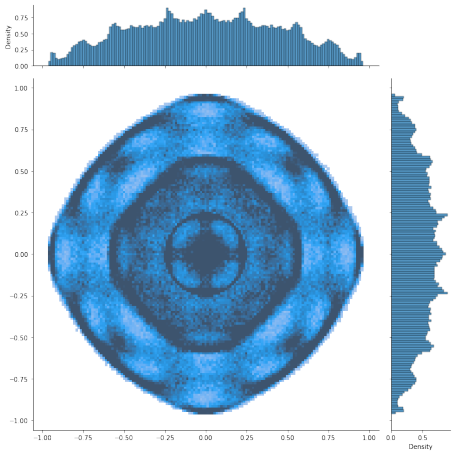
(4) μ_3, plane XY, $M = 100$, $N = 52$

Figure 3.18: Optimal transport for μ_2 and μ_3, $M = 10, 100$, $N = 27, 52$, $\beta_0 = 0$, $K = 10000$ and $\Delta t_0 = 10^{-4}$ In each graph, minimizers are represented as $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In order to better distinguish between areas of low and high particles density, plots are represented as 2D histograms.

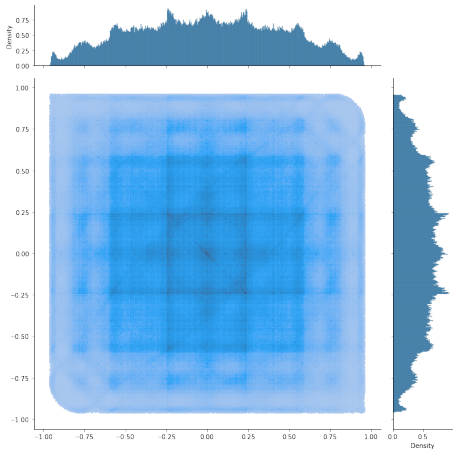
ization). Although the 1D marginal laws seem well approximated (Figures 3.19.2 and 3.19.3), planar and radial graphs (Figures 3.19.2 and 3.19.4) show that particles are concentrated on two spheres (of radius 0.6 and 1 respectively). Most of the transport takes place inside and between those two spheres.



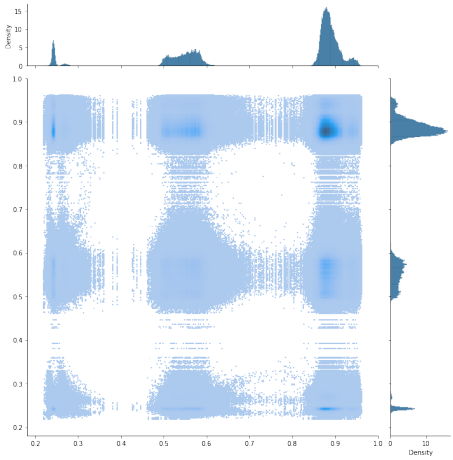
(1) Cost as a function of n



(2) plane XY



(3) X axis



(4) radial

Figure 3.19: Evolution of the cost as a function of the number of iteration n (Figure 3.19.1) and optimal transport with μ_4, $M = 100$, $N = 52$, $K = 10000$ $\beta_0 = 0$ and $\Delta t_0 = 10^{-4}$. In Figure 3.19.2 is showed $\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{x^k_{m,1},x^k_{m,2}}$. In Figure 3.19.3 is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{x^k_{m,1},x^k_{m',1}}$. In Figure 3.19.4 is showed $\frac{1}{M(M-1)K}\sum_{k=1}^{K}\sum_{m\neq m'=1}^{M}\delta_{|x^k_m|,|x^k_{m'}|}$, where $|x^k_m| = \sqrt{\sum_{i=1}^{3}(x^k_{m,i})^2}$. In order to better distinguish between areas of low and high particles density, plots are represented as 2D histograms.

3.6 Proof of Theorem 3.1

The aim of this section is to gather the proofs of our main theoretical results.

3.6.1 Tchakaloff's theorem

We present here a corollary of the so-called Tchakaloff theorem which is the backbone of our results concerning the theoretical properties of the MCOT particle problem. A general version of the Tchakaloff theorem has been proved by Bayer and Teichmann [24]. Theorem 3.4 is an immediate consequence of Tchakaloff's theorem, see Corollary 2 in [24].

Theorem 3.4. *Let π be a measure on $\mathbb{R}^d$ concentrated on a Borel set $A \in \mathcal{F}$, i.e. $\pi(\mathbb{R}^d \setminus A) = 0$. Let $N_0 \in \mathbb{N}^*$ and $\Lambda : \mathbb{R}^d \to \mathbb{R}^{N_0}$ a measurable Borel map. Assume that the first moments of $\Lambda\#\pi$ exist, i.e.*

$$\int_{\mathbb{R}^{N_0}} \|u\| \mathrm{d}\Lambda\#\pi(u) = \int_{\mathbb{R}^d} \|\Lambda(z)\| \mathrm{d}\pi(z) < \infty,$$

where $\|\cdot\|$ denotes the Euclidean norm of $\mathbb{R}^{N_0}$. Then, there exist an integer $1 \leq K \leq N_0$, points $z_1, ..., z_K \in A$ and weights $p_1, ..., p_K > 0$ such that

$$\forall 1 \leq i \leq N_0, \quad \int_{\mathbb{R}^d} \Lambda_i(z) \mathrm{d}\pi(z) = \sum_{k=1}^{K} p_k \Lambda_i(z_k),$$

where Λ_i denotes the i-th component of Λ.

We recall here that $\Lambda\#\pi$ is the push-forward of π through Λ, and is defined as $\Lambda\#\pi(A) = \pi(\Lambda^{-1}(A))$ for any Borel set $A \subset \mathbb{R}^{N_0}$.

Last, let us mention that Theorem 3.4 is a consequence of Caratheodory's theorem [287, Corollary 17.1.2] applied to $\int_{\mathbb{R}^{N_0}} u \mathrm{d}\Lambda\#\pi(u)$ which lies in the (convex) cone induced by $\mathrm{spt}(\Lambda\#\pi)$, the support of the measure $\Lambda\#\pi$.

3.6.2 Proof of Theorem 3.1

We denote here by $\mathcal{S}_K$ the set of permutations of the set $\{1, \cdots, K\}$.

Lemma 3.5. *Let $(W, Y) \in \mathcal{U}_K^N$ be such that there exists k' such that $w_{k'} = 0$. Then for any permutation $\sigma \in \mathcal{S}_K$, there exists a polygonal map $\psi : [0,1] \to \mathcal{U}_K^N$ such that $\psi(0) = (W, Y)$, $\psi(1) = (W^\sigma, Y^\sigma)$ and $\mathcal{I}(\psi(t))$ is constant, where $Y^\sigma := (X^{\sigma(k)})_{1 \leq k \leq K} \in ((\mathbb{R}^d)^M)^K$ and $W^\sigma := (w_{\sigma(k)})_{1 \leq k \leq K} \in (\mathbb{R}_+)^K$.*

Proof. For (W, Y) and (W', Y'), we will denote $[(W, Y), (W', Y')]$ the segment map $t \in [0,1] \mapsto [(1-t)W + tW', (1-t)Y + tY']$ and we will construct ψ as the concatenation of segments that are clearly in $\mathcal{U}_K^N$ and leaves $\mathcal{I}$ constant.

It is sufficient to prove this result for transpositions i.e. for σ such that there exist $i_1 < i_2$ such that $\sigma(i_1) = i_2$, $\sigma(i_2) = i_1$ and $\sigma(i) = i$ for $i \notin \{i_1, i_2\}$. We distinguish two cases.

- $k' \in \{i_1, i_2\}$, say $k' = i_2$. We then define $Y_1 = (X_1^k)_{1 \leq k \leq K}$ by $X_1^{k'} = X^{i_1}$ and $X_1^k = X^k$ for $k \neq k'$ and consider the segment $[(W, Y), (W, Y_1)]$. We then set $w_1^{k'} = w^{i_1}$, $w_1^{i_1} = 0$ and $w_1^k = w^k$ for $k \notin \{k', i_1\}$ (note that $W_1 = W^\sigma$) and consider the segment $[(W, Y_1), (W_1, Y_1)]$. Last, we define $Y_2 = (X_2^k)_{1 \leq k \leq K}$ as $X_2^{i_1} = X^{k'}$ and $X_2^k = X_1^k$ for $k \neq i_1$ (note that $Y_2 = Y^\sigma$) and consider the segment $[(W_1, Y_1), (W_1, Y_2)]$.

- $k' \notin \{i_1, i_2\}$. First, we define $Y_1 = (X_1^k)_{1\le k\le K}$ by $X_1^{k'} = X^{i_1}$ and $X_1^k = X^k$ for $k \neq k'$ and consider the segment $[(W, Y), (W, Y_1)]$. We then set $w_1^{k'} = w^{i_1}$, $w_1^{i_1} = 0$ and $w_1^k = w^k$ for $k \notin \{k', i_1\}$ and consider the segment $[(W, Y_1), (W_1, Y_1)]$. Then, we define $Y_2 = (X_2^k)_{1\le k\le K}$ as $X_2^{i_1} = X^{i_2}$ and $X_2^k = X_1^k$ for $k \neq i_1$, and consider the segment $[(W_1, Y_1), (W_1, Y_2)]$. We the set $w_2^{i_1} = w^{i_2}$, $w_2^{i_2} = 0$ and $w_2^k = w_1^k$ for $k \notin \{i_1, i_2\}$, and consider the segment $[(W_1, Y_2), (W_2, Y_2)]$. Now, we define $Y_3 = (X_3^k)_{1\le k\le K}$ by $X_3^{i_2} = X^{i_1}$, $X_3^k = X_2^k$ for $k \neq i_2$ and consider the segment $[(W_2, Y_2), (W_2, Y_3)]$. Then, we define $w_3^{i_2} = w_2^{k'} = w^{i_1}$, $w_3^{k'} = 0$ and $w_3^k = w_2^k$ for $k \notin \{i_2, k'\}$ (note that $W_3 = W^\sigma$) and consider the segment $[(W_2, Y_3), (W_3, Y_3)]$. Last, we set $Y_4 = (X_4^k)_{1\le k\le K}$ with $X_4^{k'} = X^{k'}$ and $X_4^k = X_3^k$ for $k \neq k'$ (note that $Y_4 = Y^\sigma$) and finally consider the segment $[(W_3, Y_3), (W_3, Y_4)]$, which gives the claim.

□

Proof. For $i = 0, 1$, let $W_i := (w_{k,i})_{1\le k\le K} \in \mathbb{R}_+^K$, $Y_i = (X_i^k)_{1\le k\le K} \subset (\mathbb{R}^d)^M$ and $\pi_i := \sum_{k=1}^K w_{k,i}\delta_{X_i^k} \in \mathcal{P}\left((\mathbb{R}^d)^M\right)$. Note that, for $i = 0, 1$, the support of π_i is included in the discrete set $\{X_i^k,\ 1 \le k \le K\}$.

For $i = 0, 1$, using Theorem 3.4 with $\pi = \pi_i$ and $\Lambda : (\mathbb{R}^d)^M \to \mathbb{R}^{N+3}$ the map defined such that, for all $X \in (\mathbb{R}^d)^M$,

$$\begin{aligned}&\Lambda_n(X) = \varphi_n(X), \quad \forall 1 \le n \le N,\\ &\Lambda_{N+1}(X) = 1, \quad \Lambda_{N+2}(X) = c(X) \quad \text{and} \quad \Lambda_{N+3}(X) = \vartheta(X),\end{aligned}$$

it holds that there exists a subset $J^i \subset \{1, \cdots, K\}$ such that $K_i := \#J^i \le N + 3$, and weights $(\widetilde{w}_j^i)_{j\in J^i} \subset \mathbb{R}_+$ such that

$$\forall 1 \le n \le N, \sum_{j\in J^i} \widetilde{w}_j^i \varphi_n(X_i^j) = \int_{(\mathbb{R}^d)^M} \varphi_n \, d\pi_i = \sum_{k=1}^K w_{k,i}\varphi_n(X_i^k) = \mu_n, \tag{3.49}$$

$$\sum_{j\in J^i} \widetilde{w}_j^i = \int_{(\mathbb{R}^d)^M} d\pi_i = \sum_{k=1}^K w_{k,i} = 1, \tag{3.50}$$

$$\sum_{j\in J^i} \widetilde{w}_j^i c(X_i^j) = \int_{(\mathbb{R}^d)^M} c \, d\pi_i = \sum_{k=1}^K w_{k,i}c(X_i^k) = \mathcal{I}(W_i, Y_i), \tag{3.51}$$

$$\sum_{j\in J^i} \widetilde{w}_j^i \vartheta(X_i^j) = \int_{(\mathbb{R}^d)^M} \vartheta \, d\pi_i = \sum_{k=1}^K w_{k,i}\vartheta(X_i^k) \le A. \tag{3.52}$$

Without loss of generality, by using Lemma 3.5, we can assume that $J^0 = [\![1, K_0]\!]$ where $K_0 \le N + 3$ and that $J^1 = [\![K - K_1 + 1, K]\!]$ where $K - K_1 + 1 \ge N + 4$.

We then define the weights $\widetilde{W}_0 := (\widetilde{w}_1^0, \cdots, \widetilde{w}_{K_0}^0, 0, \cdots, 0) \in \mathbb{R}_+^K$ and $\widetilde{W}_1 := (0, \cdots, 0, \widetilde{w}_{K-K_1+1}^1, \cdots, \widetilde{w}_K^1) \in \mathbb{R}_+^K$. Let us first define the applications

$$\psi_1 : \left[0, \frac{1}{5}\right] \ni t \mapsto \left(W_0 + 5t(\widetilde{W}_0 - W_0), Y_0\right)$$

and

$$\psi_5 : \left[\frac{4}{5}, 1\right] \ni t \mapsto \left(W_1 + 5(1-t)(\widetilde{W}_1 - W_1), Y_1\right)$$

so that $\psi_0(0) = (W_0, Y_0)$, $\psi_0(1/5) = (\widetilde{W}_0, Y_0)$, $\psi_1(1) = (W_1, Y_1)$, $\psi_1(4/5) = (\widetilde{W}_1, Y_1)$. Then, ψ_0 and ψ_1 are continuous applications and identities (3.49)-(3.50)-(3.51)-(3.52) implies that for all $t \in [0, 1/5]$ (respectively all $t \in [4/5, 1]$), $\psi_0(t) \in \mathcal{U}_N^K$ and $\mathcal{I}(\psi_0(t)) = \mathcal{I}(W_0, Y_0)$ (respectively $\psi_1(t) \in \mathcal{U}_N^K$ and $\mathcal{I}(\psi_1(t) = \mathcal{I}(W_1, Y_1))$.

We then define $\widetilde{Y} := \left(X_0^1, \cdots, X_0^{K_0}, 0, \cdots, 0, X_1^{K-K_1+1}, \cdots, X_1^K\right) \in ((\mathbb{R}^d)^M)^K$. We then introduce the continuous applications

$$\psi_2 : \left[\frac{1}{5}, \frac{2}{5}\right] \ni t \mapsto \left(\widetilde{W}_0, Y_0 + 5(t - 1/5)\widetilde{Y}\right)$$

and

$$\psi_4 : \left[\frac{3}{5}, \frac{4}{5}\right] \ni t \mapsto \left(\widetilde{W}_1, Y_1 + 5(4/5 - t)\widetilde{Y}\right).$$

It thus holds that $\psi_2(1/5) = (\widetilde{W}_0, Y_0)$ and $\psi_2(2/5) = (\widetilde{W}_0, \widetilde{Y})$. Similarly, $\psi_4(4/5) = (\widetilde{W}_1, Y_1)$ and $\psi_4(3/5) = (\widetilde{W}_1, \widetilde{Y})$. Let us point out here that, by the definition of $\widetilde{Y}$, for any $t \in \left[\frac{1}{5}, \frac{2}{5}\right]$, the K_0 first components of $\psi_2(t)$ are equal to $X_0^1, \cdots, X_0^{K_0}$. Thus, since $\widetilde{W}_0 := (\widetilde{w}_1^0, \cdots, \widetilde{w}_{K_0}^0, 0, \cdots, 0) \in \mathbb{R}_+^K$, this implies that for all $t \in \left[\frac{1}{5}, \frac{2}{5}\right]$, $\psi_2(t) \in \mathcal{U}_N^K$ and in addition,

$$\mathcal{I}(\psi_2(t)) = \mathcal{I}(\widetilde{W}_0, Y_0) = \mathcal{I}(\widetilde{W}_0, \widetilde{Y}) = \mathcal{I}(W_0, Y_0).$$

Similarly, for any $t \in \left[\frac{3}{5}, \frac{4}{5}\right]$, $\psi_4(t) \in \mathcal{U}_N^K$ and in addition,

$$\mathcal{I}(\psi_4(t)) = \mathcal{I}(\widetilde{W}_1, Y_1) = \mathcal{I}(\widetilde{W}_1, \widetilde{Y}) = \mathcal{I}(W_1, Y_1).$$

Notice that in particular, $\mathcal{I}$ remains constant along the paths in $\mathcal{U}_N^K$ given by the applications ψ_1, ψ_2, ψ_4 and ψ_5.

Last, we introduce the application

$$\psi_3 : \left[\frac{2}{5}, \frac{3}{5}\right] \ni t \mapsto \left(\widetilde{W}_0 + 5(t - 2/5)\widetilde{W}_1, \widetilde{Y}\right)$$

which is continuous and such that $\psi_3(2/5) = (\widetilde{W}_0, \widetilde{Y})$ and $\psi_3(3/5) = (\widetilde{W}_1, \widetilde{Y})$. Using similar arguments as above, it then holds that for all $t \in [2/5, 3/5]$, $\psi_3(t) \in \mathcal{P}_N^K$ and

$$\mathcal{I}(\psi_3(t)) = \mathcal{I}\left(\widetilde{W}_0, \widetilde{Y}\right) + 5(t - 2/5)\mathcal{I}\left(\widetilde{W}_1, \widetilde{Y}\right) = \mathcal{I}\left(W_0, Y_0\right) + 5(t - 2/5)\mathcal{I}\left(W_1, Y_1\right).$$

This implies that $\mathcal{I}$ monotonically varies along the path given by the application ψ_3.

We finally consider the application $\psi : [0, 1] \to (\mathbb{R}_+)^K \times \left((\mathbb{R}^d)^M\right)^K$ defined by

$$\forall t \in [0, 1], \quad \psi(t) = \begin{cases} \psi_1(t) & \text{if } t \in [0, 1/5], \\ \psi_2(t) & \text{if } t \in [1/5, 2/5], \\ \psi_3(t) & \text{if } t \in [2/5, 3/5], \\ \psi_4(t) & \text{if } t \in [3/5, 4/5], \\ \psi_5(t) & \text{if } t \in [4/5, 1]. \end{cases}$$

Gathering all the results we have obtained so far, it then holds that ψ is continuous, that for all $t \in [0, 1]$, $\psi(t) \in \mathcal{U}_N^K$ and that the application $\mathcal{I} \circ \psi$ is monotone. Hence the desired result. □

Appendix B

Appendix of Chapter 3

B.1 Moments computation in Normal case

Recall that for $N \sim \mathcal{N}(0,1)$,

$$\mathbb{E}\left(N^p\right) = \begin{cases} 0 & \text{if } p \text{ is odd} \\ \sigma^p(p-1) & \text{if } p \text{ is even.} \end{cases} \tag{B.1}$$

For a 3D multivariate Normal random variable

$$X \sim \mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2\end{pmatrix}\right),$$

$X = \begin{pmatrix}G_1\\G_2\\G_3\end{pmatrix}$ where $\mathrm{cov}(G_1,G_2) = \rho_{1,2}\sigma_1\sigma_2$, $\mathrm{cov}(G_1,G_3) = \rho_{1,3}\sigma_1\sigma_3$, $\mathrm{cov}(G_2,G_3) = \rho_{2,3}\sigma_2\sigma_3$ and G_1, G_2 and G_3 are not independent. Using $N_1, N_2, N_3 \sim \mathcal{N}(0,1)$ independent normal random variables, one has

$$G_1 = \sigma_1 N_1 \tag{B.2}$$
$$G_2 = \sigma_2\left(\rho_{1,2}N_1 + \sqrt{1-\rho_{1,2}^2}N_2\right) \tag{B.3}$$
$$G_3 = \sigma_3\left(\rho_{1,3}N_1 + \rho_{2,3}N_2 + \sqrt{1-\rho_{1,3}^2-\rho_{2,3}^2}N_3\right) \tag{B.4}$$

such that

$$\mathrm{Var}(G_2) = \sigma_2^2 \qquad \mathrm{Var}(G_3) = \sigma_3^2 \tag{B.5}$$
$$\mathbb{E}\left(G_1G_2\right) = \rho_{1,2}\sigma_1\sigma_2 \qquad \mathbb{E}\left(G_1G_3\right) = \rho_{1,3}\sigma_1\sigma_3 \tag{B.6}$$
$$\mathbb{E}\left(G_2G_3\right) = \rho_{2,3}\sigma_2\sigma_3 \tag{B.7}$$

Then the (k,l,m)-th moment can be computed as

$$\begin{aligned}
\mathbb{E}\left(G_1^k G_2^l G_3^m\right) &= \mathbb{E}\left(\sigma_1^k N_1^k \sigma_2^l \left(\rho_{1,2}N_1 + \sqrt{1-\rho_{1,2}^2}N_2\right)^l \right.\\
&\qquad \left. \times \sigma_3^m \left(\rho_{1,3}N_1 + \rho_{2,3}N_2 + \sqrt{1-\rho_{1,3}^2-\rho_{2,3}^2}N_3\right)^m\right)\\
&= \sigma_1^k\sigma_2^l\sigma_3^m \mathbb{E}\left(N_1^k \sum_{i=0}^{l}\binom{l}{i}\rho_{1,2}^i N_1^i \left(\sqrt{1-\rho_{1,2}^2}\right)^{l-i} N_2^{l-i}\right.\\
&\quad \left. \times \sum_{j=0}^{m}\binom{m}{j}\rho_{1,3}^j N_1^j \left(\rho_{2,3}N_2 + \sqrt{1-\rho_{1,3}^2-\rho_{2,3}^2}N_3\right)^{m-j}\right)\\
&= \sigma_1^k\sigma_2^l\sigma_3^m \sum_{i=0}^{l}\binom{l}{i}\rho_{1,2}^i \left(\sqrt{1-\rho_{1,2}^2}\right)^{l-i} \sum_{j=0}^{m}\binom{m}{j}\rho_{1,3}^j\\
&\quad \times \mathbb{E}\left(N_1^{k+i+j} N_2^{l-i} \sum_{h=0}^{m-j}\binom{m-j}{h}\rho_{2,3}^h N_2^h \left(\sqrt{1-\rho_{1,3}^2-\rho_{2,3}^2}\right)^{m-j-h} N_3^{m-j-h}\right)\\
&= \sigma_1^k\sigma_2^l\sigma_3^m \sum_{i=0}^{l}\binom{l}{i}\rho_{1,2}^i \left(\sqrt{1-\rho_{1,2}^2}\right)^{l-i} \sum_{j=0}^{m}\binom{m}{j}\rho_{1,3}^j \sum_{h=0}^{m-j}\binom{m-j}{h}\\
&\quad \times \rho_{2,3}^h \left(\sqrt{1-\rho_{1,3}^2-\rho_{2,3}^2}\right)^{m-j-h} \mathbb{E}\left(N_1^{k+i+j} N_2^{l-i+h} N_3^{m-j-h}\right)
\end{aligned} \tag{B.8}$$

Part II

Van der Waals interactions between two hydrogen atoms: The next orders

Chapter 4

Van der Waals interactions between two hydrogen atoms: The next orders

This chapter is an article written with Éric Cancès and L. Ridgway Scott and submitted to *Communications in Mathematical Sciences* [76].

Abstract

We extend a method (E. Cancès and L.R. Scott, SIAM J. Math. Anal., 50, 2018, 381–410) to compute more terms in the asymptotic expansion of the van der Waals attraction between two hydrogen atoms. These terms are obtained by solving a set of modified Slater–Kirkwood partial differential equations. The accuracy of the method is demonstrated by numerical simulations and comparison with other methods from the literature. It is also shown that the scattering states of the hydrogen atom, that are the states associated with the continuous spectrum of the Hamiltonian, have a major contribution to the C_6 coefficient of the van der Waals expansion.

4.1 Introduction

Van der Waals interactions, first introduced in 1873 to reproduce experimental results on simple gases [317], have proved to also play an essential role in complex systems in the condensed phase, such as biological molecules [21, 288] and 2D materials [153]. The quantum mechanical origin of the dispersive van der Waals interaction has been elucidated by London in the 1930s [238]. The rigorous mathematical foundations of the van der Waals interaction have been investigated in the pioneering work by Morgan and Simon [261], inspired by the one of Ahlrichs in [3], and later by Lieb and Thiring [230], followed by many authors (see in particular [13, 205] and references therein). For H_2^+, the expansion of the interaction energy as a function of the distance R between the nuclei is a diverging series – yet Borel summable, as predicted in [63] and later proved by [100, 114, 168]. Recent articles have studied this expansion for collection of atoms [14, 18], with terms up to $1/R^9$ [22], molecules [15, 16] and its differentiability [17].

In a recent paper [78], a new numerical approach was introduced to compute the leading order term $-C_6R^{-6}$ of the van der Waals interaction between hydrogen atoms separated by a distance R. Here we extend that approach to compute higher order terms $-C_nR^{-n}$, $n > 6$. The coefficients C_n have been computed by various methods. On the one hand, both [275] and [98] apparently failed to include key

components in the computation of C_{10}, computing only one component out of three that we derive here. On the other hand, our result differs by approximately 200% and agrees with [265]. One of the objects of this paper is to clarify this discrepancy.

The computation of the expansion coefficients can also be derived through techniques using polarizabilities [265] which is exact but might involve slightly different numerical computations than the perturbation method used here. In order to get the right values, one has to use a high enough order of perturbation theory. Computations using up to the second order [10, 88, 314] fail for C_{12}, C_{14} and C_{16} (with errors of approximately 1%, 5%, and 10%) for which computations up to the fourth order [257] are needed. The third order [323] is sufficient for C_{11}, C_{13} and C_{15}. Moreover, the polarizabilities method can be derived also for other atoms than hydrogen as well as for three-body interaction [88]. A comparison of the numerical results is explored in Section 4.3.1.

One can also compute the expansion coefficients using basis states as in [140]. However, this leads to a substantial error even for C_6. The discrepancy observed between the basis states method and the other methods can be interpreted as the missing contribution to the energy from the continuous spectrum.

The perturbation method of [309] is remarkable because, in the case of two hydrogen atoms, the problem splits, for any of the C_n terms, exactly into terms constituted of an angular factor and a function of two one-dimensional variables (the underlying problem is six-dimensional). The first term in this expansion has been examined in [78] and gave a value of C_6 agreeing with [265]. This article extends this analysis and allows computation of all C_n. The linearity and the nature of the angular parts allows treatment of these problems separately in a way analogous to the first term of the expansion. Although the partial differential equations (PDE) defining the functions of these two variables are not solvable in closed form, they are nevertheless easily solved by numerical techniques.

In Section 4.2, we present an extended and modified version of Slater and Kirkwood's derivation [309], in order to manipulate more suitable family of PDEs for theoretical analysis and numerical simulation. These modified Slater–Kirkwood PDEs are well posed at all orders and, when their unique solutions are multiplied by their respective angular factor, the resulting function, after summation of the terms, solves the triangular systems of six-dimensional PDEs originating from the Rayleigh–Schrödinger expansion. We finally check that the so-obtained perturbation series are asymptotic expansions of the ground state energy and wave function (after applying some "almost unitary" transform) of the hydrogen molecule in the dissociation limit. In Section 4.3, we use a Laguerre approximation [301, Section 7.3] to compute coefficients up to C_{19}, given that C_6 has been computed in [78]. Our approach also allows us to evaluate the respective contributions of the bound and scattering states of the Hamiltonian of the hydrogen atom to the C_6 coefficient of the van der Waals interaction. Numerical simulations show that the terms in the sum-over-states expansion coupling two bound states only contribute to about 60%. The mathematical proofs are gathered in Section 4.4. Lastly, some useful results on the multipolar expansion of the hydrogen molecule electrostatic potential in the dissociation limit and on the Wigner $(2n + 1)$ rule used in the computations are provided in the Appendix.

4.2 The hydrogen molecule in the dissociation limit

As usual in atomic and molecular physics, we work in atomic units: $\hbar = 1$ (reduced Planck constant), $e = 1$ (elementary charge), $m_e = 1$ (mass of the electron), $\epsilon_0 =$

$1/(4\pi)$ (dielectric permittivity of the vacuum). The length unit is the bohr (about 0.529 Ångstroms) and the energy unit is the hartree (about 4.36×10^{-18} Joules).

We study the Born-Oppenheimer approximation of a system of two hydrogen atoms, consisting of two classical point-like nuclei of charge 1 and two quantum electrons of mass 1 and charge -1. Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the positions in $\mathbb{R}^3$ of the two electrons, in a cartesian frame whose origin is the center of mass of the nuclei. We denote by $\mathbf{e}$ the unit vector pointing in the direction from one hydrogen atom to the other, and by R the distance between the two nuclei. We introduce the parameter $\epsilon = R^{-1}$ and derive expansions in ϵ of the ground state energy and wave function. Note that in [78], we use instead $\epsilon = R^{-1/3}$. The latter is well-suited to compute the lower-order coefficient C_6, but the change of variable $\epsilon = R^{-1}$ is more convenient to compute all the terms of the expansion.

Since the ground state of the hydrogen molecule is a singlet spin state [178], its wave function can be written as

$$\psi_\epsilon(\mathbf{r}_1, \mathbf{r}_2) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}, \tag{4.1}$$

where $\psi_\epsilon > 0$ is the L^2-normalized ground state of the spin-less six-dimensional Schrödinger equation

$$H_\epsilon \psi_\epsilon = \lambda_\epsilon \psi_\epsilon, \quad \|\psi_\epsilon\|_{L^2(\mathbb{R}^3 \times \mathbb{R}^3)} = 1, \tag{4.2}$$

where for $\epsilon > 0$, the Hamiltonian H_ϵ is the self-adjoint operator on $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ with domain $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ defined by

$$H_\epsilon = -\frac{1}{2}\Delta_{\mathbf{r}_1} - \frac{1}{2}\Delta_{\mathbf{r}_2} - \frac{1}{|\mathbf{r}_1 - (2\epsilon)^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_2 - (2\epsilon)^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_1 + (2\epsilon)^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_2 + (2\epsilon)^{-1}\mathbf{e}|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \epsilon,$$

where $\Delta_{\mathbf{r}_k}$ is the Laplace operator with respect to the variables $\mathbf{r}_k \in \mathbb{R}^3$. The first two terms of H_ϵ model the kinetic energy of the electrons, the next four terms the electrostatic attraction between nuclei and electrons, and the last two terms the electrostatic repulsion between, respectively, electrons and nuclei. The ground state of H_ϵ is symmetric ($\psi_\epsilon(\mathbf{r}_1, \mathbf{r}_2) = \psi_\epsilon(\mathbf{r}_2, \mathbf{r}_1)$) so that the wave function defined by (4.1) does satisfy the Pauli principle (the anti-symmetry is entirely carried by the spin component). It is well-known [14, 18, 78, 261] that

$$\lambda_\epsilon = -1 - C_6\epsilon^6 + o\left(\epsilon^6\right).$$

The computation of λ_ϵ (and ψ_ϵ) to higher order by a modified version of the Slater–Kirkwood approach, is the subject of this article.

4.2.1 Perturbation expansion

The first step is to make a change of coordinates. Introducing the translation operator

$$\tau_\epsilon f(\mathbf{r}_1, \mathbf{r}_2) = f(\mathbf{r}_1 + (2\epsilon)^{-1}\mathbf{e}, \mathbf{r}_2 - (2\epsilon)^{-1}\mathbf{e}) = f(\mathbf{r}_1 + \tfrac{1}{2}R\mathbf{e}, \mathbf{r}_2 - \tfrac{1}{2}R\mathbf{e}), \quad R = \epsilon^{-1},$$

the swapping operator $\mathcal{C}$ and the symmetrization operator $\mathcal{S}$ defined by

$$\mathcal{C}\phi(\mathbf{r}_1, \mathbf{r}_2) = \phi(\mathbf{r}_2, \mathbf{r}_1), \qquad \mathcal{S} = \frac{1}{\sqrt{2}}(\mathcal{I} + \mathcal{C}),$$

where $\mathcal{I}$ denotes the identity operator, as well as the "asymptotically unitary" operator

$$\mathcal{T}_\epsilon = \mathcal{S}\tau_\epsilon. \tag{4.3}$$

It is shown in [78] that

$$H_\epsilon \mathcal{T}_\epsilon = \mathcal{T}_\epsilon\big(H_0 + V_\epsilon\big), \tag{4.4}$$

where H_0 is the reference non-interacting Hamiltonian

$$H_0 = -\frac{1}{2}\Delta_{\mathbf{r}_1} - \frac{1}{|\mathbf{r}_1|} - \frac{1}{2}\Delta_{\mathbf{r}_2} - \frac{1}{|\mathbf{r}_2|},$$

and V_ϵ the correlation potential

$$V_\epsilon(\mathbf{r}_1, \mathbf{r}_2) = -\frac{1}{|\mathbf{r}_1 - \epsilon^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_2 + \epsilon^{-1}\mathbf{e}|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2 - \epsilon^{-1}\mathbf{e}|} + \epsilon. \tag{4.5}$$

The linear operator $\mathcal{T}_\epsilon$ is "asymptotically unitary" in the sense that for all $f, g \in L^2(\mathbb{R}^3 \times \mathbb{R}^3)$,

$$\big\langle \mathcal{T}_\epsilon f, \mathcal{T}_\epsilon g\big\rangle = \big\langle f, g\big\rangle + \big\langle \mathcal{C} f, \tau_{\epsilon/2} g\big\rangle \underset{\epsilon\to 0}{\longrightarrow} \big\langle f, g\big\rangle.$$

It follows from (4.4) that if (λ, ϕ) is a normalized eigenstate of $H_0 + V_\epsilon$, that is (λ, ϕ) satisfies

$$(H_0 + V_\epsilon)\phi = \lambda\phi, \quad \|\phi\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = 1,$$

then

$$H_\epsilon \mathcal{T}_\epsilon \phi = \lambda \mathcal{T}_\epsilon \phi.$$

In addition, we know from Zhislin's theorem [78, 325] that both H_ϵ and $H_0 + V_\epsilon$ have ground states, that their ground state eigenvalues are non-degenerate, and that their ground state wave functions are (up to replacing them by their opposites) positive everywhere in $\mathbb{R}^{3\times 3}$. Since $\mathcal{T}_\epsilon$ preserves positivity, we infer that H_ϵ and $H_0 + V_\epsilon$ share the same ground state eigenvalue λ_ϵ and that if ϕ_ϵ is the normalized positive ground state wave function of $H_0 + V_\epsilon$, then $\psi_\epsilon := \mathcal{T}_\epsilon\phi_\epsilon / \|\mathcal{T}_\epsilon\phi_\epsilon\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}$ is the normalized positive ground state wave function of H_ϵ.

The next step is to construct for $\epsilon > 0$ small enough the ground state $(\lambda_\epsilon, \phi_\epsilon)$ of $H_0 + V_\epsilon$ by the Rayleigh–Schrödinger perturbation method from the explicit ground state

$$\lambda_0 = -1, \qquad \phi_0(\mathbf{r}_1, \mathbf{r}_2) = \pi^{-1} e^{-(|\mathbf{r}_1| + |\mathbf{r}_2|)}, \tag{4.6}$$

of H_0. Using a multipolar expansion, we have

$$V_\epsilon(\mathbf{r}_1, \mathbf{r}_2) = \sum_{n=3}^{+\infty} \epsilon^n \mathcal{B}^{(n)}(\mathbf{r}_1, \mathbf{r}_2), \tag{4.7}$$

where homogeneous polynomial functions $\mathcal{B}^{(n)}$, $n \geq 3$ are specified below (see equation (4.14)), the convergence of the series being uniform on every compact subset of $\mathbb{R}^3 \times \mathbb{R}^3$. Assuming that λ_ϵ and ϕ_ϵ can be Taylor expanded as

$$\lambda_\epsilon = \lambda_0 - \sum_{n=1}^{+\infty} C_n \epsilon^n \quad \text{and} \quad \phi_\epsilon = \sum_{n=0}^{+\infty} \epsilon^n \phi_n, \qquad \text{(formal expansions)} \tag{4.8}$$

(we use the standard historical notation $-C_n$ instead of λ_n for the coefficients of the eigenvalue λ_ϵ) inserting these expansions in the equations $(H_0 + V_\epsilon)\phi_\epsilon = \lambda_\epsilon\phi_\epsilon$, $\|\phi_\epsilon\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = 1$, and identifying the terms of order n in ϵ, we obtain a triangular system of linear elliptic equations (Rayleigh–Schrödinger equations). The well-posedness of this system is given by the following lemma, whose proof is postponed until Section 4.4.2.

Lemma 4.1. *The triangular system*

$$\forall n \geq 1, \qquad (H_0 - \lambda_0)\phi_n = -\sum_{k=3}^{n} \mathcal{B}^{(k)}\phi_{n-k} - \sum_{k=1}^{n} C_k \phi_{n-k}, \tag{4.9}$$

$$\big\langle \phi_0, \phi_n \big\rangle = -\frac{1}{2}\sum_{k=1}^{n-1} \big\langle \phi_k, \phi_{n-k} \big\rangle, \tag{4.10}$$

where we use the convention $\sum_{k=m}^{n} \cdots = 0$ *if* $m > n$*, has a unique solution* $((C_n, \phi_n))_{n\in\mathbb{N}^*}$ *in* $(\mathbb{R}\times H^2(\mathbb{R}^3\times\mathbb{R}^3))^{\mathbb{N}^*}$*. In particular, we have* $(C_1, \phi_1) = (C_2, \phi_2) = 0$ *and* $C_3 = C_4 = C_5 = 0$*. In addition, the functions* ϕ_n *are real-valued.*

Note that $(C_1, \phi_1) = (C_2, \phi_2) = 0$ directly follows from the fact that the first non-vanishing term in the expansion (4.7) of V_ϵ is $\epsilon^3\mathcal{B}^{(3)}$. The formal expansions (4.8) are in fact asymptotic expansions as established in the following theorem, in which the second inequality of (4.12) has been proved to hold for H_2^+ in [261, Theorem 3.5]. Its proof is provided in Section 4.4.2.

Theorem 4.2. *Let* $\psi_\epsilon \in H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ *be the positive* $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$*-normalized ground state of* H_ϵ *and* λ_ϵ *the associated ground-state energy:*

$$H_\epsilon\psi_\epsilon = \lambda_\epsilon\psi_\epsilon, \quad \|\psi_\epsilon\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = 1, \quad \psi_\epsilon > 0 \text{ a.e. on } \mathbb{R}^3 \times \mathbb{R}^3. \tag{4.11}$$

Let (ϕ_0, λ_0) *be as in* (4.6)*,* $((C_n, \phi_n))_{n\in\mathbb{N}^*}$ *the unique solution of* (4.9) *in* $(\mathbb{R}\times H^2(\mathbb{R}^3\times\mathbb{R}^3))_{n\in\mathbb{N}^*}$*, and* $\mathcal{T}_\epsilon$ *the "almost unitary" symmetrization operator defined in* (4.3)*. Then, for all* $n \in \mathbb{N}$*, there exists* $\epsilon_n > 0$ *and* $K_n \in \mathbb{R}_+$ *such that for all* $0 < \epsilon \leq \epsilon_n$*,*

$$\big\|\psi_\epsilon - \psi_\epsilon^{(n)}\big\|_{H^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq K_n\epsilon^{n+1}, \quad \big|\lambda_\epsilon - \lambda_\epsilon^{(n)}\big| \leq K_n\epsilon^{n+1}, \quad \big|\lambda_\epsilon - \mu_\epsilon^{(n)}\big| \leq K_n\epsilon^{2(n+1)}, \tag{4.12}$$

where

$$\psi_\epsilon^{(n)} := \frac{\mathcal{T}_\epsilon\left(\phi_0 + \sum_{k=3}^{n}\epsilon^k\phi_k\right)}{\left\|\mathcal{T}_\epsilon\left(\phi_0 + \sum_{k=3}^{n}\epsilon^k\phi_k\right)\right\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}}, \quad \lambda_\epsilon^{(n)} := \lambda_0 - \sum_{k=6}^{n} C_k\epsilon^k, \quad \mu_\epsilon^{(n)} = \langle\psi_\epsilon^{(n)}|H_\epsilon|\psi_\epsilon^{(n)}\rangle.$$

Let us point out that in view of the last two bounds in (4.12), the series expansion of $\mu_\epsilon^{(n)}$ in ϵ up to order $(2n+1)$, which can be computed from the ϕ_k's for $0 \leq k \leq n$, is given by

$$\mu_\epsilon^{(n)} = \lambda_0 - \sum_{k=6}^{2n+1} C_k\epsilon^k + O(\epsilon^{2n+2}).$$

Therefore, the knowledge of the ϕ_k's up to order n allows one to compute all the C_k's up to order $(2n+1)$ (Wigner's $(2n+1)$ rule).

Remark 4.1 (van der Waals forces)**.** *It follows from the Hellmann-Feynman theorem that the van der Waals force* $\mathbf{F}_\epsilon$ *acting on the nucleus located at* $(2\epsilon)^{-1}\mathbf{e}$ *is given by*

$$\mathbf{F}_\epsilon = \int_{\mathbb{R}^3} \frac{(\mathbf{r} - (2\epsilon)^{-1}\mathbf{e})}{|\mathbf{r} - (2\epsilon)^{-1}\mathbf{e}|^3}\rho_\epsilon(\mathbf{r})\,d\mathbf{r} \text{ with } \rho_\epsilon(\mathbf{r}) = 2\int_{\mathbb{R}^3} |\psi_\epsilon(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' \text{ (electronic density).}$$

Introducing the approximation $\mathbf{F}_\epsilon^{(n)}$ *of* $\mathbf{F}_\epsilon$ *computed from* $\psi_\epsilon^{(n)}$ *as*

$$\mathbf{F}_\epsilon^{(n)} = \int_{\mathbb{R}^3} \frac{(\mathbf{r} - (2\epsilon)^{-1}\mathbf{e})}{|\mathbf{r} - (2\epsilon)^{-1}\mathbf{e}|^3}\rho_\epsilon^{(n)}(\mathbf{r})\,d\mathbf{r} \quad \text{with} \quad \rho_\epsilon^{(n)}(\mathbf{r}) = 2\int_{\mathbb{R}^3} |\psi_\epsilon^{(n)}(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}',$$

we obtain from the Cauchy-Schwarz inequality, the Hardy inequality in $\mathbb{R}^3$*, and* (4.12) *that*

$$|\mathbf{F}_\epsilon - \mathbf{F}_\epsilon^{(n)}| \leq 8\|\psi_\epsilon - \psi_\epsilon^{(n)}\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)}\|\psi_\epsilon + \psi_\epsilon^{(n)}\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)} \leq K_n'\epsilon^{n+1}$$

for some constant $K_n' \in \mathbb{R}_+$ *independent of* ϵ *and* ϵ *small enough. Since* $\mathbf{F}_\epsilon^{(n)}$ *can be Taylor expanded at* $\epsilon = 0$*, we obtain that the force* $\mathbf{F}_\epsilon$ *satisfies for all* $n \geq 6$

$$\mathbf{F}_\epsilon = -\left(\sum_{k=6}^{n} nC_n\epsilon^{n+1}\right)\mathbf{e} + O(\epsilon^{n+1}).$$

This extends the result $\mathbf{F}_\epsilon = -6C_6\epsilon^7\mathbf{e} + O(\epsilon^8)$ *proved in [17, Theorem 4] for* any *two atoms with non-degenerate ground states, to arbitrary order in the simple case of two hydrogen atoms.*

4.2.2 Computation of the perturbation series

The coefficients $\mathcal{B}^{(n)}$ are obtained by a classical multipolar expansion, detailed in Appendix C.1 for the sake of completeness. Using spherical coordinates in an orthonormal cartesian basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ of $\mathbb{R}^3$ for which $\mathbf{e}_3 = \mathbf{e}$, so that

$$\begin{aligned} \mathbf{r}_i = r_i\big(\sin(\theta_i)\cos(\phi_i)\mathbf{e}_1 + \sin(\theta_i)\sin(\phi_i)\mathbf{e}_2 + \cos(\theta_i)\mathbf{e}\big), \\ \cos(\theta_i) = \mathbf{r}_i\cdot\mathbf{e}, \quad \text{and} \quad r_i = |\mathbf{r}_i|, \quad i = 1, 2, \end{aligned} \tag{4.13}$$

it holds that for all $n \geq 3$,

$$\mathcal{B}^{(n)}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{(l_1,l_2)\in B_n} r_1^{l_1} r_2^{l_2} \sum_{-\min(l_1,l_2)\leq m\leq \min(l_1,l_2)} G_{\mathrm{c}}(l_1, l_2, m) Y_{l_1}^m(\theta_1, \phi_1) Y_{l_2}^{-m}(\theta_2, \phi_2), \tag{4.14}$$

$$= \sum_{(l_1,l_2)\in B_n} r_1^{l_1} r_2^{l_2} \sum_{-\min(l_1,l_2)\leq m\leq \min(l_1,l_2)} G_{\mathrm{r}}(l_1, l_2, m) \mathcal{Y}_{l_1}^m(\theta_1, \phi_1) \mathcal{Y}_{l_2}^m(\theta_2, \phi_2), \tag{4.15}$$

where $(Y_l^m)_{l\in\mathbb{N},\, m=-l,-l+1,\cdots,l-1,l}$ and $(\mathcal{Y}_l^m)_{l\in\mathbb{N},\, m=-l,-l+1,\cdots,l-1,l}$ are respectively the complex and real spherical harmonics , and where

$$B_n = \{(l_1, l_2) \,:\, l_1 + l_2 = n-1,\ l_1, l_2 \neq 0\} = \{(l, n-1-l) \,:\, 1 \leq l \leq n-2\}\,. \tag{4.16}$$

The coefficients $G_{\mathrm{c}}(l_1, l_2, m)$ and $G_{\mathrm{r}}(l_1, l_2, m)$ are respectively given by

$$G_{\mathrm{c}}(l_1, l_2, m) := (-1)^{l_2} \frac{4\pi(l_1 + l_2)!}{\big((2l_1+1)(2l_2+1)(l_1-m)!(l_1+m)!(l_2-m)!(l_2+m)!\big)^{1/2}}, \tag{4.17}$$

$$G_{\mathrm{r}}(l_1, l_2, m) := (-1)^m G_{\mathrm{c}}(l_1, l_2, m).$$

Both expansions (4.14) and (4.15) are useful: (4.14) will be used in the proof of Theorem 4.4 to establish formula (4.26), which has a simpler and more compact form in the complex spherical harmonics basis. On the other hand, (4.15) allows one to work with real-valued functions.

One of the main contributions of this article is to show that the functions ϕ_n, hence the real numbers λ_n, can be obtained by solving simple 2D linear elliptic boundary value problems on the quadrant

$$\Omega = \mathbb{R}_+^* \times \mathbb{R}_+^*,$$

extending the technique of Slater and Kirkwood for C_6 [309], modified in [78]. For each angular momentum quantum number $l \in \mathbb{N}$, we denote by

$$\kappa_l(r) = \frac{l(l+1)}{2r^2} - \frac{1}{r} - \frac{1}{2}\lambda_0 = \frac{l(l+1)}{2r^2} - \frac{1}{r} + \frac{1}{2}, \tag{4.18}$$

and we consider the boundary value problem: given $f \in L^2(\Omega)$

$$\begin{cases} \text{find } T \in H^1_0(\Omega) \text{ such that} \\ -\dfrac{1}{2}\Delta T(r_1, r_2) + \left(\kappa_{l_1}(r_1) + \kappa_{l_2}(r_2)\right) T = f(r_1, r_2) \quad \text{in } \mathcal{D}'(\Omega). \end{cases} \tag{4.19}$$

It follows from classical results on the radial operator $-\frac{1}{2}\frac{d^2}{dr^2} + \kappa_l$ on $L^2(0, +\infty)$ with form domain $H^1_0(0, +\infty)$ encountered in the study of the hydrogen atom (see Section 4.4.1 for details) that for all $l_1,\ l_2 \in \mathbb{N}$, $(l_1, l_2) \neq (0, 0)$, the problem (4.19) is well posed in $H^1_0(\Omega)$. For $l_1 = l_2 = 0$, this problem is well-posed in

$$\widetilde{H^1_0}(\Omega) = \left\{ v \in H^1_0(\Omega) \,:\, \int_\Omega v(r_1, r_2) e^{-r_1 - r_2}\, r_1 r_2 \, \mathrm{d}r_1 \mathrm{d}r_2 = 0 \right\},$$

provided that the compatibility condition

$$\int_\Omega f(r_1, r_2) e^{-r_1 - r_2}\, r_1 r_2 \, \mathrm{d}r_1 \mathrm{d}r_2 = 0 \tag{4.20}$$

is fulfilled. Problem (4.19) is useful to solve the Rayleigh–Schrödinger system (4.9)-(4.10) thanks to the following lemma, proved in Section 4.4.1. We denote by

$$\phi_0^\perp := \left\{ \psi \in L^2(\mathbb{R}^3 \times \mathbb{R}^3) \,:\, \langle \phi_0, \psi \rangle = 0 \right\}.$$

Note that the condition (4.20) is equivalent to $\left\langle \phi_0, \frac{f(r_1,r_2)}{r_1 r_2} \right\rangle = 0$.

Lemma 4.3. *Let $l_1, l_2 \in \mathbb{N}$, $m_1, m_2 \in \mathbb{Z}$ such that $-l_j \leq m_j \leq l_j$ for $j = 1, 2$, and $f \in L^2(\Omega)$. Consider the problem of finding $\psi \in H^2(\mathbb{R}^3 \times \mathbb{R}^3) \cap \phi_0^\perp$ solution to the equation*

$$(H_0 - \lambda_0)\psi = F \quad \textit{with} \quad F := \frac{f(r_1, r_2)}{r_1 r_2} Y_{l_1}^{m_1}(\theta_1, \phi_1) Y_{l_2}^{m_2}(\theta_2, \phi_2). \tag{4.21}$$

1. *If $(l_1, l_2) \neq (0, 0)$, then the unique solution to (4.21) in $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ is*

$$\psi = \frac{T(r_1, r_2)}{r_1 r_2} Y_{l_1}^{m_1}(\theta_1, \phi_1) Y_{l_2}^{m_2}(\theta_2, \phi_2), \tag{4.22}$$

   where T is the unique solution to (4.19) in $H^1_0(\Omega)$;

2. *If $(l_1, l_2) = (0, 0)$, and if the compatibility condition (4.20) is satisfied, then the unique solution to (4.21) in $H^2(\mathbb{R}^3 \times \mathbb{R}^3) \cap \phi_0^\perp$ is*

$$\psi = \frac{1}{4\pi} \frac{T(r_1, r_2)}{r_1 r_2},$$

   where T is the unique solution to (4.19) in $\widetilde{H}^1_0(\Omega)$.

In addition, if f decays exponentially at infinity, then so does T, hence ψ, in the following sense: for all $0 \leq \alpha < \sqrt{3/8}$, there exists a constant $C_\alpha \in \mathbb{R}_+$ such that for all $\eta > \alpha$, $l_1, l_2 \in \mathbb{N}$, $m_1, m_2 \in \mathbb{Z}$ such that $-l_j \leq m_j \leq l_j$ for $j = 1, 2$, and all $f \in L^2(\Omega)$

$$\|e^{\alpha(r_1+r_2)}T\|_{H^1(\Omega)} \leq C_\alpha \|e^{\eta(r_1+r_2)}f\|_{L^2(\Omega)}, \tag{4.23}$$

$$\|e^{\alpha(|\mathbf{r}_1|+|\mathbf{r}_2|)}\psi\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C_\alpha \|e^{\eta(|\mathbf{r}_1|+|\mathbf{r}_2|)}F\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}, \tag{4.24}$$

$$\|e^{\alpha(|\mathbf{r}_1|+|\mathbf{r}_2|)}\psi\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)} \leq C_\alpha(1 + 4l_1(l_1+1) + 4l_2(l_2+1))^{1/2} \|e^{\eta(|\mathbf{r}_1|+|\mathbf{r}_2|)}F\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}. \tag{4.25}$$

Lastly, if f is real-valued, then so is T.

The properties of the functions ϕ_n upon which our numerical method is based, are collected in the following theorem, proved in Section 4.4.2.

Theorem 4.4. *Let $((C_n, \phi_n))_{n\in\mathbb{N}^*}$ be the unique solution in $(\mathbb{R} \times H^2(\mathbb{R}^3 \times \mathbb{R}^3))_{n\in\mathbb{N}^*}$ to the Rayleigh–Schrödinger system (4.9). Then, $\phi_1 = \phi_2 = 0$, $C_n = 0$ for $1 \leq n \leq 5$ and for each $n \geq 3$, there exists a positive integer N_n such that*

$$\phi_n = \sum_{(l_1,l_2)\in\mathcal{L}_n} \frac{1}{r_1 r_2} \left(\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} T^{(n)}_{(l_1,l_2,m)}(r_1, r_2) Y_{l_1}^m(\theta_1, \phi_1) Y_{l_2}^{-m}(\theta_2, \phi_2) \right), \tag{4.26}$$

where $\mathcal{L}_n$ is a finite subset of $\mathbb{N}^2$ with cardinality $N_n < \infty$, where $T^{(n)}_{(l_1,l_2,m)}$ is the unique solution to (4.19) in $H^1(\Omega)$ (or in $\widetilde{H}^1(\Omega)$ if $l_1 = l_2 = 0$) for $f = f^{(n)}_{(l_1,l_2,m)}$, where $f^{(n)}_{(l_1,l_2,m)}$ is a real-valued function that can be computed recursively from the $T^{(n')}_{(l_1',l_2',m')}$'s, for $n' < n$ (as in (4.71)). Moreover, there exists $\alpha_n > 0$ such that

$$\|e^{\alpha_n(r_1+r_2)}T^{(n)}_{(l_1,l_2)}\|_{H^1(\Omega)} < \infty, \tag{4.27}$$

$$\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_n\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)} < \infty. \tag{4.28}$$

The number $N_n = |\mathcal{L}_n|$ (number of terms in the expansion) for $6 \leq n \leq 9$ are displayed in Table 4.1, whose construction rules are given in the proof of Theorem 4.4 (see Section 4.4.2). For $3 \leq n \leq 5$, $\mathcal{L}_n = B_n$, where the latter set is defined in (4.16), and $N_n = |B_n| = n-2$. For general n, $B_n \subset \mathcal{L}_n$. For $n \geq 6$, additional terms appear, as indicated in Table 4.1.

n	N_n	pairs of angular momentum quantum numbers (l_1, l_2) in $\mathcal{L}_n \backslash B_n$
6	8	(0,2;0,2)
7	13	(0,2;1,3), (1,3;0,2)
8	18	(0,2;0,2,4), (1,3;1,3), (0,2,4;0,2)
9	27	(0,2;1,3,5), (1,3;0,2,4), (1,3,5;0,2), (0,2,4;1,3), (1,3;1,3)

Table 4.1: Additional spherical harmonics appearing in each ϕ_n for $6 \leq n \leq 9$. N_n is the number of terms in the spherical harmonics expansion (4.26). The condensed notation $(l_1, l_1'; l_2, l_2')$ (resp. $(l_1, l_1'; l_2, l_2', l_2'')$ or $(l_1, l_1', l_1''; l_2, l_2')$) stands for the four (resp. six) pairs (l_1, l_2), (l_1', l_2), (l_1, l_2'), etc.

Table 4.1 can be read using the following rule: for a given n, if (l_1, l_2) appears in the corresponding row of the table, then there may exist m such that

$Y_{l_1}^m(\theta_1,\phi_1)Y_{l_2}^{-m}(\theta_2,\phi_2)$ might appear with a non-zero function $T^{(n)}_{(l_1,l_2,m)}$ in the spherical harmonics expansion (4.26) of ϕ_n. Conversely, if a given (l_1,l_2) does not appear in the table, then $\big\langle \phi_n, \frac{v(r_1,r_2)}{r_1r_2}Y_{l_1}^{m_1}(\theta_1,\phi_1)Y_{l_2}^{m_2}(\theta_2,\phi_2)\big\rangle = 0$, for all m_1, m_2 and all $v \in L^2(\Omega)$. The relative complexity of Table 4.1 is due to fact the first term in the right-hand side of (4.9) is a sum of bilinear terms in $\mathcal{B}^{(k)}$ and ϕ_{n-k}. The angular parts of both $\mathcal{B}^{(k)}$ and ϕ_{n-k} are finite linear combinations of angular basis functions $Y_{l_1}^m \otimes Y_{l_2}^{-m}$. When multiplied, they give rise to a still finite but longer linear combination of $Y_{l_1}^m \otimes Y_{l_2}^{-m}$'s (see (4.69)). By contrast, the corresponding table for the $\mathcal{B}^{(n)}$'s is quite simple, since all the rows have the same structure: for all $n \geq 3$, we have

$$n \quad | \quad n-2 \quad | \quad (k, n-k) \quad \text{for } 1 \leq k \leq n-2. \tag{4.29}$$

From $(\phi_k)_{0\leq k\leq n}$, we can obtain the coefficients λ_j up to $j = 2n+1$ using Wigner's $(2n+1)$ rule. Another, more direct, way to compute recursively the λ_n's is to take the inner product of ϕ_0 with each side of (4.9) and use the fact that $\langle \phi_0, (H_0-\lambda_0)\phi_n\rangle = \langle (H_0-\lambda_0)\phi_0, \phi_n\rangle = 0$. Since $(C_1,\phi_1) = (C_2,\phi_2) = 0$, we thus obtain

$$C_n = -\sum_{k=3}^{n-3} \big\langle \phi_0, \mathcal{B}^{(k)}\phi_{n-k}\big\rangle - \sum_{k=3}^{n-3} C_k \big\langle \phi_0, \phi_{n-k}\big\rangle, \tag{4.30}$$

where we use the convention $\sum_{k=m}^n \ldots = 0$ if $m > n$. It follows that $C_3 = C_4 = C_5 = 0$.

Using (4.14), (4.26) and the orthonormality properties of the complex spherical harmonics, the terms $\big\langle \psi_0, \mathcal{B}^{(k)}\phi_{n-k}\big\rangle$ in (4.30) can be written as

$$\begin{aligned}
\big\langle \phi_0, \mathcal{B}^{(k)}\phi_{n-k}\big\rangle &= \big\langle \mathcal{B}^{(k)}\phi_0, \phi_{n-k}\big\rangle \\
&= \Big\langle \sum_{(l_1,l_2)\in B_k} r_1^{l_1} r_2^{l_2} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} G_{\rm c}(l_1,l_2,m) Y_{l_1}^m(\theta_1,\phi_1) Y_{l_2}^{-m}(\theta_2,\phi_2)\pi^{-1}e^{-(r_1+r_2)}, \\
&\qquad \sum_{(l_1',l_2')\in\mathcal{L}_{n-k}} \frac{1}{r_1r_2} \sum_{m'=-\min(l_1',l_2')}^{\min(l_1',l_2')} T^{(n-k)}_{(l_1',l_2',m')}(r_1,r_2) Y_{l_1'}^{m'}(\theta_1,\phi_1) Y_{l_2'}^{-m'}(\theta_2,\phi_2)\Big\rangle \\
&= -\sum_{(l_1,l_2)\in\mathcal{L}_{n-k}\cap B_k} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} \beta^{(n-k)}_{(l_1,l_2,m)} t^{(n-k)}_{l_1,l_2,m},
\end{aligned} \tag{4.31}$$

where

$$\beta^{(n)}_{(l_1,l_2,m)} := -\pi^{-1} G_{\rm c}(l_1,l_2,m) \tag{4.32}$$

$$t^{(n)}_{(l_1,l_2,m)} := \int_\Omega r_1^{l_1+1} r_2^{l_2+1} e^{-(r_1+r_2)} T^{(n)}_{(l_1,l_2,m)}(r_1,r_2)\, dr_1\, dr_2, \tag{4.33}$$

with the convention that $t^{(n)}_{(l_1,l_2,m)} = 0$ if $(l_1,l_2) \notin \mathcal{L}_n$. In view of Table 4.1, we see in particular that since the sum in (4.31) is empty

$$\big\langle \phi_0, \mathcal{B}^{(k)}\phi_n\big\rangle = 0 \quad \forall\, k, n = 3,4,5,\ k \neq n, \tag{4.34}$$

and that many other vanish, e.g.

$$\big\langle \phi_0, \mathcal{B}^{(3)}\phi_6\big\rangle = 0, \quad \big\langle \phi_0, \mathcal{B}^{(4)}\phi_5\big\rangle = 0, \quad \big\langle \phi_0, \mathcal{B}^{(5)}\phi_4\big\rangle = 0, \quad \big\langle \phi_0, \mathcal{B}^{(6)}\phi_3\big\rangle = 0. \tag{4.35}$$

Additional pairs k, n can be examined by comparing the sets B_k and $\mathcal{L}_{n-k}$.

Furthermore, if the chosen numerical method to solve the boundary value problem (4.19) giving the radial function $T^{n-k}_{l_1',l_2',m'}$ is a Galerkin method using as basis functions of the approximation space tensor products of 1D Laguerre functions (that are, polynomials in r times e^{-r}), then the computation of $t^{(n)}_{l_1,l_2,m}$ can be done explicitly, at least for the approximate solution [301, Section 7.3]. Using the fact that

$$\phi_0 = 4e^{-(r_1+r_2)}Y_0^0(\theta_1,\phi_1)Y_0^0(\theta_2,\phi_2), \tag{4.36}$$

we then have

$$\begin{aligned}\left\langle \phi_0, \phi_n \right\rangle = \Bigg\langle & 4e^{-(r_1+r_2)}Y_0^0(\theta_1,\phi_1)Y_0^0(\theta_2,\phi_2), \\ & \sum_{(l_1',l_2')\in\mathcal{L}_n} \frac{1}{r_1 r_2} \sum_{m'=-\min(l_1',l_2')}^{\min(l_1',l_2')} T^{(n)}_{(l_1',l_2',m')}(r_1,r_2)Y_{l_1'}^{m'}(\theta_1,\phi_1)Y_{l_2'}^{-m'}(\theta_2,\phi_2)\Bigg\rangle \\ & = 4t^{(n)}_{(0,0,0)}.\end{aligned} \tag{4.37}$$

As a consequence, $\left\langle \phi_0, \phi_n \right\rangle = 0$ if $(0,0) \notin \mathcal{L}_n$, so that in particular

$$\left\langle \phi_0, \phi_3 \right\rangle = \left\langle \phi_0, \phi_4 \right\rangle = \left\langle \phi_0, \phi_5 \right\rangle = 0. \tag{4.38}$$

Then, C_n can be computed from (4.30) as

$$C_n = \sum_{k=3}^{n-3} \sum_{\substack{(l_1,l_2)\in\mathcal{L}_{n-k} \\ l_1+l_2=k-1 \\ l_1,l_2\neq 0}} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} \beta^{(n-k)}_{(l_1,l_2,m)} t^{(n-k)}_{(l_1,l_2,m)} - 4\sum_{k=6}^{n-3} C_k t^{(n-k)}_{(0,0,0)}. \tag{4.39}$$

4.2.3 Practical computation of the lowest order terms

We detail in this section the practical computation of ϕ_3 (already done in [78]), ϕ_4 and ϕ_5, as well as C_n for $n \leq 11$. Recall that $\phi_1 = \phi_2 = 0$, and $C_n = 0$ for $n \leq 5$.

Computation of ϕ_3. We have

$$\mathcal{B}^{(3)} = r_1 r_2 \left(\sum_{m=-1}^{1} G_{\mathrm{c}}(1,1,m))Y_1^m(\theta_1,\phi_1)Y_1^{-m}(\theta_2,\phi_2) \right), \tag{4.40}$$

$$(H_0 - \lambda_0)\phi_3 = -\mathcal{B}^{(3)}\phi_0, \tag{4.41}$$

$$\left\langle \phi_0, \phi_3 \right\rangle = 0, \tag{4.42}$$

with $G_{\mathrm{c}}(1,1,m) = -\frac{\pi}{3}(8-4|m|)$ and therefore

$$(H_0 - \lambda_0)\phi_3 = -r_1 r_2 e^{-(r_1+r_2)} \left(\sum_{m=-1}^{1} \pi^{-1} G_{\mathrm{c}}(1,1,m) Y_1^m(\theta_1,\phi_1)Y_1^{-m}(\theta_2,\phi_2) \right),$$

$$\left\langle \phi_0, \phi_3 \right\rangle = 0.$$

As a consequence, using Lemma 4.3, it holds that $\mathcal{L}_3 = \{(1,1)\}$,

$$\phi_3 = \frac{T^{(3)}_{(1,1)}(r_1,r_2)}{r_1 r_2} \left(\sum_{m=-1}^{1} \alpha^{(3)}_{(1,1,m)} Y_1^m(\theta_1,\phi_1)Y_1^{-m}(\theta_2,\phi_2) \right), \tag{4.43}$$

where $\alpha^{(3)}_{(1,1,m)} = -\pi^{-1}G_{\mathrm{c}}(1,1,m) = -\frac{1}{3}(8-4|m|)$ and where $T^{(3)}_{(1,1)} \in H^1_0(\Omega)$ can be numerically computed by solving the 2D boundary value problem

$$-\frac{1}{2}\Delta T^{(3)}_{(1,1)} + \left(\kappa_1(r_1) + \kappa_1(r_2)\right) T^{(3)}_{(1,1)} = r_1^2 r_2^2 e^{-(r_1+r_2)} \quad \text{in } \Omega$$

with homogeneous Dirichlet boundary conditions.

Computation of ϕ_4. To compute the next order, we first expand $\mathcal{B}^{(4)}$ as

$$\mathcal{B}^{(4)} = r_1 r_2^2 \sum_{m=-1}^{1} G_{\mathrm{c}}(1,2,m) Y_1^m(\theta_1,\phi_1) Y_2^{-m}(\theta_2,\phi_2) + r_1^2 r_2 \sum_{m=-2}^{2} G_{\mathrm{c}}(2,1,m) Y_1^m(\theta_1,\phi_1) Y_2^{-m}(\theta_2,\phi_2),$$

with $G_{\mathrm{c}}(1,2,1) = G_{\mathrm{c}}(1,2,-1) = 4\pi/\sqrt{5}$, $G_{\mathrm{c}}(1,2,0) = 4\pi\sqrt{3}/\sqrt{5}$, $G_{\mathrm{c}}(2,1,m) = -G_{\mathrm{c}}(1,2,m)$. From (4.9)-(4.10), we get

$$\begin{aligned} (H_0 - \lambda_0)\phi_4 &= -\mathcal{B}^{(3)}\phi_1 - \mathcal{B}^{(4)}\phi_0, \\ \left\langle \phi_0, \phi_4 \right\rangle &= 0, \end{aligned}$$

since $\phi_1 = \phi_2 = 0$ and $C_k = 0$ for $1 \le k \le 5$. We therefore have $\mathcal{L}_4 = \{(1,2),(2,1)\}$ and

$$\phi_4 = \frac{T^{(4)}_{(1,2)}(r_1,r_2)}{r_1 r_2} \sum_{m=-1}^{1} \alpha^{(4)}_{(1,2,m)} Y_1^m(\theta_1,\phi_1) Y_2^{-m}(\theta_2,\phi_2) + \frac{T^{(4)}_{(2,1)}(r_1,r_2)}{r_1 r_2} \sum_{m=-1}^{1} \alpha^{(4)}_{(2,1,m)} Y_2^m(\theta_1,\phi_1) Y_1^{-m}(\theta_2,\phi_2),$$

where $\alpha^{(4)}_{(l_1,l_2,m)} = -\pi^{-1}G_{\mathrm{c}}(l_1,l_2,m)$, $T^{(4)}_{(2,1)} \in H^1_0(\Omega)$ solves

$$-\frac{1}{2}\Delta_2 T^{(4)}_{(2,1)}(r_1,r_2) + \left(\kappa_2(r_1) + \kappa_1(r_2)\right) T^{(4)}_{(2,1)} = r_1^3 r_2^2 e^{-r_1-r_2} \quad \text{in } \Omega, \tag{4.44}$$

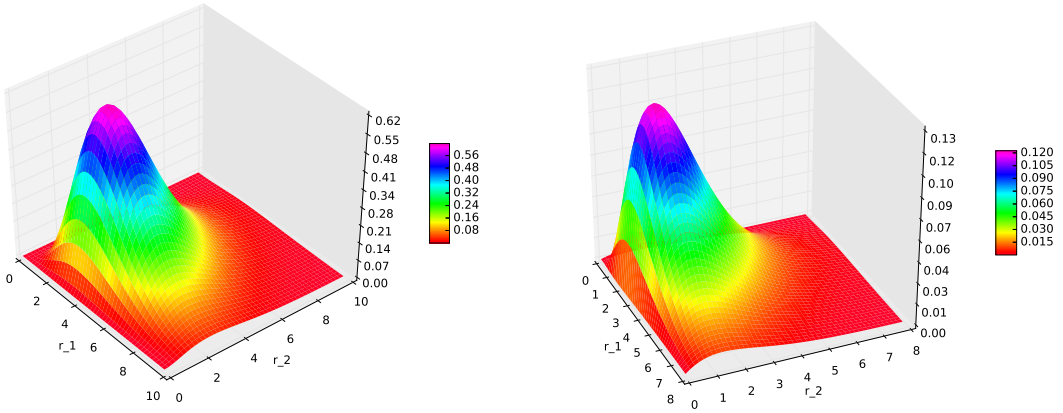
and $T^{(4)}_{(1,2)}(r_1,r_2) = T^{(4)}_{(2,1)}(r_2,r_1)$. A representation of $T^{(4)}_{(2,1)}$ can be seen in Figure 4.1.

Computation of ϕ_5. We have

$$\mathcal{B}^{(5)} = r_1 r_2^3 \sum_{m=-1}^{1} G_{\mathrm{c}}(1,3,m) Y_1^m(\theta_1,\phi_1) Y_3^{-m}(\theta_2,\phi_2) + r_1^2 r_2^2 \sum_{m=-2}^{2} G_{\mathrm{c}}(2,2,m) Y_2^m(\theta_1,\phi_1) Y_2^{-m}(\theta_2,\phi_2) + r_1^3 r_2^1 \sum_{m=-1}^{1} G_{\mathrm{c}}(3,1,m) Y_3^m(\theta_1,\phi_1) Y_1^{-m}(\theta_2,\phi_2),$$

and

$$\begin{aligned} (H_0 - \lambda_0)\phi_5 &= -\mathcal{B}^{(5)}\phi_0, \\ \left\langle \phi_0, \phi_5 \right\rangle &= 0, \end{aligned}$$



(1) (2)

Figure 4.1: Shape of $T^{(4)}_{(2,1)}$ 1 and $T^{(4)}_{(2,1)}(r_1, r_2)/(r_1 r_2)$ 2, using the Laguerre function approximation scheme [301, Section 7.3].

since $\phi_1 = \phi_2 = 0$ and $C_k = 0$ for $1 \le k \le 5$. We thus have $\mathcal{L}_5 = \{(1,3),(2,2),(3,1)\}$ and

$$\begin{aligned}\psi^{(5)} = \frac{T^{(5)}_{(1,3)}(r_1,r_2)}{r_1 r_2} \sum_{m=-1}^{1} \alpha^{(5)}_{(1,3,m)} Y_1^m(\theta_1,\phi_1) Y_3^{-m}(\theta_2,\phi_2) \\ + \frac{T^{(5)}_{(2,2)}(r_1,r_2)}{r_1 r_2} \sum_{m=-2}^{2} \alpha^{(5)}_{(2,2,m)} Y_2^m(\theta_1,\phi_1) Y_2^{-m}(\theta_2,\phi_2) \\ + \frac{T^{(5)}_{(3,1)}(r_1,r_2)}{r_1 r_2} \sum_{m=-1}^{1} \alpha^{(5)}_{(3,1,m)} Y_3^m(\theta_1,\phi_1) Y_1^{-m}(\theta_2,\phi_2), \quad (4.45)\end{aligned}$$

where $\alpha^{(5)}_{(l_1,l_2,m)} = -\pi^{-1} G_c(l_1,l_2,m)$, $T^{(5)}_{(1,3)} \in H^1_0(\Omega)$ solves

$$-\frac{1}{2}\Delta_2 T^{(5)}_{(1,3)}(r_1,r_2) + (\kappa_1(r_1) + \kappa_3(r_2))\, T^{(5)}_{(1,3)} = r_1^2 r_2^4 e^{-(r_1+r_2)}, \quad (4.46)$$

$T^{(5)}_{(2,3)} \in H^1_0(\Omega)$ solves

$$-\frac{1}{2}\Delta_2 T^{(5)}_{(2,2)}(r_1,r_2) + (\kappa_2(r_1) + \kappa_2(r_2))\, T^{(5)}_{(2,2)} = r_1^3 r_2^3 e^{-(r_1+r_2)}, \quad (4.47)$$

and $T^{(5)}_{(3,1)}(r_1,r_2) = T^{(5)}_{(1,3)}(r_2,r_1)$.

Computation of λ_n for $6 \le n \le 11$. Let us define for $n = 3,4,5$,

$$\begin{aligned}\beta^{(n)}_{(l_1,l_2)} &:= -\pi^{-1} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} \alpha^{(n)}_{(l_1,l_2,m)} G_c(l_1,l_2,m) \\ t^{(n)}_{(l_1,l_2)} &:= \int_\Omega r_1^{l_1+1} r_2^{l_2+1} e^{-(r_1+r_2)} T^{(n)}_{(l_1,l_2)}(r_1,r_2)\, dr_1\, dr_2,\end{aligned}$$

with the convention that $\beta^{(n)}_{(l_1,l_2)} = t^{(n)}_{(l_1,l_2)} = 0$ if $(l_1,l_2) \notin \mathcal{L}_n$. From (4.30) and the fact that $C_n = 0$ for $3 \le n \le 5$, we obtain, using (4.31), (4.38), (4.39), Table 4.1,

and the symmetries of the coefficients $\beta^{(n)}_{(l_1,l_2)}$ and $t^{(n)}_{(l_1,l_2)}$,

$$
\begin{aligned}
C_6 &= -\big\langle \phi_0, \mathcal{B}^{(3)}\phi_3 \big\rangle = \beta^{(3)}_{(1,1)} t^{(3)}_{(1,1)}, \qquad (4.48)\\
C_7 &= -\big\langle \phi_0, \mathcal{B}^{(3)}\phi_4 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(4)}\phi_3 \big\rangle = 0,\\
C_8 &= -\big\langle \phi_0, \mathcal{B}^{(3)}\phi_5 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(4)}\phi_4 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(5)}\phi_3 \big\rangle = -\big\langle \phi_0, \mathcal{B}^{(4)}\phi_4 \big\rangle\\
&= \beta^{(4)}_{(1,2)} t^{(4)}_{(1,2)} + \beta^{(4)}_{(2,1)} t^{(4)}_{(2,1)} = 2\beta^{(4)}_{(1,2)} t^{(4)}_{(1,2)},\\
C_9 &= -\big\langle \phi_0, \mathcal{B}^{(3)}\phi_6 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(4)}\phi_5 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(5)}\phi_4 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(6)}\phi_3 \big\rangle - C_6 \big\langle \phi_0, \phi_3 \big\rangle = 0,\\
C_{10} &= -\sum_{k=3}^{7} \big\langle \phi_0, \mathcal{B}^{(k)}\phi_{10-k} \big\rangle - \sum_{k=6}^{7} C_k \big\langle \phi_0, \phi_{10-k} \big\rangle = -\big\langle \phi_0, \mathcal{B}^{(5)}\phi_5 \big\rangle\\
&= \beta^{(5)}_{(1,3)} t^{(5)}_{(1,3)} + \beta^{(5)}_{(2,2)} t^{(5)}_{(2,2)} + \beta^{(5)}_{(3,1)} t^{(5)}_{(3,1)} = 2\beta^{(5)}_{(1,3)} t^{(5)}_{(1,3)} + \beta^{(5)}_{(2,2)} t^{(5)}_{(2,2)}, \qquad (4.49)\\
C_{11} &= -\sum_{k=3}^{8} \big\langle \phi_0, \mathcal{B}^{(k)}\phi_{11-k} \big\rangle - \sum_{k=6}^{8} C_k \big\langle \phi_0, \phi_{11-k} \big\rangle = -\big\langle \phi_0, \mathcal{B}^{(4)}\phi_7 \big\rangle - \big\langle \phi_0, \mathcal{B}^{(5)}\phi_6 \big\rangle\\
&= \sum_{m=-1}^{1} \Big[\beta^{(7)}_{(1,2,m)} t^{(7)}_{(1,2,m)} + \beta^{(7)}_{(2,1,m)} t^{(7)}_{(2,1,m)} \Big] + \sum_{m=-2}^{2} \beta^{(6)}_{(2,2,m)} t^{(6)}_{(2,2,m)}. \qquad (4.50)
\end{aligned}
$$

As $\alpha^{(n)}_{(l_1,l_2,m)} = -\pi^{-1} G_c(l_1,l_2,m)$ for $n = 3,4,5$, $(l_1,l_2) \in \mathcal{L}_n$ and $-\min(l_1,l_2) \le m \le \min(l_1,l_2)$, we obtain, using (4.17), that

$$
\big(\alpha^{(n)}_{(l_1,l_2,m)}\big)^2 = \frac{16\,((l_1+l_2)!)^2}{(2l_1+1)(2l_2+1)(l_1-m)!(l_1+m)!(l_2-m)!(l_2+m)!},
$$

and therefore

$$
\begin{aligned}
\beta^{(3)}_{(1,1)} &= \sum_{m=-1}^{1} (\alpha^{(3)}_{(1,1,m)})^2 = \frac{16}{9} + \frac{64}{9} + \frac{16}{9} = \frac{32}{3},\\
\beta^{(4)}_{(1,2)} &= \beta^{(4)}_{(2,1)} = \sum_{m=-1}^{1} (\alpha^{(4)}_{(1,2,m)})^2 = \frac{16}{5} + 3 \times \frac{16}{5} + \frac{16}{5} = 16,\\
\beta^{(5)}_{(1,3)} &= \beta^{(5)}_{(3,1)} = \sum_{m=-1}^{1} (\alpha^{(5)}_{(1,3,m)})^2 = \frac{64}{3}, \qquad \beta^{(5)}_{(2,2)} = \sum_{m=-2}^{2} (\alpha^{(5)}_{(2,2,m)})^2 = \frac{224}{5},
\end{aligned}
$$

so that

$$
C_6 = \frac{32}{3} t^{(3)}_{(1,1)}, \quad C_7 = 0, \quad C_8 = 32 t^{(4)}_{(1,2)}, \quad C_9 = 0, \qquad C_{10} = \frac{128}{3} t^{(5)}_{(1,3)} + \frac{224}{5} t^{(5)}_{(2,2)}. \qquad (4.51)
$$

It is optimal to use (4.51) to compute C_6, C_8, C_{10} since only ϕ_n is needed to compute C_{2n}. On the other hand, computing C_{11} using (4.50) requires computing ϕ_6 and ϕ_7, and it is therefore preferable to use Wigner's $(2n+1)$ rule that allows computing C_{11} from ϕ_3, ϕ_4 and ϕ_5.

Computation of higher-order terms. For $n \ge 6$, the right-hand side of (4.9) contains terms of the form $\mathcal{B}^{(k)}\phi_{n-k}$ with $k \ge 3$ and $n-k \ge 1$. The computation of ϕ_n therefore requires solving 2D boundary value problems of the form

$$
-\frac{1}{2}\Delta T + (\kappa_{l_1}(r_1) + \kappa_{l_2}(r_2))\, T = r_1^{l_1'} r_2^{l_2'} T^{(n-k)}_{(l_1'',l_2'',m'')}
$$

for some $(l_1, l_2) \in \mathcal{L}_n$, $l_1' + l_2' = k - 1$, $(l_1'', l_2'') \in \mathcal{L}_{n-k}$ and $-\min(l_1'', l_2'') \leq m'' \leq \min(l_1'', l_2'')$. The right-hand side of this equation is not explicit, but the above equation can nevertheless be solved numerically since $T^{(n-k)}_{(l_1'', l_2'', m'')}$ has been previously computed numerically during the calculation of ϕ_{n-k}. An analogous procedure was used by Morgan and Simon for H_2^+ and can be found in the Appendix of [261].

4.3 Numerical results

4.3.1 Comparison between different approaches

The following tables contain the results of the approximated values of the C_n coefficients computed by Ovsiannikov and Mitroy [265], by Choy [98], by Pauling and Beach [275], and by the techniques described in this paper. The latter consist in solving recursively the Modified Slater–Kirkwood boundary value problems of type (4.9) using a Galerkin scheme in finite-dimensional approximation spaces constructed from tensor products of 1D Laguerre functions with degrees lower of equal to k. With basic double-precision floating-point arithmetics, the latter approach is numerical stable up to $k = 11$ and provides results with excellent precision (relative error lower than 10^{-9}). It is well-known that the conditioning of spectral methods for PDEs using orthogonal polynomial spaces grows exponentially. However, in the present case, the entries of the Galerkin matrix are square roots of rational numbers so that arbitrary precision can be obtained using symbolic computation. The method of Choy [98] is based on the Slater–Kirkwood algorithm [309], whereas the method of Pauling and Beach [275] is different. Although Slater and Kirkwood are referenced in [275], Pauling and Beach were motivated by a method of S. C. Wang [322].

Method	C_6	C_8	C_{10}	C_{11}
[275]	6.49903	124.399	1135.21	
[98]	6.4990267	124.3990835	1135.2140398	
This work	6.49902670540 [78]	124.399083	3285.82841	-3474.89803
[265]	6.499026705406	124.3990835836	3285.828414967	-3474.898037882

Table 4.2: Comparison of the coefficients C_6 to C_{11} between various papers and the basis states method and our method based on numerical solutions of boundary value problems of type (4.19) in tensor products of Laguerre functions up to degree 11 (for which round-off error is suitably controlled). These results agree at least to 9 digits with the results in [88, 257, 265, 314, 323].

The discrepancy between the Choy [98] and Pauling–Beach [275] results (who agree to the digits given) and the other methods for C_{10} has the following origin. According to (4.49), we have

$$C_{10} = 2\beta^{(5)}_{(1,3)} t^{(5)}_{(1,3)} + \beta^{(5)}_{(2,2)} t^{(5)}_{(2,2)}.$$

It appears that Choy in [98], who also was guided by [309], only computed the second term

$$\beta^{(5)}_{(2,2)} t^{(5)}_{(2,2)} = 1135.214\ldots \tag{4.52}$$

Method	C_{12}	C_{13}	C_{14}	C_{15}
This work	122727.608	-326986.924	6361736.04	-28395580.6
[265]	122727.6087007	-326986.9240441	6361736.045092	-28395580.6

Table 4.3: Comparison of the C_n coefficients C_{12} to C_{15} between [265] and our method based on numerical solutions of boundary value problems of type (4.19) in tensor products of Laguerre functions up to degree 11 (for which round-off error is suitably controlled). These results agree at least to 9 digits with the results in [257, 265, 323] for C_{13} and C_{15} and [257, 265] for C_{12} and C_{14}.

Method	C_{16}	$C_{17} \times 10^{-9}$	$C_{18} \times 10^{-10}$	$C_{19} \times 10^{-11}$
This work	441205192	-2.73928165	3.93524773	-3.07082459
[265]	441205192.2739	-2.739281653140	3.93524773346	-3.07082459389

Table 4.4: Comparison of the C_n coefficients C_{16} to C_{19} between [265] and our method based on numerical solutions of boundary value problems of type (4.19) in tensor products of Laguerre functions up to degree 11 (for which round-off error is suitably controlled). These results agree at least to 9 digits with the results in [257, 265].

4.3.2 Role of continuous spectra in sum-over-state formulae

It follows from (4.41), (4.42) and (4.48) that the leading coefficient C_6 of the van der Waals expansion can be written as

$$C_6 = \langle \mathcal{B}^{(3)}\phi_0, (H_0 - \lambda_0)^{-1}_{\phi_0^\perp} \mathcal{B}^{(3)}\phi_0 \rangle,$$

where $(H_0-\lambda_0)^{-1}_{\phi_0^\perp}$ is the inverse of the restriction to $H_0-\lambda_0$ to the invariant subspace $\phi_0^\perp$ (which is well-defined since λ_0 is a non-degenerate eigenvalue of the self-adjoint operator H_0. This expression is sometimes wrongly rewritten as a sum-over-state formula

$$C_6 = \sum_j \frac{|\langle \psi_j, \mathcal{B}^{(3)}\psi_0\rangle|^2}{E_j - E_0} \quad \text{(wrong)}, \tag{4.53}$$

with $\psi_0 := \phi_0$, $E_0 := \lambda_0 = -1$, where the ψ_j's form an orthonormal family of excited states of H_0 associated with the eigenvalues E_j. This is not possible because H_0 has a non-empty continuous spectrum. Using (4.53) with a sum running over the excited states of H_0 (and omitting an integral over the scattering states of H_0) leads to an error that we are going to estimate. We have

$$C_6' := \sum_j \frac{|\langle \psi_j, \mathcal{B}^{(3)}\psi_0\rangle|^2}{E_j - E_0} = -\langle \mathcal{B}^{(3)}\phi_0, \phi_{3,\mathrm{pp}} \rangle,$$

where $\phi_{3,\mathrm{pp}}$ is the projection of ϕ_3 on the Hilbert space spanned by the eigenfunctions of H_0. Recall that the eigenvalues and associated eigenfunctions of the hydrogen atom Hamiltonian $h_0 := -\frac{1}{2}\Delta - \frac{1}{|\mathbf{r}|}$, which is a self-adjoint operator on $L^2(\mathbb{R}^3)$, are of the form

$$\varepsilon_n = -\frac{1}{2n^2}, \quad \psi_{n,l,m}(\mathbf{r}) = \varphi_{n,l}(r) Y_l^m(\theta,\phi), \quad n \in \mathbb{N}^*, \quad 0 \le l \le n-1, \quad -l \le m \le l, \tag{4.54}$$

with

$$\varphi_{n,1} = \sqrt{\left(\frac{2}{n}\right)^3 \frac{(n-2)!}{2n(n+1)!}} \left(\frac{2r}{n}\right) L_{n-2}^{(3)}\left(\frac{2r}{n}\right) e^{-r/n}, \tag{4.55}$$

where the associated Laguerre polynomials of the second type $L_n^{(m)}$, $n, m \in \mathbb{N}$, are defined from the Laguerre polynomial L_n and are given by

$$L_n^{(m)}(x) = (-1)^m \frac{\mathrm{d}^m L_{n+m}}{\mathrm{d}x^m}(x) = \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!} \binom{n+m}{n-k} (-x)^k. \tag{4.56}$$

The eigenvalues and associated eigenfunctions of H_0 are therefore given by

$$\mathcal{E}_{n_1,n_2} = \varepsilon_{n_1} + \varepsilon_{n_2} = -\frac{1}{2n_1^2} - \frac{1}{2n_2^2}, \qquad \Psi_{n_1,l_1,m_1;n_2,l_2,m_2} = \psi_{n_1,l_1,m_1} \otimes \psi_{n_2,l_2,m_2},$$

for $n_j \in \mathbb{N}^*$, $0 \leq l_j \leq n_j - 1$, $-l_j \leq m_j \leq l_j$. Note that $\phi_0 = \Psi_{1,0,0;1,0,0}$. We therefore have

$$C_6' = \sum_{(n_1,n_2)\in(\mathbb{N}^*\times\mathbb{N}^*)\setminus\{(1,1)\}} \sum_{l_1=0}^{n_1-1} \sum_{l_2=0}^{n_2-1} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \frac{|\langle \Psi_{n_1,l_1,m_1;n_2,l_2,m_2}, \mathcal{B}^{(3)}\psi_0\rangle|^2}{\varepsilon_{n_1} + \varepsilon_{n_2} + 1},$$

Using (4.40) and the $L^2(\mathbb{S}^2)$-orthonormality of the spherical harmonics, we get

$$\langle \Psi_{n_1,l_1,m_1;n_2,l_2,m_2}, \mathcal{B}^{(3)}\psi_0\rangle = \pi^{-1} S_{n_1} S_{n_2} \sum_{m=-1}^{1} G_{\mathrm{c}}(1,1,m)\delta_{l_1,1}\delta_{l_2,1}\delta_{m,m_1}\delta_{-m,m_2},$$

where

$$S_n := \int_0^{+\infty} r^3 e^{-r} \phi_{n,1}(r)\,\mathrm{d}r = 8n^3 \frac{(n-1)^{n-3}}{(n+1)^{n+3}} \sqrt{\frac{(n+1)!}{(n-2)!}}. \tag{4.57}$$

The latter expression is derived in Appendix C.3. We finally obtain

$$C_6' = \pi^{-2} \sum_{m=-1}^{1} |G_{\mathrm{c}}(1,1,m)|^2 \sum_{n_1,n_2\geq 2} \frac{S_{n_1}^2 S_{n_2}^2}{1 - \frac{1}{2n_1^2} - \frac{1}{2n_2^2}} = \frac{32}{3} \sum_{n_1,n_2\geq 2} \frac{S_{n_1}^2 S_{n_2}^2}{1 - \frac{1}{2n_1^2} - \frac{1}{2n_2^2}}. \tag{4.58}$$

Summing up the terms of the above series for $n_1, n_2 \leq 300$ (note that $S_n \sim_{n\to\infty} \frac{8}{e^2 n^{3/2}}$), we obtain the approximate value

$$C_6' \simeq 3.923$$

which shows that the continuous spectrum plays a major role in the sum-over-state evaluation of the C_6 coefficient of the hydrogen molecule (recall that $C_6 \simeq 6.499$).

4.4 Proofs

We now establish the results stated above, starting from Lemma 4.3.

4.4.1 Proof of Lemma 4.3

Recall that the Hydrogen atom Hamiltonian $h_0 = -\frac{1}{2}\Delta - \frac{1}{|\mathbf{r}|}$ introduced in the previous section is a self-adjoint operator on $L^2(\mathbb{R}^3)$ with domain $H^2(\mathbb{R}^3)$, and that its ground state is non-degenerate:

$$h_0\psi_{1,0,0} = -\frac{1}{2}\psi_{1,0,0} \quad \text{with} \quad \psi_{1,0,0} = \varphi_{1,0}(r)Y_0^0(\theta,\phi) = \pi^{-1/2}e^{-r},\ \|\psi_{1,0,0}\|_{L^2(\mathbb{R}^3)} = 1.$$

Since $H_0 = h_0 \otimes \mathbb{1}_{L^2(\mathbb{R}^3)} + \mathbb{1}_{L^2(\mathbb{R}^3)} \otimes h_0$, H_0 is a self-adjoint operator on $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ with domain $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ and it also has a non-degenerate ground state

$$H_0\phi_0 = \lambda_0\phi_0 \text{ with } \phi_0 = \psi_{1,0,0} \otimes \psi_{1,0,0} = \pi^{-1}e^{-(r_1+r_2)}, \ \|\phi_0\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = 1 \text{ and } \lambda_0 = -1.$$

Given $(\alpha, F) \in \mathbb{R} \times L^2(\mathbb{R}^3 \times \mathbb{R}^3)$, the problem consisting of seeking $(\mu, \Psi) \in \mathbb{R} \times H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ such that

$$(H_0 - \lambda_0)\Psi = F - \mu\phi_0, \quad \langle \phi_0, \Psi \rangle = \alpha, \tag{4.59}$$

is well-posed and its unique solution is given

$$\Psi = (H_0 - \lambda_0)|^{-1}_{\phi_0^\perp} \Pi_{\phi_0^\perp} F + \alpha\phi_0, \quad \mu = \langle \phi_0, F \rangle,$$

where $(H_0 - \lambda_0)|^{-1}_{\phi_0^\perp}$ is the inverse of $H_0 - \lambda_0$ on the invariant subspace $\phi_0^\perp$ and $\Pi_{\phi_0^\perp} F := F - \langle \phi_0, F \rangle \phi_0$ the orthogonal projection of F on $\phi_0^\perp$. Consider the unitary map

$$\mathcal{U} : L^2(\Omega) \otimes L^2(\mathbb{S}^2) \otimes L^2(\mathbb{S}^2) \to L^2(\mathbb{R}^3 \times \mathbb{R}^3) \equiv L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$$

induced by the spherical coordinates defined for all $f \in L^2(\Omega)$, $l_1, l_2 \in \mathbb{N}$, $-l_j \leq m_j \leq l_j$ by

$$(\mathcal{U}(f \otimes s_1 \otimes s_2))(\mathbf{r}_1, \mathbf{r}_2) = \frac{f(|\mathbf{r}_1|, |\mathbf{r}_2|)}{|\mathbf{r}_1|\,|\mathbf{r}_2|}\, s_1\left(\frac{\mathbf{r}_1}{|\mathbf{r}_1|}\right)\, s_2\left(\frac{\mathbf{r}_2}{|\mathbf{r}_2|}\right).$$

Since $(Y_l^m)_{l\in\mathbb{N}, -l\leq m\leq l}$ is an orthonormal basis of $L^2(\mathbb{S}^2)$, we have

$$L^2(\Omega) \otimes L^2(\mathbb{S}^2) \otimes L^2(\mathbb{S}^2) = \bigoplus_{l_1,l_2\in\mathbb{N}} \bigoplus_{m_1=-l_1}^{l_1} \bigoplus_{m_2=-l_2}^{l_2} \mathcal{H}_{l_1,l_2}^{m_1,m_2}$$

where $\mathcal{H}_{l_1,l_2}^{m_1,m_2} := L^2(\Omega) \otimes \mathbb{C}Y_{l_1}^{m_1} \otimes \mathbb{C}Y_{l_2}^{m_2}$. It follows from classical results for Schrödinger operators on $L^2(\mathbb{R}^3)$ with central potentials (see e.g. [286, Section XIII.3.B]) that each $\mathcal{H}_{l_1,l_2}^{m_1,m_2}$ is an invariant subspace for $\mathcal{U}^* H_0 \mathcal{U}$ and that

$$\mathcal{U}^* H_0 \mathcal{U}|_{\mathcal{H}_{l_1,l_2}^{m_1,m_2}} = H_{l_1,l_2} \otimes \mathbb{1}_{\mathbb{C}Y_{l_1}^{m_1}} \otimes \mathbb{1}_{\mathbb{C}Y_{l_2}^{m_2}},$$

where the expression of H_{l_1,l_2} can be derived by adapted the arguments in [78, Section 3], that we do not detail here for the sake of brevity: H_{l_1,l_2} is the self-adjoint operator on $L^2(\Omega)$ with form domain $H_1^0(\Omega)$ defined by

$$H_{l_1,l_2} = -\frac{1}{2}\Delta + \kappa_{l_1}(r_1) + \kappa_{l_2}(r_2) + \lambda_0. \tag{4.60}$$

Note that the operator H_{l_1,l_2} on $L^2(\Omega) \equiv L^2(0,+\infty) \otimes L^2(0,+\infty)$ can itself be decomposed as

$$H_{l_1,l_2} = h_{l_1} \otimes \mathbb{1}_{L^2(0,+\infty)} + \mathbb{1}_{L^2(0,+\infty)} \otimes h_{l_2} \geq -\frac{1}{2(l_1+1)^2} - \frac{1}{2(l_2+1)^2},$$

where for each $l \in \mathbb{N}$, h_l is the self-adjoint operator on $L^2(0,+\infty)$ with form domain $H_0^1(0,+\infty)$ defined by

$$h_l := -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{1}{r} = -\frac{1}{2}\frac{d^2}{dr^2} + \kappa_l - \frac{1}{2}.$$

This well-known operator allows one to construct the bound-states of hydrogen atom with orbital quantum number l. It satisfies $h_l \geq -\frac{1}{2(l+1)^2}$ and its ground state eigenvalue $-\frac{1}{2(l+1)^2}$ is non-degenerate. It follows from this bound that

$$H_{l_1,l_2} - \lambda_0 = H_{l_1,l_2} + 1 \geq \frac{3}{8} \quad \text{for all } (l_1, l_2) \in \mathbb{N}^2 \setminus \{(0,0\}. \tag{4.61}$$

Choosing $\alpha = 0$ in (4.59) amounts to enforcing that the solution Ψ is in $\phi_0^\perp$. Taking $\alpha = 0$ and $F = \frac{f(r_1,r_2)}{r_1 r_2} Y_{l_1}^{m_1}(\theta_1,\phi_1) Y_{l_2}^{m_2}(\theta_2,\phi_2) = \mathcal{U}(f \otimes Y_{l_1}^{m_1} \otimes Y_{l_2}^{m_2})$, with $f \in L^2(\Omega)$, it follows that (4.21) has a unique solution in $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ if and only if $\mu = \langle \phi_0, F \rangle = 0$, that is

$$\delta_{(l_1,l_2)=(0,0)} \int_\Omega f(r_1,r_2) e^{-(r_1+r_2)} r_1 r_2 \, dr_1 \, dr_2 = 0,$$

in which case the solution is given by $\Psi = \mathcal{U}(T \otimes Y_{l_1}^{m_1} \otimes Y_{l_2}^{m_2})$ where

$$\begin{aligned} T &:= (H_{l_1,l_2} - \lambda_0)^{-1} f && \text{if } (l_1,l_2) \neq (0,0), \\ T &:= (H_{0,0} - \lambda_0)|^{-1}_{(r_1 r_2 e^{-(r_1+r_2)})^\perp} f && \text{if } (l_1,l_2) = (0,0). \end{aligned}$$

We therefore have

$$\psi = \frac{T(r_1,r_2)}{r_1 r_2} Y_{l_1}^{m_1}(\theta_1,\phi_1) Y_{l_2}^{m_2}(\theta_2,\phi_2),$$

where T is the unique solution to (4.19) in $H_0^1(\Omega)$ if $(l_1,l_2) \neq (0,0)$ and T is the unique solution to (4.19) in $\widetilde{H}_0^1(\Omega) = H_0^1(\Omega) \cap (r_1 r_2 e^{-(r_1+r_2)})^\perp$ if $(l_1,l_2) = 0$.

The fact that if f decays exponentially at infinity, then so does T, hence ψ, is a consequence of the following result, whose proof follows the same lines as in [78, Section 3.3] where this result is established for the special case when $(l_1,l_2) = (1,1)$ and $f = r_1^2 r_2^2 e^{-(r_1+r_2)}$.

Lemma 4.5. *If the function f of* (4.19) *decays exponentially at infinity at a rate $\eta > 0$, in the sense that*

$$\|e^{\eta(r_1+r_2)} f\|_{L^2(\Omega)} < \infty, \tag{4.62}$$

then the unique solution T of (4.19) *also decays exponentially at infinity. More precisely, for all $0 \leq \alpha < \sqrt{3/8}$, there exists a constant $C_\alpha \in \mathbb{R}_+$ such that for all $\eta > \alpha$ and all $f \in L^2(\Omega)$ satisfying* (4.62)*, it holds*

$$\|e^{\alpha(r_1+r_2)} T\|_{H^1(\Omega)} \leq C_\alpha \|e^{\eta(r_1+r_2)} f\|_{L^2(\Omega)}. \tag{4.63}$$

Proof. We limit ourselves to the case when $(l_1,l_2) \neq (0,0)$. The special case $(l_1,l_2) = (0,0)$ can be dealt with similarly, by replacing $H_0^1(\Omega)$ by $\widetilde{H}_0^1(\Omega)$. Let a be the continuous bilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$ associated with the positive self-adjoint operator $H_{l_1,l_2} - \lambda_0$:

$$\forall u, v \in H_0^1(\Omega), \quad a(u,v) = \frac{1}{2} \int_\Omega \nabla u \cdot \nabla v + \int_\Omega (\kappa_{l_1}(r_1) + \kappa_{l_2}(r_2)) u(r_1,r_2) v(r_1,r_2) \, dr_1 \, dr_2.$$

Recall that the continuity of a can be shown directly (without using the fact that $H_0^1(\Omega)$ is the form domain of $H_{l_1 l_2}$) as a straightforward consequence of the one-dimensional Hardy inequality

$$\forall g \in H_0^1(0,+\infty), \quad \int_0^\infty (g(r)/r)^2 \mathrm{d}r \leq 4 \int_0^\infty g'(r)^2 \, \mathrm{d}r. \tag{4.64}$$

It follows from (4.61) that $a \geq \frac{3}{8}$ (in the sense of quadratic forms on $L^2(\Omega)$). For $0 \leq \alpha < \sqrt{3/8}$, we introduce the continuous bilinear form a_α on $H^1_0(\Omega) \times H^1_0(\Omega)$ defined by

$$\forall u, v \in H^1_0(\Omega), \quad a_\alpha(u,v) = a(u,v) - \int_\Omega \alpha u(\mathbf{r}) \left(\frac{\partial v}{\partial r_1}(\mathbf{r}) + \frac{\partial v}{\partial r_2}(\mathbf{r}) \right) \mathrm{d}\mathbf{r} - \int_\Omega \alpha^2 u(\mathbf{r}) v(\mathbf{r}) \mathrm{d}\mathbf{r},$$

for which

$$\forall v \in H^1_0(\Omega), \quad a_\alpha(v,v) = a(v,v) - \alpha^2 \|v\|^2_{L^2(\Omega)} \geq \underbrace{\left(\frac{3}{8} - \alpha^2 \right)}_{>0} \|v\|^2_{L^2(\Omega)}.$$

Using either the fact that $\kappa_l(r) \geq \frac{1}{4}$ (for $l \geq 1$) or the Hardy inequality (4.64) (for $l = 0$), we also have

$$\forall v \in H^1_0(\Omega), \quad a_\alpha(v,v) = a(v,v) - \alpha^2 \|v\|^2_{L^2(\Omega)} \geq \frac{1}{4} \int_\Omega |\nabla v|^2 - 2\|v\|^2_{L^2}.$$

Since $a \geq \frac{3}{8}$ and $a_\alpha \geq \left(\frac{3}{8} - \alpha^2 \right) > 0$, the above bound implies that a and a_α are both continuous and coercive on $H^1_0(\Omega)$. The function $T \in H^1_0(\Omega)$ solution to (4.19) is also the unique solution to the variational equation

$$\forall w \in H^1_0(\Omega), \quad a(T,w) = \int_\Omega f w.$$

Proceeding as in [78, Section 3.3], we obtain that for all $u \in H^1_0(\Omega)$ such that $e^{\alpha(r_1+r_2)} u \in H^1_0(\Omega)$ and $w \in C^\infty_{\mathrm{c}}(\Omega)$, we have

$$a_\alpha(e^{\alpha(r_1+r_2)} u, w) = a(u, e^{\alpha(r_1+r_2)} w). \tag{4.65}$$

Consider now $f \in L^2(\Omega)$ satisfying (4.62) for some $\eta > \alpha$. The function $e^{\alpha(r_1+r_2)} f$ is in $L^2(\Omega)$, so that the problem of finding $v \in H^1(\Omega)$ such that

$$\forall w \in H^1_0(\Omega), \quad a_\alpha(v,w) = \int_\Omega e^{\alpha(r_1+r_2)} f w$$

has a unique solution v, satisfying $\|v\|_{H^1(\Omega)} \leq C_\alpha \|e^{\alpha(r_1+r_2)} f\|_{L^2(\Omega)} \leq C_\alpha \|e^{\eta(r_1+r_2)} f\|_{L^2(\Omega)}$, where $C_\alpha \geq 1$ is the ratio between the continuity constant and the coercivity constant of a_α. Let $u = e^{-\alpha(r_1+r_2)} v \in H^1_0(\Omega)$. In view of (4.65), we have

$$\forall w \in C^\infty_{\mathrm{c}}(\Omega), \quad a(u, e^{\alpha(r_1+r_2)} w) = a_\alpha(v,w) = \int_\Omega e^{\alpha(r_1+r_2)} f w = a(T, e^{\alpha(r_1+r_2)} w).$$

Hence, $T = u$ and $\|e^{\alpha(r_1+r_2)} T\|_{H^1(\Omega)} = \|e^{\alpha(r_1+r_2)} u\|_{H^1(\Omega)} = \|v\|_{H^1(\Omega)} \leq C_\alpha \|e^{\eta(r_1+r_2)} f\|_{L^2(\Omega)}$. □

As a consequence, we have

$$\begin{aligned} \|e^{\alpha(|\mathbf{r}_1|+|\mathbf{r}_2|)} \psi\|_{L^2(\mathbb{R}^3 \times \mathbb{R}^3)} &= \|e^{\alpha(r_1+r_2)} T\|_{L^2(\Omega)} \leq \|e^{\alpha(r_1+r_2)} T\|_{H^1(\Omega)} \\ &\leq C_\alpha \|e^{\eta(r_1+r_2)} f\|_{L^2(\Omega)} = C_\alpha \|e^{\eta(|\mathbf{r}_1|+|\mathbf{r}_2|)} F\|_{L^2(\mathbb{R}^6)}, \end{aligned}$$

which proves (4.25). In addition, a simple calculation using (4.64) shows that for all $g \in H^1_0(\Omega)$

$$\begin{aligned} \left\| \frac{g}{r_1 r_2} \otimes Y^{m_1}_{l_1} \otimes Y^{m_2}_{l_2} \right\|^2_{H^1(\mathbb{R}^3 \times \mathbb{R}^3)} &= \|g\|^2_{H^1(\Omega)} + l_1(l_1+1) \left\| \frac{g}{r_1} \right\|^2_{L^2(\Omega)} + l_2(l_2+1) \left\| \frac{g}{r_2} \right\|^2_{L^2(\Omega)} \\ &\leq (1 + 4l_1(l_1+1) + 4l_2(l_2+1)) \|g\|^2_{H^1}, \end{aligned}$$

yielding

$$\begin{aligned}\|e^{\alpha(|\mathbf{r}_1|+|\mathbf{r}_2|)}\psi\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)} &\le (1+4l_1(l_1+1)+4l_2(l_2+1))^{1/2}\|e^{\alpha(r_1+r_2)}T\|_{H^1(\Omega)}\\ &\le C_\alpha(1+4l_1(l_1+1)+4l_2(l_2+1))^{1/2}\|e^{\eta(|\mathbf{r}_1|+|\mathbf{r}_2|)}F\|_{L^2(\Omega)}.\end{aligned}$$

Lastly, since H_{l_1,l_2} is a real operator in the sense that $\overline{H_{l_1,l_2}\phi} = H_{l_1,l_2}\overline{\phi}$ for all $\phi \in D(H_{l_1,l_2})$, it is obvious that T is real-valued, whenever f is.

4.4.2 Proof of Lemma 4.1 and Theorem 4.4

We have seen in the previous section that for each $(\alpha, F) \in \mathbb{R}\times L^2(\mathbb{R}^3\times\mathbb{R}^3)$, (4.59) has a unique solution (μ,ψ) in $\mathbb{R}\times H^2(\mathbb{R}^3\times\mathbb{R}^3)$. For $n=1$, we have

$$(H_0-\lambda_0)\phi_1 = -C_1\phi_0, \quad \langle\phi_0,\phi_1\rangle = 0,$$

and it is clear that $(C_1,\phi_1)=(0,0)$ is *a* solution, hence *the* solution, to this system. Likewise, for $n=2$, we have

$$(H_0-\lambda_0)\phi_2 = -C_1\phi_1 - C_2\phi_0 = -C_2\phi_2, \quad \langle\phi_0,\phi_2\rangle = -\frac{1}{2}\langle\phi_1,\phi_1\rangle = 0,$$

so that $(C_2,\phi_2)=(0,0)$. To prove that the Rayleigh–Schrödinger triangular system (4.9)-(4.10) is well-posed and that ϕ_n is of the form (4.26), we proceed by induction on n. It is proven in [78] that for $n=3$,

$$\phi_3 = \frac{T^{(3)}_{(1,1)}(r_1,r_2)}{r_1r_2}\sum_{m=-1}^{1}\alpha^{(3)}_{(1,1,m)}Y_1^m(\theta_1,\phi_1)Y_1^{-m}(\theta_2,\phi_2),$$

with $\alpha^{(3)}_{(1,1,m)} = -\pi G_{\rm c}(1,1,m)$ and $\|T^{(3)}_{(1,1)}(r_1,r_2)e^{\eta^3_{1,1}(r_1+r_2)}\|_{H_1(\Omega)} =: C^3_{1,1} < \infty$. Let $\mathcal{L}_3 = \{(1,1)\}$ and assume that for some $n\ge 3$ the following recursion hypotheses are satisfied (this is the case for $n=3$): for all $3\le k\le n$,

$$\phi_k = \sum_{(l_1,l_2)\in\mathcal{L}_k}\frac{1}{r_1r_2}\left(\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)}T^{(k)}_{(l_1,l_2,m)}(r_1,r_2)Y_{l_1}^m(\theta_1,\phi_1)Y_{l_2}^{-m}(\theta_2,\phi_2)\right), \quad (4.66)$$

for some finite set $\mathcal{L}_k\subset\mathbb{N}^2$ with cardinality $N_k<\infty$, where $T^{(k)}_{(l_1,l_2,m)}$ is the unique solution to (4.19) in $H^1(\Omega)$ (or in $\widetilde{H}^1(\Omega)$ if $l_1=l_2=0$) for $f = f^{(k)}_{(l_1,l_2,m)}\in L^2(\Omega)$ and that for all $(l_1,l_2)\in\mathcal{L}_k$ and $-\min(l_1,l_2)\le m\le\min(l_1,l_2)$ there exists $\eta^k_{l_1,l_2,m}>0$ such that

$$\|T^{(k)}_{(l_1,l_2,m)}(r_1,r_2)e^{\eta^k_{l_1,l_2,m}(r_1+r_2)}\|_{H^1(\Omega)} =: C^k_{l_1,l_2,m} < \infty. \quad (4.67)$$

From (4.14), the fact that $\phi_1=\phi_2=0$ and the recursion hypothesis (4.66), we obtain that for all $3\le k\le n+1$,

$$\begin{aligned}\mathcal{B}^{(k)}\phi_{n+1-k} = \sum_{\substack{l_1+l_2=k-1\\ l_1,l_2\neq 0}}\sum_{(l_1',l_2')\in\mathcal{L}_{n+1-k}}\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)}\sum_{m'=-\min(l_1',l_2')}^{\min(l_1',l_2')}\\ \mathcal{U}\left(f^{m,m'}_{n-k+1,l_1,l_1',l_2,l_2'}\otimes Y_{l_1}^mY_{l_1'}^{m'}\otimes Y_{l_2}^{-m}Y_{l_2'}^{-m'}\right), \quad (4.68)\end{aligned}$$

where

$$f^{m,m'}_{j,l_1,l_1',l_2,l_2'}(r_1,r_2) := G_{\rm c}(l_1,l_2,m)r_1^{l_1}r_2^{l_2}T^{(j)}_{(l_1',l_2',m')}(r_1,r_2).$$

In addition, we have

$$Y_l^m Y_{l'}^{m'} = \sum_{l''=|l-l'|}^{l+l'} \zeta_{l,l',l''}^{m,m'} Y_{l''}^{m+m'} \quad \text{where} \quad \zeta_{l,l',l''} = 0 \text{ if } l+l'+l'' \notin 2\mathbb{N}, \tag{4.69}$$

where the coefficients $\zeta_{l,l',l''}^{m,m'} \in \mathbb{R}$ can be computed explicitly using Wigner's 3-j symbols [64, p. 146]:

$$\zeta_{l,l',l''}^{m,m'} = (-1)^{m+m'} \sqrt{\frac{(2l+1)(2l'+1)(2l''+1)}{4\pi}} \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ m & m' & -m-m' \end{pmatrix}.$$

This implies that

$$\begin{aligned} &-\sum_{k=3}^{n+1} \mathcal{B}^{(k)} \phi_{n+1-k}, -\sum_{k=1}^{n+1} C_k \phi_{n+1-k} \\ &\qquad = \sum_{(l_1,l_2)\in\mathcal{L}_{n+1}} \frac{1}{r_1 r_2} \left(\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} f_{(l_1,l_2,m)}^{(n+1)}(r_1,r_2) Y_{l_1}^m(\theta_1,\phi_1) Y_{l_2}^{-m}(\theta_2,\phi_2) \right), \end{aligned} \tag{4.70}$$

for some $\mathcal{L}_{n+1} \subset \mathbb{N}^2$ with finite cardinality, where for $(l_1,l_2) \in \mathcal{L}_{n+1}$, $-\min(l_1,l_2) \leq m \leq \min(l_1,l_2)$,

$$\begin{aligned} f_{(l_1,l_2,m)}^{(n+1)} = -\sum_{k=3}^{n+1} \Bigg[& C_k T_{(l_1,l_2,m)}^{(n-k+1)}(r_1,r_2) \\ & + \sum_{m'+m''=m} \sum_{\substack{l_1'+l_2'=k-1 \\ l_1',l_2'\neq 0 \\ \min(l_1',l_2')\geq|m'|}} \sum_{\substack{(l_1'',l_2'')\in\mathcal{L}_{n+1-k} \\ \min(l_1'',l_2'')\geq|m''|}} r_1^{l_1'} r_2^{l_2'} T_{(l_1'',l_2'',m'')}^{(n-k+1)}(r_1,r_2) G_{\mathrm{c}}(l_1',l_2',m') \zeta_{l_1',l_1'',l_1}^{m',m''} \zeta_{l_2',l_2'',l_2}^{m',m''} \Bigg], \end{aligned} \tag{4.71}$$

is a linear combinations of the functions $r_1^{l_1'} r_2^{l_2'} T_{(l_1'',l_2'',m'')}^{(j)} \in L^2(\Omega)$, $3 \leq j \leq n$, $l_1'', l_2'' \in \mathcal{L}_j$, $l_1'+l_2'+j \leq n+1$, $-\min(l_1'',l_2'') \leq m'' \leq \min(l_1'',l_2'')$ and therefore satisfies in view of (4.67)

$$\| f_{(l_1,l_2,m)}^{(n+1)}(r_1,r_2) e^{\xi_{l_1,l_2,m}^{n+1}(r_1+r_2)} \|_{H_1(\Omega)} < \infty \tag{4.72}$$

for some $\xi_{l_1,l_2,m}^{n+1} > 0$. Therefore the problem consisting in seeking $(C_{n+1}, \phi_{n+1}) \in \mathbb{R} \times H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ satisfying

$$(H_0 - \lambda_0)\phi_{n+1} = -\sum_{k=3}^{n+1} \mathcal{B}^{(k)} \phi_{n+1-k}, -\sum_{k=1}^{n+1} C_k \phi_{n+1-k}, \quad \left\langle \phi_0, \phi_{n+1} \right\rangle = -\frac{1}{2} \sum_{k=1}^{n} \left\langle \phi_k, \phi_{n+1-k} \right\rangle$$

is well-posed and we deduce from Lemma 4.3 that

$$\phi_{n+1} := \sum_{(l_1,l_2)\in\mathcal{L}_{n+1}} \frac{1}{r_1 r_2} \left(\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} T_{(l_1,l_2,m)}^{(n+1)}(r_1,r_2) Y_{l_1}^m(\theta_1,\phi_1) Y_{l_2}^{-m}(\theta_2,\phi_2) \right),$$

where $T_{(l_1,l_2,m)}^{(n+1)}$ is the unique solution to (4.19) in $H^1(\Omega)$ (or in $\widetilde{H}^1(\Omega)$ if $l_1 = l_2 = 0$) for $f = f_{(l_1,l_2,m)}^{(n+1)}$. In addition, it follows from (4.72) that (4.67) holds true for

$k = n+1$. Therefore, the Rayleigh–Schrödinger triangular system (4.9)-(4.10) is well-posed and the $T^{(n)}_{(l_1,l_2,m)}$'s decay exponentially at infinity in the sense of (4.67). From (4.66) we obtain that for $\alpha_n = \min_{\substack{(l_1,l_2)\in\mathcal{L}_n \\ -\min(l_1,l_2)\le m\le \min(l_1,l_2)}} (\eta^n_{l_1,l_2,m}) > 0$, we have

$$\begin{aligned}\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_n\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)} &\le C_n \sum_{(l_1,l_2)\in\mathcal{L}_n} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} \|e^{\alpha_n(r_1+r_2)} T^{(n)}_{(l_1,l_2,m)}\|_{H^1(\Omega)} \\ &\le C_n \sum_{(l_1,l_2)\in\mathcal{L}_n} \sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} \|e^{\eta^n_{(l_1,l_2,m)}(r_1+r_2)} T^{(n)}_{(l_1,l_2,m)}\|_{H^1(\Omega)} < \infty,\end{aligned}$$

for some $C_n \in \mathbb{R}_+$, so that ϕ_n decays exponentially at infinity in the sense of (4.28). Lastly, we infer from Wigner's $(2n+1)$ rule and the fact that $\phi_1 = \phi_2 = 0$, that $C_n = 0$ for $1 \le n \le 5$. This completes the proof of both Lemma 4.1 and Theorem 4.4.

Let us finally explain how to construct Table 4.1. We have already shown that $\mathcal{L}_3 = \{(1,1)\}$, and from (4.68)-(4.70) and the fact that $\phi_1 = \phi_2 = 0$, we see that

$$\mathcal{L}_{n+1} \subset \left(\bigcup_{k=3}^{n-2} \mathcal{M}_{k,n+1-k} \right) \bigcup \mathcal{M}_{n+1,0} \bigcup \left(\bigcup_{3\le k\le n-5 \mid C_{n+1-k}\neq 0} \mathcal{L}_k \right),$$

where for $k, n \ge 3$,

$$\begin{aligned}\mathcal{M}_{k,0} &= \{(l_1,l_2) \in \mathbb{N}^*\times\mathbb{N}^* \mid l_1+l_2 = k-1\} = \{(1,k-2),\cdots,(k-2,1)\}, \\ \mathcal{M}_{k,n} &= \big\{(l_1,l_2)\in\mathbb{N}\times\mathbb{N} \mid \exists (l_1',l_2')\in\mathcal{M}_{k,0},\ \exists (l_1'',l_2'')\in\mathcal{L}_n \text{ s.t.} \\ &\qquad |l_j'-l_j''| \le l_j \le l_j'+l_j'',\ l_j+l_j'+l_j''\in 2\mathbb{N},\ j=1,2\big\}.\end{aligned}$$

Consequently, we have

$\mathcal{L}_4 = \mathcal{M}_{4,0}$;
$\mathcal{L}_5 = \mathcal{M}_{5,0}$;
$\mathcal{L}_6 = \mathcal{M}_{3,3} \cup \mathcal{M}_{6,0}$ with $\mathcal{M}_{3,3} = \{(0,2;0,2)\}$;
$\mathcal{L}_7 = \mathcal{M}_{3,4} \cup \mathcal{M}_{4,3} \cup \mathcal{M}_{7,0}$ with $\mathcal{M}_{3,4} = \mathcal{M}_{4,3} = \{(0,2;1,3),(1,3;0,2)\}$;
$\mathcal{L}_8 = \mathcal{M}_{3,5} \cup \mathcal{M}_{4,4} \cup \mathcal{M}_{5,3} \cup \mathcal{M}_{8,0}$ with
$\mathcal{M}_{3,5} = \mathcal{M}_{5,3} = \{(0,2;2,4),(1,3;1,3),(2,4;0,2)\}$,
$\mathcal{M}_{4,4} = \{(0,2;0,2,4),(0,2,4;0,2),(1,3;1,3)\}$
$\mathcal{L}_9 = \mathcal{M}_{3,6} \cup \mathcal{M}_{4,5} \cup \mathcal{M}_{5,4} \cup \mathcal{M}_{6,3} \cup \mathcal{M}_{9,0} \cup \mathcal{L}_3$ with
$\mathcal{M}_{6,3} \subsetneq \mathcal{M}_{3,6} = \{(0,2;3,5),(1,3;2,4),(2,4;1,3),(3,5;0,2),(1,3;1,3)\}$,
$\mathcal{M}_{4,5} = \mathcal{M}_{5,4}$
$= \{(0,2;1,3,5),(1,3;0,2,4),(2,4;1,3),(1,3;2,4),(0,2,4;1,3),(1,3,5;0,2)\}$,

where we recall that $(l_1,l_1';l_2,l_2')$ (resp. $(l_1,l_1';l_2,l_2',l_2'')$, $(l_1,l_1',l_1'';l_2,l_2')$) stands for the four (resp. six) pairs (l_1,l_2), (l_1',l_2), (l_1,l_2'), etc. After eliminating redundancies, we obtain Table 4.1.

4.4.3 Proof of Theorem 4.2

As in [78], we introduce the space

$$\mathcal{V} = \left\{ v \in L^2(\mathbb{R}^3\times\mathbb{R}^3) \,:\, v(\mathbf{r}_1,\mathbf{r}_2) = v(\mathbf{r}_2,\mathbf{r}_1)\ \forall \mathbf{r}_1,\mathbf{r}_2 \in \mathbb{R}^3 \right\}, \tag{4.73}$$

the functions $\psi_\epsilon^{(n)} \in \mathcal{V} \cap H^2(\mathbb{R}^3 \times \mathbb{R}^3)$ normalized in $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$,

$$\psi_\epsilon^{(n)} := m_\epsilon^{(n)} \mathcal{T}_\epsilon\left(\phi_\epsilon^{(n)}\right) \quad \text{where} \quad \phi_\epsilon^{(n)} := \phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k \text{ and } m_\epsilon^{(n)} = \left\| \mathcal{T}_\epsilon\left(\phi_\epsilon^{(n)}\right) \right\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}^{-1}, \tag{4.74}$$

as well as the Rayleigh quotient

$$\mu_\epsilon^{(n)} = \left\langle \psi_\epsilon^{(n)}, H_\epsilon \psi_\epsilon^{(n)} \right\rangle \tag{4.75}$$

and the approximation

$$\lambda_\epsilon^{(n)} = \lambda_0 - \sum_{k=6}^{n} C_n \epsilon^n$$

of λ_ϵ. When $\epsilon \to 0$, we have $\mathcal{T}_\epsilon(\phi_0) \to 1$ and therefore $m_\epsilon^{(n)} \to 1$. We know from [78, Section 2.4] that there exists a constant $C \in \mathbb{R}_+$ such that for $\epsilon > 0$ small enough

$$\|\psi_\epsilon - \psi_\epsilon^{(3)}\|_{H^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C\epsilon^4, \quad |\lambda_\epsilon - \mu_\epsilon^{(3)}| \leq C\epsilon^8, \quad \text{and} \quad |\lambda_\epsilon - \lambda_\epsilon^{(6)}| \leq C\epsilon^7.$$

It follows from Theorem 4.4 that the ϕ_n's are in $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$. Since $\mathcal{T}_\epsilon$ continuous on this space, we obtain that for all $n \geq 3$, there exists $c_n \in \mathbb{R}$, such that for $\epsilon > 0$ small enough

$$\|\psi_\epsilon - \psi_\epsilon^{(n)}\|_{H^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n \epsilon^4.$$

We infer from [78, Lemma 2.2 and Appendix A] that there exists a constant $C \in \mathbb{R}_+$ such that for all $n \geq 3$ there exists $\epsilon > 0$ such that for all $0 < \epsilon \leq \epsilon_n$,

$$|\lambda_\epsilon - \mu_\epsilon^{(n)}| \leq C\|H_\epsilon \psi_\epsilon^{(n)} - \mu_\epsilon^{(n)} \psi_\epsilon^{(n)}\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}, \tag{4.76}$$

$$\|\psi_\epsilon - \psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C\|H_\epsilon \psi_\epsilon^{(n)} - \mu_\epsilon^{(n)} \psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \tag{4.77}$$

(the first estimate above follows from the Kato-Temple inequality [190]). To proceed further, we need to evaluate the L^2-norm of the residual $r_\epsilon^{(n)} := H_\epsilon \psi_\epsilon^{(n)} - \mu_\epsilon^{(n)} \psi_\epsilon^{(n)}$. We have

$$\begin{aligned} H_\epsilon \psi_\epsilon^{(n)} &= m_\epsilon^{(n)} H_\epsilon \mathcal{T}_\epsilon(\phi_\epsilon^{(n)}) = m_\epsilon^{(n)} \mathcal{T}_\epsilon\left[(H_0 + V_\epsilon)\phi_\epsilon^{(n)})\right] \\ &= m_\epsilon^{(n)} \mathcal{T}_\epsilon\left[(H_0 + V_\epsilon)(\phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k)\right], \end{aligned}$$

and thus,

$$\begin{aligned} r_\epsilon^{(n)} &= m_\epsilon^{(n)} \mathcal{T}_\epsilon\left[(H_0 + V_\epsilon)\phi_\epsilon^{(n)} - \mu_\epsilon^{(n)} \phi_\epsilon^{(n)}\right] \\ &= m_\epsilon^{(n)} \mathcal{T}_\epsilon\Bigg[(H_0 + V_\epsilon)(\phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k) - (\lambda_0 - \sum_{k=3}^{n} C_k \epsilon^k)(\phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k) \\ &\qquad + (\lambda_\epsilon^{(n)} - \mu_\epsilon^{(n)})\phi_\epsilon^{(n)}\Bigg] \\ &= m_\epsilon^{(n)} \mathcal{T}_\epsilon\Bigg[\left(H_0 + \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)}\right)(\phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k) - (\lambda_0 - \sum_{k=3}^{n} C_k \epsilon^k)(\phi_0 + \sum_{k=3}^{n} \epsilon^k \phi_k) \\ &\qquad + (\lambda_\epsilon^{(n)} - \mu_\epsilon^{(n)})\phi_\epsilon^{(n)} + (V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)})\phi_\epsilon^{(n)}\Bigg]. \end{aligned}$$

Using (4.9), we get

$$\begin{aligned}(H_0+\sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)})(\phi_0+\sum_{k=3}^{n}\epsilon^k\phi_k)-(\lambda_0-\sum_{k=3}^{n}C_k\epsilon^k)(\phi_0+\sum_{k=3}^{n}\epsilon^k\phi_k)\\ =\epsilon^n\sum_{k=1}^{n}\epsilon^k\left(\sum_{j=k}^{n}\mathcal{B}^{(j)}\phi_{n+k-j}+\sum_{j=k}^{n}C_j\phi_{n+k-j}\right).\end{aligned}\tag{4.78}$$

Since $\mathcal{B}^{(j)}$ are degree $(j-1)$ homogeneous functions (in cartesian coordinates) and the ϕ_n's decay exponentially in the sense of (4.28), there exists $K_n\in\mathbb{R}_+$ and $\epsilon_n>0$ such that for all $0<\epsilon\leq\epsilon_n$,

$$\left\|(H_0+\sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)})(\phi_0+\sum_{k=3}^{n}\epsilon^k\phi_k)-(\lambda_0-\sum_{k=3}^{n}C_k\epsilon^k)(\phi_0+\sum_{k=3}^{n}\epsilon^k\phi_k)\right\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}\leq K_n\epsilon^{n+1}.\tag{4.79}$$

It remains to bound $\|(V_\epsilon-\sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)})\psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}$. From (4.6), (4.28) and (4.74), there exists $\epsilon_n>0$, $\alpha_n>0$ and $M_n\in\mathbb{R}_+$ such that for all $0<\epsilon\leq\epsilon_n$

$$\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)}\leq M_n.$$

Introducing

$$\Omega_\epsilon=\left\{(\mathbf{r}_1,\mathbf{r}_2)\in\mathbb{R}^3\times\mathbb{R}^3\,:\,|\mathbf{r}_1|+|\mathbf{r}_2|<(2\epsilon)^{-1}\right\}.\tag{4.80}$$

and the potentials defined by

$$v_\epsilon^{(1)}(\mathbf{r}_1,\mathbf{r}_2):=|\mathbf{r}_1-\epsilon^{-1}\mathbf{e}|^{-1},\ v_\epsilon^{(2)}(\mathbf{r}_1,\mathbf{r}_2):=|\mathbf{r}_2+\epsilon^{-1}\mathbf{e}|^{-1},\ v_\epsilon^{(3)}(\mathbf{r}_1,\mathbf{r}_2):=|\mathbf{r}_1-\mathbf{r}_2-\epsilon^{-1}\mathbf{e}|^{-1},\tag{4.81}$$

we have,

$$\begin{aligned}\|(V_\epsilon-\sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)})\phi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}&\leq\|(V_\epsilon-\sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)})\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon)}+\sum_{k=3}^{n}\epsilon^k\|\mathcal{B}^{(k)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}\\&\quad+\sum_{j=1}^{3}\|v_\epsilon^{(j)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}+\epsilon\|\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}.\end{aligned}$$

We first see that

$$\|\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}\leq e^{-\alpha_n(2\epsilon)^{-1}}\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}\leq M_ne^{-\alpha_n(2\epsilon)^{-1}}.$$

Next, as $\mathcal{B}^{(k)}$ is a polynomial function, there exists a constant B_n such as for all $0<\epsilon\leq\epsilon_n$,

$$\begin{aligned}\sum_{k=3}^{n}\epsilon^k\|\mathcal{B}^{(k)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}&\leq\sum_{k=3}^{n}\epsilon^k\|\mathcal{B}^{(k)}e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\|_{L^\infty(\Omega_\epsilon^c)}\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}\\&\leq M_n\sum_{k=3}^{n}\epsilon^k\|\mathcal{B}^{(k)}e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\|_{L^\infty(\Omega_\epsilon^c)}\leq B_n\epsilon^3e^{-\alpha_n(2\epsilon)^{-1}}.\end{aligned}$$

In addition, we have

$$\begin{aligned}\sum_{j=1}^{3}\|v_\epsilon^{(j)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}&\leq\sum_{j=1}^{3}e^{-\alpha_n(2\epsilon)^{-1}}\|v_\epsilon^{(j)}e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon^c)}\\&\leq\sum_{j=1}^{3}e^{-\alpha_n(2\epsilon)^{-1}}\|v_\epsilon^{(j)}e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}\\&\leq 8e^{-\alpha_n(2\epsilon)^{-1}}\|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{H^1(\mathbb{R}^3\times\mathbb{R}^3)}=8e^{-\alpha_n(2\epsilon)^{-1}}M_n,\end{aligned}$$

where we have used the Hardy inequality in dimension 3

$$\forall \phi \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \frac{|\phi(\mathbf{r})|^2}{|\mathbf{r}|^2}\, d\mathbf{r} \leq 4 \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{r})|^2\, d\mathbf{r}$$

to show that for any $\psi \in H^1(\mathbb{R}^3 \times \mathbb{R}^3)$,

$$\begin{aligned}\|v_\epsilon^{(j)}\psi\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} &= \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} \frac{|\psi(\mathbf{r}_1,\mathbf{r}_2)|^2}{|\mathbf{r}_j + (-1)^j \epsilon^{-1}\mathbf{e}|^2}\, d\mathbf{r}_j \right) d\mathbf{r}_{3-j} \\ &\leq \int_{\mathbb{R}^3} 4 \left(\int_{\mathbb{R}^3} |\nabla_{\mathbf{r}_j}\psi(\mathbf{r}_1,\mathbf{r}_2)|^2\, d\mathbf{r}_j \right) d\mathbf{r}_{3-j} \leq 4\|\nabla_{\mathbf{r}_j}\psi\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)},\end{aligned}$$

for $j = 1, 2$, and

$$\begin{aligned}\|v_\epsilon^{(3)}\psi\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{|\psi(\mathbf{r}_1,\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2 - \epsilon^{-1}\mathbf{e}|^2}\, d\mathbf{r}_1\, d\mathbf{r}_2 \\ &= \frac{1}{8}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{|\psi\left(\mathbf{r}_1' + \mathbf{r}_2', \mathbf{r}_1' - \mathbf{r}_2'\right)|^2}{|\mathbf{r}_2' - \epsilon^{-1}\mathbf{e}|^2}\, d\mathbf{r}_1'\, d\mathbf{r}_2' \\ &\leq \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |(\nabla_{\mathbf{r}_1} - \nabla_{\mathbf{r}_2})\psi\left(\mathbf{r}_1' + \mathbf{r}_2', \mathbf{r}_1' - \mathbf{r}_2'\right)|^2\, d\mathbf{r}_1'\, d\mathbf{r}_2' \\ &= 4\|(\nabla_{\mathbf{r}_1} - \nabla_{\mathbf{r}_2})\psi\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = 8\|\nabla\psi\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}.\end{aligned}$$

From the multipolar expansion of V_ϵ, we know that there exist $c_n \in \mathbb{R}_+$

$$\left| V_\epsilon(\mathbf{r}_1,\mathbf{r}_2) - \sum_{i=3}^{n} \epsilon^i \mathcal{B}^{(i)}(\mathbf{r}_1,\mathbf{r}_2) \right| \leq c_n K^n \epsilon^{n+1}, \quad \text{whenever } |\mathbf{r}_1| + |\mathbf{r}_2| \leq K \leq (2\epsilon)^{-1}. \tag{4.82}$$

Let us now show that (4.82) implies that there exists $\widetilde{c}_n \in \mathbb{R}_+$ such that for all $0 \leq K \leq (2\epsilon)^{-1}$,

$$\sup_{|\mathbf{r}_1|+|\mathbf{r}_2|\leq K} \left| V_\epsilon(\mathbf{r}_1,\mathbf{r}_2) - \sum_{i=3}^{n} \epsilon^i \mathcal{B}^{(i)}(\mathbf{r}_1,\mathbf{r}_2) \right| e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)} \leq \widetilde{c}_n \epsilon^{n+1}, \tag{4.83}$$

This is immediate from (4.82) for $K \leq 1$, taking $\widetilde{c}_n = c_n$. Now we let $K > 1$. Then (4.82) implies

$$\sup_{(K/2)\leq(|\mathbf{r}_1|+|\mathbf{r}_2|)\leq K} \left| V_\epsilon(\mathbf{r}_1,\mathbf{r}_2) - \sum_{i=3}^{n} \epsilon^i \mathcal{B}^{(i)}(\mathbf{r}_1,\mathbf{r}_2) \right| e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)} \leq c_n e^{-\alpha_n K/2} K^n \epsilon^{n+1}.$$

Applying this repeatedly for $2^{-j}K$ replacing K until $2^{-j}K < 1$ yields (4.83), with

$$\widetilde{c}_n = c_n \sup_{t\geq 0} t^n e^{-\alpha_n t/2}.$$

Applying (4.83) for $K = (2\epsilon)^{-1}$ yields

$$\|(V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)}) e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\|_{L^\infty(\Omega_\epsilon)} \leq \widetilde{c}_n \epsilon^{n+1},$$

from which we obtain

$$\begin{aligned}&\|(V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)})\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon)} \\ &\quad\leq \|(V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)}) e^{-\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\|_{L^\infty(\Omega_\epsilon)} \|e^{\alpha_n(|\mathbf{r}_1|+|\mathbf{r}_2|)}\phi_\epsilon^{(n)}\|_{L^2(\Omega_\epsilon)} \\ &\quad\leq \widetilde{c}_n M_n \epsilon^{n+1}.\end{aligned}$$

Finally, we get

$$\|(V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)})\phi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq \widetilde{c}_n M_n \epsilon^{n+1} + (8 + \epsilon + B_n\epsilon^3) M_n e^{-\alpha_n (2\epsilon)^{-1}}, \quad (4.84)$$

Together with (4.79), this proves that there exists $c_n'' \in \mathbb{R}_+$ such that for all $0 < \epsilon \leq \epsilon_n$,

$$\|r_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = \|H_\epsilon \psi_\epsilon^{(n)} - \mu_\epsilon^{(n)} \psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n'' \epsilon^{n+1}. \quad (4.85)$$

It follows from (4.76)-(4.77) that for $n \geq 3$ fixed, there exists $C \in \mathbb{R}_+$ such that for all $0 < \epsilon \leq \epsilon_n$,

$$|\lambda_\epsilon - \mu_\epsilon^{(n)}| \leq C\epsilon^{2(n+1)} \quad \text{and} \quad \|\psi_\epsilon - \psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C\epsilon^{n+1}. \quad (4.86)$$

Then,

$$\begin{aligned}
&\mu_\epsilon^{(n)} - \lambda_\epsilon^{(n)} \\
&\quad = \left\langle \psi_\epsilon^{(n)}, H_\epsilon \psi_\epsilon^{(n)} - \lambda_\epsilon^{(n)} \psi_\epsilon^{(n)} \right\rangle \\
&\quad = m_\epsilon^{(n)} \left\langle \psi_\epsilon^{(n)}, \mathcal{T}_\epsilon \Big[(V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)})\phi_\epsilon^{(n)} + \epsilon^n \sum_{k=1}^{n} \epsilon^k \Big(\sum_{j=k}^{n} \mathcal{B}^{(j)} \phi_{n+k-j} + \sum_{j=k}^{n} C_j \phi_{n+k-j} \Big) \Big] \right\rangle
\end{aligned}$$

so that there exists a constant c_n such that for $0 < \epsilon \leq \epsilon_n$,

$$\begin{aligned}
&\left|\mu_\epsilon^{(n)} - \lambda_\epsilon^{(n)}\right| \\
&\qquad \leq 2 \left\| (V_\epsilon - \sum_{k=3}^{n} \epsilon^k \mathcal{B}^{(k)})\phi_\epsilon^{(n)} + \epsilon^n \sum_{k=1}^{n} \epsilon^k \Big(\sum_{j=k}^{n} \mathcal{B}^{(j)} \phi_{n+k-j} + \sum_{j=k}^{n} C_j \phi_{n+k-j} \Big) \right\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \\
&\qquad \leq c_n \epsilon^{n+1}.
\end{aligned}$$

The error bounds on the eigenvalue errors in (4.12) follow from (4.86) and the above inequality.

Finally, the error $\xi_\epsilon^{(n)} = \psi_\epsilon - \psi_\epsilon^{(n)}$, as defined in [78], satisfies

$$H_\epsilon \xi_\epsilon^{(n)} = \lambda_\epsilon \psi_\epsilon - H_\epsilon \psi_\epsilon^{(n)} = \lambda_\epsilon - \mu_\epsilon^{(n)} - r_\epsilon^{(n)} =: \eta_\epsilon^{(n)}.$$

From (4.85)-(4.86), there exists a constant $c_n \in \mathbb{R}_+$ such that for all $0 < \epsilon \leq \epsilon_n$,

$$\|\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n \epsilon^{n+1} \quad \text{and} \quad \|\eta_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n \epsilon^{n+1}.$$

In addition,

$$-\frac{1}{2}\Delta \xi_\epsilon^{(n)} = -W_\epsilon \xi_\epsilon^{(n)} + \eta_\epsilon^{(n)}, \quad (4.87)$$

where

$$\begin{aligned}
W_\epsilon(\mathbf{r}_1, \mathbf{r}_2) := &- \frac{1}{|\mathbf{r}_1 - (2\epsilon)^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_2 - (2\epsilon)^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_1 + (2\epsilon)^{-1}\mathbf{e}|} \\
&- \frac{1}{|\mathbf{r}_2 + (2\epsilon)^{-1}\mathbf{e}|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \epsilon.
\end{aligned}$$

Proceeding as in [78, Section 2.4], we use the Hardy inequality in $\mathbb{R}^3$ and the Cauchy-Schwarz inequality to obtain that

$$\begin{aligned}
&\frac{1}{2}\|\nabla \xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}^2 \\
&\qquad = \langle \xi_\epsilon^{(n)}, -W_\epsilon \xi_\epsilon^{(n)} + \eta_\epsilon^{(n)} \rangle \\
&\qquad \leq (10\|\nabla \xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} + \epsilon\|\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} + \|\eta_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)})\|\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)},
\end{aligned}$$

$$\frac{1}{2}\|\Delta\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} = \| - W_\epsilon\xi_\epsilon^{(n)} + \eta_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}$$
$$\leq 10\|\nabla\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} + \epsilon\|\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} + \|\eta_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}.$$

It follows from (4.87) that there exists a constant $c_n \in \mathbb{R}_+$ such that for all $0 < \epsilon \leq \epsilon_n$, $\|\Delta\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n\epsilon^{n+1}$, and thus $\|\xi_\epsilon^{(n)}\|_{H^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq c_n\epsilon^{n+1}$.

Appendix C

Appendix of Chapter 4

C.1 Multipolar expansion of V_ϵ

We start from the well-known multipolar expansion of $\frac{1}{|\mathbf{r}-R\mathbf{e}|}$ in terms of Legendre polynomials

$$\frac{1}{|\mathbf{r}-R\mathbf{e}|} = \frac{1}{R}\left(\sum_{k=0}^{\infty} P_k\Big(\frac{\mathbf{r}\cdot\mathbf{e}}{|\mathbf{r}|}\Big)\left(\frac{|\mathbf{r}|}{R}\right)^k\right), \quad \text{for } |\mathbf{r}| < R, \tag{C.1}$$

which is a straightforward consequence of the definition of Legendre polynomials via their generating function [321]

$$\forall -1 \le x \le 1, \quad \big(1-2xt+t^2\big)^{-1/2} = \sum_{k=0}^{\infty} P_k(x)t^k, \tag{C.2}$$

taking

$$-1 \le x = \frac{\mathbf{r}\cdot\mathbf{e}}{|\mathbf{r}|} \le 1, \qquad t = \frac{|\mathbf{r}|}{R}.$$

Since the Legendre polynomials are at most 1 in magnitude on the interval $[-1,1]$, the sum in (C.2) converges absolutely for all $|t|<1$, and

$$\Big|\sum_{k=n}^{\infty} P_k(x)t^k\Big| \le \sum_{k=n}^{\infty} t^k = \frac{t^n}{1-t} \le 2t^n, \quad \text{for all } |t| \le \tfrac{1}{2}.$$

Consequently,

$$\left|\frac{1}{|\mathbf{r}-R\mathbf{e}|} - \frac{1}{R}\left(\sum_{k=0}^{n-1} P_k\Big(\frac{\mathbf{r}\cdot\mathbf{e}}{|\mathbf{r}|}\Big)\left(\frac{|\mathbf{r}|}{R}\right)^k\right)\right| \le 2\frac{|\mathbf{r}|^n}{R^{n+1}}, \quad \text{for all } |\mathbf{r}| \le R/2. \tag{C.3}$$

Recalling that $P_0(x)=1$, $P_1(x)=x$ and

$$V_\epsilon(\mathbf{r}_1,\mathbf{r}_2) = -\frac{1}{|\mathbf{r}_1-\epsilon^{-1}\mathbf{e}|} - \frac{1}{|\mathbf{r}_2+\epsilon^{-1}\mathbf{e}|} + \frac{1}{|\mathbf{r}_1-\mathbf{r}_2-\epsilon^{-1}\mathbf{e}|} + \epsilon.$$

with $\epsilon = R^{-1}$, we deduce from (C.3) that

$$\left|V_\epsilon(\mathbf{r}_1,\mathbf{r}_2) - \sum_{k=3}^{n}\epsilon^k\mathcal{B}^{(k)}(\mathbf{r}_1,\mathbf{r}_2)\right| \le 6K^n\epsilon^{n+1}, \quad \text{whenever } |\mathbf{r}_1|+|\mathbf{r}_2| \le K \le (2\epsilon)^{-1}, \tag{C.4}$$

where the polynomial functions $\mathcal{B}^{(k)}$ are given by

$$\mathcal{B}^{(k)}(\mathbf{r}_1,\mathbf{r}_2) := P_{k-1}\left(\frac{(\mathbf{r}_1-\mathbf{r}_2)\cdot\mathbf{e}}{|\mathbf{r}_1-\mathbf{r}_2|}\right)|\mathbf{r}_1-\mathbf{r}_2|^{k-1} - P_{k-1}\left(\frac{\mathbf{r}_1\cdot\mathbf{e}}{|\mathbf{r}_1|}\right)|\mathbf{r}_1|^{k-1} - P_{k-1}\left(-\frac{\mathbf{r}_2\cdot\mathbf{e}}{|\mathbf{r}_2|}\right)|\mathbf{r}_2|^{k-1}.$$

This proves (4.82). To derive the expression (4.14) for the $\mathcal{B}^{(k)}$'s, we first use the identities

$$P_l(\sigma\cdot\sigma') = \left(\frac{4\pi}{2l+1}\right)\sum_{m=-l}^{l}(-1)^m Y_l^m(\sigma)Y_l^m(\sigma'), \qquad \sqrt{\frac{4\pi}{2l+1}}Y_l^m(\mathbf{e}) = \delta_{m,0},$$

valid for all $l\in\mathbb{N}$, $-l\le m\le l$, $\sigma,\sigma'\in\mathbb{S}^2$ (recall that $\mathbf{e}$ is the unit vector of the z-axis), and get

$$\mathcal{B}^{(k)}(\mathbf{r}_1,\mathbf{r}_2) := \sqrt{\frac{4\pi}{2k-1}}\left(Y_{k-1}^0\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{|\mathbf{r}_1-\mathbf{r}_2|}\right)|\mathbf{r}_1-\mathbf{r}_2|^{k-1} - Y_{k-1}^0\left(\frac{\mathbf{r}_1}{|\mathbf{r}_1|}\right)|\mathbf{r}_1|^{k-1} - Y_{k-1}^0\left(-\frac{\mathbf{r}_2}{|\mathbf{r}_2|}\right)|\mathbf{r}_2|^{k-1}\right).$$

We next use the addition formula [316] stating that for $l\in\mathbb{N}$, $\mathbf{r}_1,\mathbf{r}_2\in\mathbb{R}^3$,

$$\sqrt{\frac{4\pi}{2l+1}}Y_l^0\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{|\mathbf{r}_1-\mathbf{r}_2|}\right)|\mathbf{r}_1-\mathbf{r}_2|^l = \sum_{l_1+l_2=l}\sum_{m=-\min(l_1,l_2)}^{\min(l_1,l_2)} G_{\rm c}(l_1,l_2,m)r_1^{l_1}Y_{l_1}^m\left(\frac{\mathbf{r}_1}{|\mathbf{r}_1|}\right)r_2^{l_2}Y_{l_2}^{-m}\left(\frac{\mathbf{r}_2}{|\mathbf{r}_2|}\right),$$

where

$$G_{\rm c}(l_1,l_2,m) = (-1)^{l_2}\frac{4\pi}{((2l_1+1)(2l_2+1))^{1/2}}\binom{l_1+l_2}{l_1+m}^{1/2}\binom{l_1+l_2}{l_1-m}^{1/2}, = (-1)^{l_2}\frac{4\pi(l_1+l_2)!}{((2l_1+1)(2l_2+1)(l_1+m)!(l_2+m)!(l_1-m)!(l_2-m)!)^{1/2}}.$$

As for $G_{\rm c}(l,0,0) = G_{\rm c}(0,l,0) = \frac{4\pi}{(2l+1)^{1/2}}$ and $Y_0^0 = \frac{1}{\sqrt{4\pi}}$, we finally obtain (4.14).

C.2 Wigner $(2n+1)$ rule

Using the notation in (4.74), we consider the Rayleigh quotients

$$\mu_\epsilon^{(n)} = \langle\psi_\epsilon^{(n)}, H_\epsilon\psi_\epsilon^{(n)}\rangle \quad\text{and}\quad \widetilde{\mu}_\epsilon^{(n)} = \frac{\left\langle\phi_\epsilon^{(n)}, \left(H_0+\sum_{i=3}^{2n+1}\epsilon^i\mathcal{B}^{(i)}\right)\phi_\epsilon^{(n)}\right\rangle}{\|\phi_\epsilon^{(n)}\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}}$$

(recall that $\|\psi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}=1$). Let

$$\eta_\epsilon^{(n)} := (H_0+V_\epsilon)\phi_\epsilon^{(n)}, \quad \upsilon_\epsilon^{(n)} := (V_\epsilon - \sum_{i=3}^{2n+1}\epsilon^i\mathcal{B}^{(i)})\phi_\epsilon^{(n)} \quad\text{and}\quad \xi_\epsilon^{(n)} := (\mathcal{T}_\epsilon^*\mathcal{T}_\epsilon - 1)\phi_\epsilon^{(n)}.$$

We deduce from the boundedness of the ϕ_n's in $H^2(\mathbb{R}^3 \times \mathbb{R}^3)$, the Hardy inequality in $\mathbb{R}^3$, and the estimates (4.28) and (4.82), that there exist $C \in \mathbb{R}_+$, $\beta_n > 0$ and $\epsilon_n > 0$ such that for all $0 \leq \epsilon \leq \epsilon_n$

$$\begin{aligned} &\|\phi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq 2, && \|\eta_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C, \\ &\|\upsilon_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq C\epsilon^{2n+2}, && \|\xi_\epsilon^{(n)}\|_{L^2(\mathbb{R}^3\times\mathbb{R}^3)} \leq Ce^{-\beta_n\epsilon}, \end{aligned}$$

proceeding as in the proof of (4.84) to establish the third inequality. It follows from (4.12) and the above bounds that

$$\begin{aligned} \widetilde{\mu}_\epsilon^{(n)} &= \lambda_\epsilon + \widetilde{\mu}_\epsilon^{(n)} - \mu_\epsilon^{(n)} + O(\epsilon^{2n+2}) \\ &= \lambda_\epsilon + \frac{\left\langle \phi_\epsilon^{(n)}, \left(H_0 + \sum_{i=3}^{2n+1} \epsilon^i \mathcal{B}^{(i)}\right) \phi_\epsilon^{(n)} \right\rangle}{\|\phi_\epsilon^{(n)}\|^2_{L^2(\mathbb{R}^3\times\mathbb{R}^3)}} - \frac{\left\langle \mathcal{T}_\epsilon^* \mathcal{T}_\epsilon \phi_\epsilon^{(n)}, (H_0 + V_\epsilon)\phi_\epsilon^{(n)} \right\rangle}{\left\langle \mathcal{T}_\epsilon^* \mathcal{T}_\epsilon \phi_\epsilon^{(n)}, \phi_\epsilon^{(n)} \right\rangle} + O(\epsilon^{2n+2}) \\ &= \lambda_\epsilon - \frac{\left\langle \phi_\epsilon^{(n)}, \upsilon_\epsilon^{(n)} \right\rangle}{\left\langle \phi_\epsilon^{(n)}, \phi_\epsilon^{(n)} \right\rangle} + \frac{\left\langle \xi_\epsilon^{(n)}, \eta_\epsilon^{(n)} \right\rangle - \left\langle \xi_\epsilon^{(n)}, \phi_\epsilon^{(n)} \right\rangle \left\langle \phi_\epsilon^{(n)}, \eta_\epsilon^{(n)} \right\rangle}{\left\langle \phi_\epsilon^{(n)}, \phi_\epsilon^{(n)} \right\rangle + \left\langle \xi_\epsilon^{(n)}, \phi_\epsilon^{(n)} \right\rangle} + O(\epsilon^{2n+2}) \\ &= \lambda_\epsilon + O(\epsilon^{2n+2}) = -1 - \sum_{k=6}^{2n+1} C_k \epsilon^k + O(\epsilon^{2n+2}). \end{aligned}$$

Thus, the coefficients C_k for $k \leq 2n+1$ can be computed from the Taylor expansion of $\widetilde{\mu}_\epsilon^{(n)}$ up to order $(2n+1)$, which only involves the ϕ_k's for $k \leq n$, and the $\mathcal{B}^{(k)}$'s for $k \leq (2n+1)$. To obtain a computable expression of the coefficients C_{2n} and C_{2n+1}, we first use Equation (4.9), which can be rewritten as

$$H_0\phi_k + \sum_{j=3}^{k} \mathcal{B}^{(j)}\phi_{k-j} = -C_0\phi_k - \sum_{j=6}^{k} C_j\phi_{k-j} = -\sum_{j=0}^{k} C_j\phi_{k-j}, \tag{C.5}$$

with $C_0 = 1$ and $C_i = 0$ for $i = 1, ..., 5$, to get that for all $n \geq 1$

$$\begin{aligned} \nu_\epsilon^{(n)} :&= \left\langle \phi_\epsilon^{(n)}, \left(H_0 + \sum_{i=3}^{2n+1} \epsilon^i \mathcal{B}^{(i)}\right) \phi_\epsilon^{(n)} \right\rangle \\ &= -\sum_{l=0}^{n} \epsilon^l \sum_{i=0}^{l} \left\langle \phi_i, \sum_{j=0}^{l-i} C_j \phi_{l-i-j} \right\rangle + \epsilon^n \sum_{l=1}^{n} \epsilon^l \left(-\sum_{i=l}^{n} \left\langle \phi_i, \sum_{j=0}^{n+l-i} C_j \phi_{n+l-i-j} \right\rangle \right. \\ &\quad \left. + \sum_{i=0}^{l-1} \left\langle \phi_i, \sum_{j=0}^{n} \mathcal{B}^{(n+l-i-j)} \phi_j \right\rangle \right) + \epsilon^{2n+1} \sum_{i=0}^{n} \left\langle \phi_i, \sum_{j=0}^{n} \mathcal{B}^{(2n+1-i-j)} \phi_j \right\rangle + O(\epsilon^{2n+2}). \end{aligned} \tag{C.6}$$

In addition, we have

$$\|\phi_\epsilon^{(n)}\|^2 = \left\langle \sum_{i=0}^{n} \epsilon^i \phi_i, \sum_{i=0}^{n} \epsilon^i \phi_j \right\rangle = 1 + \sum_{k=1}^{n} \epsilon^k \sum_{i=0}^{k} \left\langle \phi_i, \phi_{k-i} \right\rangle + \epsilon^n \sum_{k=1}^{n} \epsilon^k \sum_{i=k}^{n} \left\langle \phi_i, \phi_{n+k-i} \right\rangle,$$

and, using the relation $\sum_{i=0}^{k} \left\langle \phi_i, \phi_{k-i} \right\rangle = 0$ derived from (4.10), we get

$$\|\phi_\epsilon^{(n)}\|^2 = 1 + \epsilon^n \sum_{k=1}^{n} \epsilon^k \sum_{i=k}^{n} \left\langle \phi_i, \phi_{n+k-i} \right\rangle. \tag{C.7}$$

It follows from (C.6)-(C.7) that

$$\widetilde{\mu}_\epsilon^{(n)} = \frac{\nu_\epsilon^{(n)}}{\|\phi_\epsilon^{(n)}\|^2} = -\sum_{k=0}^{2n+1} C_k \epsilon^k + O(\epsilon^{2n+2}),$$

with

$$C_{2n} = \big\langle \phi_n, \sum_{j=0}^{n} C_j \phi_{n-j} \big\rangle - \sum_{i=0}^{n-1} \big\langle \phi_i, \sum_{j=0}^{n} \mathcal{B}^{(2n-i-j)} \phi_j \big\rangle \\ - \sum_{k=1}^{n} \left(\sum_{i=k}^{n} \big\langle \phi_i, \phi_{n+k-i} \big\rangle \right) \sum_{i=0}^{n-k} \big\langle \phi_i, \sum_{j=0}^{n-k-i} C_j \phi_{n-k-i-j} \big\rangle,$$

and

$$C_{2n+1} = - \sum_{k=1}^{n} \left(\sum_{i=k}^{n} \big\langle \phi_i, \phi_{n+k-i} \big\rangle \right) \sum_{i=0}^{n+1-k} \big\langle \phi_i, \sum_{j=0}^{n+1-k-i} C_j \phi_{n+1-k-i-j} \big\rangle \\ - \sum_{i=0}^{n} \big\langle \phi_i, \sum_{j=0}^{n} \mathcal{B}^{(2n+1-i-j)} \phi_j \big\rangle.$$

C.3 Computation of the integrals S_n in (4.57)

Recall that

$$S_n = \int_0^{+\infty} r^3 e^{-r} \varphi_{n,1}(r) \mathrm{d}r,$$

where

$$\varphi_{n,1} = \sqrt{\left(\frac{2}{n}\right)^3 \frac{(n-2)!}{2n(n+1)!}} \left(\frac{2r}{n}\right) L_{n-2}^{(3)} \left(\frac{2r}{n}\right) e^{-r/n},$$

where the associated Laguerre polynomials of the second kind $L_n^{(m)}$, $n, m \in \mathbb{N}$, satisfy the following properties [1, Section 22.2]:

- for all $k, k', m \in \mathbb{N}$,

$$\int_0^{\infty} x^m L_k^{(m)}(x) L_{k'}^{(m)}(x) e^{-x} \, \mathrm{d}x = \frac{(k+m)!}{k!} \delta_{k,k'}; \tag{C.8}$$

- for all $\gamma \in \mathbb{C}$ such that $\Re(\gamma) > -\frac{1}{2}$, and $m \in \mathbb{N}$,

$$e^{-\gamma x} = \sum_{k=0}^{+\infty} \frac{\gamma^k}{(1+\gamma)^{k+m+1}} L_k^{(m)}(x); \tag{C.9}$$

- for all $k, m \in \mathbb{N}$,

$$x L_k^{(m+1)}(x) = (k+m+1) L_k^{(m)}(x) - (k+1) L_{k+1}^{(m)}(x). \tag{C.10}$$

By a change of variable, we obtain

$$S_n = \frac{n^2}{8} \sqrt{\frac{(n-2)!}{(n+1)!}} I_n \quad \text{with} \quad I_n := \int_0^{+\infty} x^4 L_{n-2}^{(3)} e^{-\frac{n-1}{2} x} e^{-x} \, dx.$$

Applying (C.9) for $\gamma = \frac{n-1}{2}$ and $m = 4$, then (C.10) for $m = 3$, and finally (C.8) for $m = 3$, we obtain

$$\begin{aligned}
I_n &= \int_0^{+\infty} x^4 L_{n-2}^{(3)} \left(\sum_{k=0}^{+\infty} \frac{2^5 (n-1)^k}{(n+1)^{k+5}} L_k^{(4)}(x) \right) e^{-x}\, dx \\
&= \int_0^{+\infty} x^3 L_{n-2}^{(3)} \left(\sum_{k=0}^{+\infty} \frac{2^5 (n-1)^k}{(n+1)^{k+5}} \left((k+4) L_k^{(3)}(x) - (k+1) L_{k+1}^{(3)}(x) \right) \right) e^{-x}\, dx \\
&= \sum_{k=0}^{+\infty} \frac{2^5 (n-1)^k}{(n+1)^{k+5}} \left((k+4) \frac{(k+3)!}{k!} \delta_{k,n-2} - (k+1) \frac{(k+4)!}{(k+1)!} \delta_{k+1,n-2} \right) \\
&= \frac{2^5 (n-1)^{n-2}}{(n+1)^{n+3}} (n+2) \frac{(n+1)!}{(n-2)!} - \frac{2^5 (n-1)^{n-3}}{(n+1)^{n+2}} (n-2) \frac{(n+1)!}{(n-2)!} \\
&= \frac{2^6 n (n-1)^{n-3}}{(n+1)^{n+3}} \frac{(n+1)!}{(n-2)!}.
\end{aligned}$$

Finally, we get

$$S_n = 8n^3 \frac{(n-1)^{n-3}}{(n+1)^{n+3}} \sqrt{\frac{(n+1)!}{(n-2)!}}.$$

Bibliography

[1] Milton Abramowitz and Irene A Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55. US Government Printing Office, 1970.

[2] Martial Agueh and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM J. Math. Anal.*, 43(2):904–924, 2011.

[3] Reinhart Ahlrichs. Convergence properties of the intermolecular force series (1/R-expansion). *Theoretica chimica acta*, 41(1):7–15, 1976.

[4] Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. *Network flows*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1993. Theory, algorithms, and applications.

[5] Aurélien Alfonsi, Jacopo Corbetta, and Benjamin Jourdain. Sampling of one-dimensional probability measures in the convex order and computation of robust option price bounds. *International Journal of Theoretical and Applied Finance*, 0(0):1950002, 0.

[6] Aurélien Alfonsi, Jacopo Corbetta, and Benjamin Jourdain. Sampling of probability measures in the convex order by Wasserstein projection. *arXiv e-prints*, page arXiv:1709.05287, Sep 2017.

[7] Aurélien Alfonsi, Rafaël Coyaud, and Virginie Ehrlacher. Constrained overdamped langevin dynamics for symmetric multimarginal optimal transportation. *arXiv preprint arXiv:2102.03091*, 2021.

[8] Aurélien Alfonsi, Rafaël Coyaud, Virginie Ehrlacher, and Damiano Lombardi. Approximation of optimal transport problems with marginal moments constraints. *Math. Comp.*, 90(328):689–737, 2021.

[9] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[10] Neemias Alves de Lima. Van der Waals density functional from multipole dispersion interactions. *The Journal of chemical physics*, 132(1):014110, 2010.

[11] Luigi Ambrosio. Lecture notes on optimal transport problems. In *Mathematical aspects of evolving interfaces (Funchal, 2000)*, volume 1812 of *Lecture Notes in Math.*, pages 1–52. Springer, Berlin, 2003.

[12] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[13] Ioannis Anapolitanos. *On van der Waals forces*. PhD thesis, University of Toronto, 2011.

[14] Ioannis Anapolitanos. Remainder estimates for the long range behavior of the van der waals interaction energy. In *Annales Henri Poincaré*, volume 17, pages 1209–1261. Springer, 2016.

[15] Ioannis Anapolitanos, Mariam Badalyan, and Dirk Hundertmark. On the van der Waals interaction between a molecule and a half-infinite plate. *arXiv:2004.04771*, 2020.

[16] Ioannis Anapolitanos and Mathieu Lewin. Compactness of molecular reaction paths in quantum mechanics. *Arch. Ration. Mech. Anal.*, 236(2):505–576, 2020.

[17] Ioannis Anapolitanos, Mathieu Lewin, and Matthias Roth. Differentiability of the van der Waals interaction between two atoms. *arXiv:1902.06683*, 2019.

[18] Ioannis Anapolitanos and Israel Michael Sigal. Long-range behavior of the van der Waals force. *Communications on Pure and Applied Mathematics*, 70(9):1633–1671, 2017.

[19] Sigurd Angenent, Steven Haker, and Allen Tannenbaum. Minimizing flows for the Monge-Kantorovich problem. *SIAM J. Math. Anal.*, 35(1):61–97, 2003.

[20] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein gan. *arXiv preprint arXiv:1701.07875*, 2017.

[21] Robert L Baldwin. Energetics of protein folding. *Journal of Molecular Biology*, 371(2):283–301, 2007.

[22] Jean-Marie Barbaroux, Michael Hartig, Dirk Hundertmark, and Semjon Vugalter. Van der Waals-London interaction of atoms with pseudo-relativistic kinetic energy. *arXiv:1902.09222*, 2019.

[23] Federico Bassetti, Stefano Gualandi, and Marco Veneroni. On the computation of kantorovich–wasserstein distances between two-dimensional histograms by uncapacitated minimum cost flows. *SIAM Journal on Optimization*, 30(3):2441–2469, 2020.

[24] Christian Bayer and Josef Teichmann. The proof of Tchakaloff's theorem. *Proceedings of the American mathematical society*, 134(10):3035–3040, 2006.

[25] Mathias Beiglböck, Alexander M. G. Cox, and Martin Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math.*, 208(2):327–400, 2017.

[26] Mathias Beiglböck, Martin Goldstern, Gabriel Maresch, and Walter Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009.

[27] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013.

[28] Mathias Beiglböck, Pierre Henry-Labordère, and Nizar Touzi. Monotone martingale transport plans and Skorokhod embedding. *Stochastic Process. Appl.*, 127(9):3005–3013, 2017.

[29] Mathias Beiglböck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016.

[30] Mathias Beiglböck, Christian Léonard, and Walter Schachermayer. A general duality theorem for the Monge-Kantorovich transport problem. *Studia Math.*, 209(2):151–167, 2012.

[31] Mathias Beiglböck, Christian Léonard, and Walter Schachermayer. A generalized dual maximizer for the Monge-Kantorovich transport problem. *ESAIM Probab. Stat.*, 16:306–323, 2012.

[32] Mathias Beiglböck, Christian Léonard, and Walter Schachermayer. On the duality theory for the Monge-Kantorovich transport problem. In *Optimal transportation*, volume 413 of *London Math. Soc. Lecture Note Ser.*, pages 216–265. Cambridge Univ. Press, Cambridge, 2014.

[33] Mathias Beiglböck, Tongseok Lim, and Jan Obł ój. Dual attainment for the martingale transport problem. *Bernoulli*, 25(3):1640–1658, 2019.

[34] Mathias Beiglböck and Marcel Nutz. Martingale inequalities and deterministic counterparts. *Electron. J. Probab.*, 19:no. 95, 15, 2014.

[35] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi. Complete duality for martingale optimal transport on the line. *Ann. Probab.*, 45(5):3038–3074, 2017.

[36] J. D. Benamou and Y. Brenier. Mixed L^2-Wasserstein optimal mapping between prescribed density functions. *J. Optim. Theory Appl.*, 111(2):255–271, 2001.

[37] J.-D. Benamou, Y. Brenier, and K. Guittet. The Monge-Kantorovitch mass transfer and its computational fluid mechanics formulation. volume 40, pages 21–30. 2002. ICFD Conference on Numerical Methods for Fluid Dynamics (Oxford, 2001).

[38] Jean-David Benamou. Numerical resolution of an "unbalanced" mass transport problem. *M2AN Math. Model. Numer. Anal.*, 37(5):851–868, 2003.

[39] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[40] Jean-David Benamou and Guillaume Carlier. Augmented lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, 2015.

[41] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.*, 37(2):A1111–A1138, 2015.

[42] Jean-David Benamou, Guillaume Carlier, and Luca Nenna. A numerical method to solve multi-marginal optimal transport problems with Coulomb cost. In *Splitting methods in communication, imaging, science, and engineering*, Sci. Comput., pages 577–601. Springer, Cham, 2016.

[43] Jean-David Benamou, Guillaume Carlier, and Luca Nenna. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm. *Numer. Math.*, 142(1):33–54, 2019.

[44] Jean-David Benamou, Brittany D. Froese, and Adam M. Oberman. Numerical solution of the optimal transportation problem using the Monge-Ampère equation. *J. Comput. Phys.*, 260:107–126, 2014.

[45] Jean-David Benamou, Thomas O. Gallouët, and François-Xavier Vialard. Second-order models for optimal transport and cubic splines on the Wasserstein space. *Found. Comput. Math.*, 19(5):1113–1143, 2019.

[46] Georg Berschneider and Zoltán Sasvári. On a theorem of karhunen and related moment problems and quadrature formulae. In *Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equations*, pages 173–187. Springer, 2012.

[47] D. P. Bertsekas. The auction algorithm: a distributed relaxation method for the assignment problem. *Ann. Oper. Res.*, 14(1-4):105–123, 1988.

[48] Dimitri P Bertsekas and David A Castanon. The auction algorithm for the transportation problem. *Annals of Operations Research*, 20(1):67–96, 1989.

[49] Ugo Bindini. Smoothing operators in multi-marginal optimal transport. *Math. Phys. Anal. Geom.*, 23(2):Paper No. 21, 27, 2020.

[50] Ugo Bindini and Luigi De Pascale. Optimal transport with Coulomb cost and the semiclassical limit of density functional theory. *J. Éc. polytech. Math.*, 4:909–934, 2017.

[51] Adrien Blanchet and Guillaume Carlier. Optimal transport and Cournot-Nash equilibria. *Math. Oper. Res.*, 41(1):125–145, 2016.

[52] Guy Bouchitté, Giuseppe Buttazzo, Thierry Champion, and Luigi De Pascale. Dissociating limit in density functional theory with coulomb optimal transport cost. *arXiv preprint arXiv:1811.12085*, 2018.

[53] Guy Bouchitté, Giuseppe Buttazzo, Thierry Champion, and Luigi De Pascale. Relaxed multi-marginal costs and quantization effects. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*. Elsevier, 2020.

[54] David P Bourne, Bernhard Schmitzer, and Benedikt Wirth. Semi-discrete unbalanced optimal transport and quantization. *arXiv preprint arXiv:1808.01962*, 2018.

[55] Andrea Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[56] Y. Brenier. The dual least action problem for an ideal, incompressible fluid. *Arch. Rational Mech. Anal.*, 122(4):323–351, 1993.

[57] Yann Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math.*, 305:805–808, 1987.

[58] Yann Brenier. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Math. Soc.*, 2(2):225–255, 1989.

[59] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

[60] Yann Brenier. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math.*, 52(4):411–452, 1999.

[61] Yann Brenier. Generalized solutions and hydrostatic approximation of the Euler equations. *Phys. D*, 237(14-17):1982–1988, 2008.

[62] Yann Brenier, Uriel Frisch, Michel Hénon, Grégoire Loeper, Sabino Matarrese, Roya Mohayaee, and A Sobolevskiĭ. Reconstruction of the early universe as a convex optimization problem. *Monthly Notices of the Royal Astronomical Society*, 346(2):501–524, 2003.

[63] E Brezin and J Zinn-Justin. Expansion of the H_2^+ ground state energy in inverse powers of the distance between the two protons. *Journal de Physique Lettres*, 40(19):511–512, 1979.

[64] DM Brink and GR Satchler. *Angular momentum*. Oxford University Press, 1968.

[65] G. Buttazzo, C. Jimenez, and E. Oudet. An optimization problem for mass transportation with congested dynamics. *SIAM J. Control Optim.*, 48(3):1961–1976, 2009.

[66] Giuseppe Buttazzo, Thierry Champion, and Luigi De Pascale. Continuity and estimates for multimarginal optimal transportation problems with singular costs. *Appl. Math. Optim.*, 78(1):185–200, 2018.

[67] Giuseppe Buttazzo, Luigi De Pascale, and Paola Gori-Giorgi. Optimal-transport formulation of electronic density-functional theory. *Physical Review A*, 85(6):062502, 2012.

[68] Giuseppe Buttazzo, Aldo Pratelli, Sergio Solimini, and Eugene Stepanov. *Optimal urban networks via mass transportation*, volume 1961 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[69] Giuseppe Buttazzo and Filippo Santambrogio. A model for the optimal planning of an urban area. *SIAM J. Math. Anal.*, 37(2):514–530, 2005.

[70] Libor Buš and Pavel Tvrdík. Towards auction algorithms for large dense assignment problems. *Comput. Optim. Appl.*, 43(3):411–436, 2009.

[71] Luis Caffarelli, Mikhail Feldman, and Robert McCann. Constructing optimal maps for monge's transport problem as a limit of strictly convex costs. *Journal of the American Mathematical Society*, 15(1):1–26, 2002.

[72] Luis A Caffarelli. Interior w2, p estimates for solutions of the monge-ampere equation. *Annals of Mathematics*, pages 135–150, 1990.

[73] Luis A Caffarelli. A localization property of viscosity solutions to the monge-ampere equation and their strict convexity. *Annals of mathematics*, 131(1):129–134, 1990.

[74] Luis A Caffarelli. Some regularity properties of solutions of monge ampere equation. *Communications on pure and applied mathematics*, 44(8-9):965–969, 1991.

[75] Luis A. Caffarelli and Robert J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math. (2)*, 171(2):673–730, 2010.

[76] Eric Cancès, Rafaël Coyaud, and L Ridgway Scott. Van der waals interactions between two hydrogen atoms: The next orders. *arXiv preprint arXiv:2007.04227*, 2020.

[77] Eric Cancès, Mireille Defranceschi, Werner Kutzelnigg, Claude Le Bris, and Yvon Maday. Computational quantum chemistry: a primer. In *Handbook of numerical analysis, Vol. X*, Handb. Numer. Anal., X, pages 3–270. North-Holland, Amsterdam, 2003.

[78] Eric Cancès and L. Ridgway Scott. Van der Waals interactions between two hydrogen atoms: The Slater-Kirkwood method revisited. *SIAM Journal on Mathematical Analysis*, 50(1):381–410, 2018.

[79] P. Cardaliaguet, G. Carlier, and B. Nazaret. Geodesics for a class of distances in the space of probability measures. *Calc. Var. Partial Differential Equations*, 48(3-4):395–420, 2013.

[80] G. Carlier and I. Ekeland. The structure of cities. *J. Global Optim.*, 29(4):371–376, 2004.

[81] G. Carlier and I. Ekeland. Matching for teams. *Econom. Theory*, 42(2):397–418, 2010.

[82] Guillaume Carlier. Optimal transportation and economic applications. *Lecture Notes*, 2012.

[83] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.*, 49(2):1385–1418, 2017.

[84] Guillaume Carlier and Maxime Laborde. A differential approach to the multi-marginal Schrödinger system. *SIAM J. Math. Anal.*, 52(1):709–717, 2020.

[85] Guillaume Carlier and Bruno Nazaret. Optimal transportation for the determinant. *ESAIM Control Optim. Calc. Var.*, 14(4):678–698, 2008.

[86] Guillaume Carlier, Adam Oberman, and Edouard Oudet. Numerical methods for matching for teams and Wasserstein barycenters. *ESAIM Math. Model. Numer. Anal.*, 49(6):1621–1642, 2015.

[87] José Antonio Carrillo, Maria Pia Gualdani, and Giuseppe Toscani. Finite speed of propagation in porous media by mass transportation methods. *C. R. Math. Acad. Sci. Paris*, 338(10):815–818, 2004.

[88] M. A. Cebim, Mauro Masili, and J. J. De Groote. High precision calculation of multipolar dynamic polarizabilities and two- and three-body dispersion coefficients of atomic hydrogen. *Few-Body Systems*, 46(2):75–85, 2009.

[89] Thierry Champion and Luigi De Pascale. On the twist condition and c-monotone transport plans. *Discrete Contin. Dyn. Syst.*, 34(4):1339–1353, 2014.

[90] Thierry Champion, Luigi De Pascale, and Petri Juutinen. The ∞-Wasserstein distance: local solutions and existence of optimal transport maps. *SIAM J. Math. Anal.*, 40(1):1–20, 2008.

[91] Huajie Chen and Gero Friesecke. Pair densities in density functional theory. *Multiscale Modeling & Simulation*, 13(4):1259–1289, 2015.

[92] Huajie Chen, Gero Friesecke, and Christian B Mendl. Numerical methods for a kohn–sham density functional model based on optimal transport. *Journal of chemical theory and computation*, 10(10):4360–4368, 2014.

[93] Patrick Cheridito, Matti Kiiski, David J Prömel, and H Mete Soner. Martingale optimal transport duality. *Mathematische Annalen*, pages 1–28, 2020.

[94] Pierre-André Chiappori, Robert J. McCann, and Lars P. Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Econom. Theory*, 42(2):317–354, 2010.

[95] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fisher-Rao metrics. *Found. Comput. Math.*, 18(1):1–44, 2018.

[96] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Scaling algorithms for unbalanced optimal transport problems. *Math. Comp.*, 87(314):2563–2609, 2018.

[97] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: dynamic and Kantorovich formulations. *J. Funct. Anal.*, 274(11):3090–3123, 2018.

[98] T. C. Choy. Van der Waals interaction of the hydrogen molecule: An exact implicit energy density functional. *Phys. Rev. A*, 62(1):1–10, 2000.

[99] Giovanni Ciccotti, Tony Lelievre, and Eric Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 61(3):371–408, 2008.

[100] Jiří Cížek, Robert J Damburg, Sandro Graffi, Vincenzo Grecchi, Evans M Harrell, Johathan G Harris, Sachiko Nakai, Josef Paldus, Rafail Kh Propin, Harris J Silverstone, et al. $1/R$ expansion for H_2^+: Calculation of exponentially small terms and asymptotics. *Physical Review A*, 33(1):12, 1986.

[101] Sebastian Claici, Edward Chien, and Justin Solomon. Stochastic wasserstein barycenters. *arXiv preprint arXiv:1802.05757*, 2018.

[102] Maria Colombo, Luigi De Pascale, and Simone Di Marino. Multimarginal optimal transport maps for one–dimensional repulsive costs. *Canadian Journal of Mathematics*, 67(2):350–368, 2015.

[103] Maria Colombo and Simone Di Marino. Equality between Monge and Kantorovich multimarginal problems with Coulomb cost. *Ann. Mat. Pura Appl. (4)*, 194(2):307–320, 2015.

[104] Maria Colombo, Simone Di Marino, and Federico Stra. Continuity of multimarginal optimal transport with repulsive cost. *SIAM J. Math. Anal.*, 51(4):2903–2926, 2019.

[105] Maria Colombo and Federico Stra. Counterexamples in multimarginal optimal transport with Coulomb cost and spherically symmetric data. *Math. Models Methods Appl. Sci.*, 26(6):1025–1049, 2016.

[106] Codina Cotar, Gero Friesecke, and Claudia Klüppelberg. Density functional theory and optimal transportation with coulomb cost. *Communications on Pure and Applied Mathematics*, 66(4):548–599, 2013.

[107] Codina Cotar, Gero Friesecke, and Claudia Klüppelberg. Smoothing of transport plans with fixed marginals and rigorous semiclassical limit of the hohenberg–kohn functional. *Archive for Rational Mechanics and Analysis*, 228(3):891–922, 2018.

[108] Codina Cotar, Gero Friesecke, and Brendan Pass. Infinite-body optimal transport with coulomb cost. *Calculus of Variations and Partial Differential Equations*, 54(1):717–742, 2015.

[109] Codina Cotar and Mircea Petrache. Next-order asymptotic expansion for N-marginal optimal transport with Coulomb and Riesz costs. *Adv. Math.*, 344:137–233, 2019.

[110] Christopher J Cramer and Donald G Truhlar. Density functional theory for transition metals and transition metal chemistry. *Physical Chemistry Chemical Physics*, 11(46):10757–10816, 2009.

[111] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*, pages 2292–2300, 2013.

[112] Marco Cuturi and Arnaud Doucet. Fast computation of wasserstein barycenters. 2014.

[113] Dinh Dũng, Vladimir Temlyakov, and Tino Ullrich. *Hyperbolic cross approximation*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser/Springer, Cham, 2018. Edited and with a foreword by Sergey Tikhonov.

[114] Robert J Damburg, Rafail Kh Propin, Sandro Graffi, Vincenzo Grecchi, Evans M Harrell, Jiří Čížek, Josef Paldus, Harris J Silverstone, et al. $1/R$ expansion for H_2^+: Analyticity, summability, asymptotics, and calculation of exponentially small terms. *Physical review letters*, 52(13):1112, 1984.

[115] George B Dantzig. Maximization of a linear function of variables subject to linear inequalities. *Activity analysis of production and allocation*, 13:339–347, 1951.

[116] George B. Dantzig. *Linear programming and extensions*. Princeton University Press, Princeton, N.J., 1963.

[117] Fernando De Goes, David Cohen-Steiner, Pierre Alliez, and Mathieu Desbrun. An optimal transport approach to robust reconstruction and simplification of 2d shapes. In *Computer Graphics Forum*, volume 30, pages 1593–1602. Wiley Online Library, 2011.

[118] Hadrien De March. Entropic approximation for multi-dimensional martingale optimal transport. *arXiv preprint arXiv:1812.11104*, 2018.

[119] Hadrien De March and Nizar Touzi. Irreducible convex paving for decomposition of multidimensional martingale transport plans. *Ann. Probab.*, 47(3):1726–1774, 2019.

[120] Stefano De Marco and Pierre Henry-Labordère. Linking vanillas and VIX options: a constrained Martingale optimal transport problem. *SIAM J. Financial Math.*, 6(1):1171–1194, 2015.

[121] Luigi De Pascale. Optimal transport with Coulomb cost. Approximation and duality. *ESAIM Math. Model. Numer. Anal.*, 49(6):1643–1657, 2015.

[122] Jean-François Delmas and Benjamin Jourdain. *Modèles aléatoires*, volume 57 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006. Applications aux sciences de l'ingénieur et du vivant. [Applications to engineering and the life sciences].

[123] Simone Di Marino, Augusto Gerolin, and Luca Nenna. Optimal transportation theory with repulsive costs. In *Topological optimization and optimal transport*, volume 17 of *Radon Ser. Comput. Appl. Math.*, pages 204–256. De Gruyter, Berlin, 2017.

[124] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

[125] Yan Dolinsky and H. Mete Soner. Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Related Fields*, 160(1-2):391–427, 2014.

[126] D. C. Dowson and B. V. Landau. The Fréchet distance between multivariate normal distributions. *J. Multivariate Anal.*, 12(3):450–455, 1982.

[127] Lawrence C. Evans. Partial differential equations and Monge-Kantorovich mass transfer. In *Current developments in mathematics, 1997 (Cambridge, MA)*, pages 65–126. Int. Press, Boston, MA, 1999.

[128] Arash Fahim and Yu-Jui Huang. Model-independent superhedging under portfolio constraints. *Finance Stoch.*, 20(1):51–81, 2016.

[129] Albert Fathi and Alessio Figalli. Optimal transportation on non-compact manifolds. *Israel J. Math.*, 175:1–59, 2010.

[130] Sira Ferradans, Nicolas Papadakis, Gabriel Peyré, and Jean-François Aujol. Regularized discrete optimal transport. *SIAM J. Imaging Sci.*, 7(3):1853–1882, 2014.

[131] Sira Ferradans, Gui-Song Xia, Gabriel Peyré, and Jean-François Aujol. Static and dynamic texture mixing using optimal transport. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 137–148. Springer, 2013.

[132] Jean Feydy, Benjamin Charlier, François-Xavier Vialard, and Gabriel Peyré. Optimal transport for diffeomorphic registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 291–299. Springer, 2017.

[133] Jean Feydy, Thibault Séjourné, François-Xavier Vialard, Shun-ichi Amari, Alain Trouvé, and Gabriel Peyré. Interpolating between optimal transport and mmd using sinkhorn divergences. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 2681–2690, 2019.

[134] Alessio Figalli. The optimal partial transport problem. *Arch. Ration. Mech. Anal.*, 195(2):533–560, 2010.

[135] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *Journal de mathématiques pures et appliquées*, 94(2):107–130, 2010.

[136] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. *J. Math. Pures Appl. (9)*, 94(2):107–130, 2010.

[137] Alessio Figalli and Nicolas Juillet. Absolute continuity of wasserstein geodesics in the heisenberg group. *Journal of Functional Analysis*, 255(1):133–141, 2008.

[138] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Inventiones mathematicae*, 182(1):167–211, 2010.

[139] Alessio Figalli, Ludovic Rifford, and Cédric Villani. Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds. *Tohoku Math. J. (2)*, 63(4):855–876, 2011.

[140] Lindsay Forestell and Frank Marsiglio. The importance of basis states: an example using the hydrogen basis. *Canadian Journal of Physics*, 93(10):1009–1014, 2015.

[141] Gero Friesecke. A simple counterexample to the Monge ansatz in multi-marginal optimal transport, convex geometry of the set of Kantorovich plans, and the Frenkel-Kontorova model. *SIAM J. Math. Anal.*, 51(6):4332–4355, 2019.

[142] Gero Friesecke, Christian B Mendl, Brendan Pass, Codina Cotar, and Claudia Klüppelberg. N-density representability and the optimal transport limit of the hohenberg-kohn functional. *The Journal of chemical physics*, 139(16):164109, 2013.

[143] Gero Friesecke and Daniela Vögler. Breaking the curse of dimension in multi-marginal kantorovich optimal transport on finite state spaces. *SIAM Journal on Mathematical Analysis*, 50(4):3996–4019, 2018.

[144] Uriel Frisch, Sabino Matarrese, Roya Mohayaee, and Andrei Sobolevski. A reconstruction of the initial conditions of the universe by optimal mass transportation. *Nature*, 417(6886):260–262, 2002.

[145] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1):312–336, 2014.

[146] Alfred Galichon. A survey of some recent applications of optimal transport methods to econometrics. *The Econometrics Journal*, 20(2):C1–C11, 2017.

[147] Alfred Galichon. *Optimal transport methods in economics*. Princeton University Press, 2018.

[148] Thomas Gallouët and Quentin Mérigot. A lagrangian scheme for the incompressible euler equation using optimal transport. *arXiv preprint arXiv:1605.00568*, 2016.

[149] Wilfrid Gangbo. An elementary proof of the polar factorization of vector-valued functions. *Archive for rational mechanics and analysis*, 128(4):381–399, 1994.

[150] Wilfrid Gangbo and Robert J McCann. Optimal maps in monge's mass transport problem. *Comptes Rendus de l'Academie des Sciences-Serie I-Mathematique*, 321(12):1653, 1995.

[151] Wilfrid Gangbo and Robert J McCann. The geometry of optimal transportation. *Acta Mathematica*, 177(2):113–161, 1996.

[152] Wilfrid Gangbo and Andrzej Święch. Optimal maps for the multidimensional Monge-Kantorovich problem. *Comm. Pure Appl. Math.*, 51(1):23–45, 1998.

[153] Andre K Geim and Irina V Grigorieva. Van der Waals heterostructures. *Nature*, 499(7459):419–425, 2013.

[154] Aude Genevay, Marco Cuturi, Gabriel Peyré, and Francis Bach. Stochastic optimization for large-scale optimal transport. In *Advances in neural information processing systems*, pages 3440–3448, 2016.

[155] Aude Genevay, Gabriel Peyré, and Marco Cuturi. Learning generative models with sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617, 2018.

[156] Augusto Gerolin, Juri Grossi, and Paola Gori-Giorgi. Kinetic correlation functionals from the entropic regularization of the strictly correlated electrons problem. *Journal of Chemical Theory and Computation*, 16(1):488–498, 2019.

[157] Augusto Gerolin, Anna Kausamo, and Tapio Rajala. Duality theory for multimarginal optimal transport with repulsive costs in metric spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 25:62, 2019.

[158] Augusto Gerolin, Anna Kausamo, and Tapio Rajala. Nonexistence of optimal transport maps for the multimarginal repulsive harmonic cost. *SIAM J. Math. Anal.*, 51(3):2359–2371, 2019.

[159] Augusto Gerolin, Anna Kausamo, and Tapio Rajala. Multi-marginal entropy-transport with repulsive cost. *Calc. Var. Partial Differential Equations*, 59(3):Paper No. 90, 20, 2020.

[160] Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim. Structure of optimal martingale transport plans in general dimensions. *Ann. Probab.*, 47(1):109–164, 2019.

[161] Nassif Ghoussoub and Bernard Maurey. Remarks on multi-marginal symmetric Monge-Kantorovich problems. *Discrete Contin. Dyn. Syst.*, 34(4):1465–1480, 2014.

[162] Nassif Ghoussoub and Abbas Moameni. Symmetric Monge-Kantorovich problems and polar decompositions of vector fields. *Geom. Funct. Anal.*, 24(4):1129–1166, 2014.

[163] Nassif Ghoussoub and Brendan Pass. Decoupling of DeGiorgi-type systems via multi-marginal optimal transport. *Comm. Partial Differential Equations*, 39(6):1032–1047, 2014.

[164] Andrew V. Goldberg and Robert E. Tarjan. Finding minimum-cost circulations by successive approximation. *Math. Oper. Res.*, 15(3):430–466, 1990.

[165] Paola Gori-Giorgi and Michael Seidl. Density functional theory for strongly-interacting electrons: perspectives for physics and chemistry. *Physical Chemistry Chemical Physics*, 12(43):14405–14419, 2010.

[166] Paola Gori-Giorgi, Michael Seidl, and Giovanni Vignale. Density-functional theory for strongly interacting electrons. *Physical review letters*, 103(16):166402, 2009.

[167] Paola Gori-Giorgi, Giovanni Vignale, and Michael Seidl. Electronic zero-point oscillations in the strong-interaction limit of density functional theory. *Journal of chemical theory and computation*, 5(4):743–753, 2009.

[168] S Graffi, V Grecchi, EM Harrell II, and HJ Silverstone. The $1/R$ expansion for H_2^+: Analyticity, summability, and asymptotics. *Annals of Physics*, 165(2):441–483, 1985.

[169] Juri Grossi, Derk P Kooi, Klaas JH Giesbertz, Michael Seidl, Aron J Cohen, Paula Mori-Sánchez, and Paola Gori-Giorgi. Fermionic statistics in the strongly correlated limit of density functional theory. *Journal of chemical theory and computation*, 13(12):6089–6100, 2017.

[170] Juri Grossi, Michael Seidl, Paola Gori-Giorgi, and Klaas JH Giesbertz. Functional derivative of the zero-point-energy functional from the strong-interaction limit of density-functional theory. *Physical Review A*, 99(5):052504, 2019.

[171] Gaoyue Guo and Jan Obloj. Computational Methods for Martingale Optimal Transport problems. *arXiv e-prints*, page arXiv:1710.07911, Oct 2017.

[172] Gaoyue Guo, Xiaolu Tan, and Nizar Touzi. Optimal Skorokhod embedding under finitely many marginal constraints. *SIAM J. Control Optim.*, 54(4):2174–2201, 2016.

[173] Julien Guyon. The joint s&p 500/vix smile calibration puzzle solved. *Risk, April*, 2020.

[174] Julien Guyon, Romain Menegaux, and Marcel Nutz. Bounds for vix futures given s&p 500 smiles. *Finance and Stochastics*, 21(3):593–630, 2017.

[175] Eldad Haber, Tauseef Rehman, and Allen Tannenbaum. An efficient numerical method for the solution of the L_2 optimal mass transfer problem. *SIAM J. Sci. Comput.*, 32(1):197–211, 2010.

[176] Steven Haker, Lei Zhu, Allen Tannenbaum, and Sigurd Angenent. Optimal mass transport for registration and warping. *International Journal of computer vision*, 60(3):225–240, 2004.

[177] Valentin Hartmann and Dominic Schuhmacher. Semi-discrete optimal transport: a solution procedure for the unsquared euclidean distance case. *Mathematical Methods of Operations Research*, pages 1–31, 2020.

[178] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2013.

[179] Didier Henrion, Jean-Bernard Lasserre, and Johan Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. *Optimization Methods & Software*, 24(4-5):761–779, 2009.

[180] Pierre Henry-Labordère. *Model-free hedging: A martingale optimal transport viewpoint*. Chapman and Hall/CRC, 2017.

[181] Pierre Henry-Labordère, Xiaolu Tan, and Nizar Touzi. An explicit martingale version of the one-dimensional Brenier's theorem with full marginals constraint. *Stochastic Process. Appl.*, 126(9):2800–2834, 2016.

[182] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. (2)*, 136:B864–B871, 1964.

[183] Gao Huang, Chuan Guo, Matt J Kusner, Yu Sun, Fei Sha, and Kilian Q Weinberger. Supervised word mover's distance. In *Advances in Neural Information Processing Systems*, pages 4862–4870, 2016.

[184] Angelo Iollo and Damiano Lombardi. A Lagrangian scheme for the solution of the optimal mass transfer problem. *J. Comput. Phys.*, 230(9):3430–3442, 2011.

[185] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017.

[186] C. Jimenez. Dynamic formulation of optimal transport problems. *J. Convex Anal.*, 15(3):593–622, 2008.

[187] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[188] Benjamin Jourdain and Julien Reygner. Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(3):455–506, Sep 2013.

[189] Leonid Vitalievich Kantorovich. On the translocation of masses. In *Dokl. Akad. Nauk. USSR (NS)*, volume 37, pages 199–201, 1942.

[190] T. Kato. On the upper and lower bounds of eigenvalues. *Journal of the Physical Society of Japan*, 4(4-6):334–339, 1949.

[191] Hans G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984.

[192] Yuehaw Khoo, Lin Lin, Michael Lindsey, and Lexing Ying. Semidefinite relaxation of multimarginal optimal transport for strictly correlated electrons in second quantization. *SIAM J. Sci. Comput.*, 42(6):B1462–B1489, 2020.

[193] Yuehaw Khoo and Lexing Ying. Convex relaxation approaches for strictly correlated density functional theory. *SIAM Journal on Scientific Computing*, 41(4):B773–B795, 2019.

[194] Young-Heon Kim and Robert J. McCann. Continuity, curvature, and the general covariance of optimal transportation. *J. Eur. Math. Soc. (JEMS)*, 12(4):1009–1040, 2010.

[195] Young-Heon Kim and Brendan Pass. A general condition for Monge solutions in the multi-marginal optimal transport problem. *SIAM J. Math. Anal.*, 46(2):1538–1550, 2014.

[196] Young-Heon Kim and Brendan Pass. Multi-marginal optimal transport on Riemannian manifolds. *Amer. J. Math.*, 137(4):1045–1060, 2015.

[197] Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a newton algorithm for semi-discrete optimal transport. *arXiv preprint arXiv:1603.05579*, 2016.

[198] Jun Kitagawa and Brendan Pass. The multi-marginal optimal partial transport problem. In *Forum of Mathematics, Sigma*, volume 3. Cambridge University Press, 2015.

[199] Marcel Klatt, Carla Tameling, and Axel Munk. Empirical regularized optimal transport: Statistical theory and applications. *SIAM Journal on Mathematics of Data Science*, 2(2):419–443, 2020.

[200] Philip A. Knight. The Sinkhorn-Knopp algorithm: convergence and applications. *SIAM J. Matrix Anal. Appl.*, 30(1):261–275, 2008.

[201] Patrice Koehl, Marc Delarue, and Henri Orland. Optimal transport at finite temperature. *Physical Review E*, 100(1):013310, 2019.

[202] Soheil Kolouri, Se Rim Park, Matthew Thorpe, Dejan Slepcev, and Gustavo K Rohde. Optimal mass transport: Signal processing and machine-learning applications. *IEEE signal processing magazine*, 34(4):43–59, 2017.

[203] Soheil Kolouri, Akif B Tosun, John A Ozolek, and Gustavo K Rohde. A continuous linear optimal transport approach for pattern analysis in image datasets. *Pattern recognition*, 51:453–462, 2016.

[204] Derk P Kooi and Paola Gori-Giorgi. A variational approach to london dispersion interactions without density distortion. *The journal of physical chemistry letters*, 10(7):1537–1541, 2019.

[205] Mario Johannes Koppen et al. *Van der Waals forces in the context of non-relativistic quantum electrodynamics*. PhD thesis, PhD thesis, Technische Universität München, Fukultät für Mathematik, 2010.

[206] Jeffrey J Kosowsky and Alan L Yuille. The invisible hand algorithm: Solving the assignment problem with statistical physics. *Neural networks*, 7(3):477–490, 1994.

[207] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Res. Logist. Quart.*, 2:83–97, 1955.

[208] M. Laborde. Nonlinear systems coupled through multi-marginal transport problems. *European J. Appl. Math.*, 31(3):450–469, 2020.

[209] Giovanna Lani, Simone Di Marino, Augusto Gerolin, Robert van Leeuwen, and Paola Gori-Giorgi. The adiabatic strictly-correlated-electrons functional: kernel and exact properties. *Physical Chemistry Chemical Physics*, 18(31):21092–21101, 2016.

[210] Jean B Lasserre. A semidefinite programming approach to the generalized problem of moments. *Mathematical Programming*, 112(1):65–92, 2008.

[211] Jean-Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2010.

[212] Hugo Lavenant. Unconditional convergence for discretizations of dynamical optimal transport. *Mathematics of Computation*, 2020.

[213] Hugo Lavenant, Sebastian Claici, Edward Chien, and Justin Solomon. Dynamical optimal transport on discrete surfaces. *ACM Transactions on Graphics (TOG)*, 37(6):1–16, 2018.

[214] Charles L Lawson and Richard J Hanson. *Solving least squares problems*. SIAM, 1995.

[215] Arthur Leclaire and Julien Rabin. A stochastic multi-layer algorithm for semi-discrete optimal transport with applications to texture synthesis and style transfer. *Journal of Mathematical Imaging and Vision*, pages 1–27, 2020.

[216] Hugo Leclerc, Quentin Mérigot, Filippo Santambrogio, and Federico Stra. Lagrangian discretization of crowd motion and linear diffusion. *SIAM J. Numer. Anal.*, 58(4):2093–2118, 2020.

[217] Benedict Leimkuhler and Sebastian Reich. *Simulating hamiltonian dynamics*, volume 14. Cambridge university press, 2004.

[218] Tony Lelievre, Mathias Rousset, and Gabriel Stoltz. *Free energy computations: A mathematical perspective*. World Scientific, 2010.

[219] Tony Lelievre, Mathias Rousset, and Gabriel Stoltz. Langevin dynamics with constraints and computation of free energy differences. *Mathematics of computation*, 81(280):2071–2125, 2012.

[220] Tony Lelièvre, Mathias Rousset, and Gabriel Stoltz. Hybrid monte carlo methods for sampling probability measures on submanifolds. *Numerische Mathematik*, 143(2):379–421, 2019.

[221] Christian Léonard. From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.*, 262(4):1879–1920, 2012.

[222] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[223] Bruno Lévy. A numerical algorithm for L_2 semi-discrete optimal transport in 3D. *ESAIM Math. Model. Numer. Anal.*, 49(6):1693–1715, 2015.

[224] Bruno Lévy and Erica L Schwindt. Notions of optimal transport theory and how to implement them on a computer. *Computers & Graphics*, 72:135–148, 2018.

[225] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Nat. Acad. Sci. U.S.A.*, 76(12):6062–6065, 1979.

[226] Mathieu Lewin. Semi-classical limit of the Levy-Lieb functional in density functional theory. *C. R. Math. Acad. Sci. Paris*, 356(4):449–455, 2018.

[227] Mathieu Lewin, Elliott H. Lieb, and Robert Seiringer. Statistical mechanics of the uniform electron gas. *J. Éc. polytech. Math.*, 5:79–116, 2018.

[228] Mathieu Lewin, Elliott H Lieb, and Robert Seiringer. Universal functionals in density functional theory. In Éric Cancès, Gero Friesecke, and Lin Lin, editors, *Density Functional Theory*. 2019. arXiv preprint arXiv:1912.10424.

[229] Elliott H Lieb. Density functionals for coulomb systems. In *Inequalities*, pages 269–303. Springer, 2002.

[230] Elliott H Lieb and Walter E Thirring. Universal nature of van der Waals forces for Coulomb systems. *Physical Review A*, 34(1):40–46, 1986.

[231] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *SIAM J. Math. Anal.*, 48(4):2869–2911, 2016.

[232] Tianyi Lin, Nhat Ho, Marco Cuturi, and Michael I Jordan. On the complexity of approximating multimarginal optimal transport. *arXiv preprint arXiv:1910.00152*, 2019.

[233] Tianyi Lin, Nhat Ho, and Michael I Jordan. On efficient optimal transport: An analysis of greedy and accelerated mirror descent algorithms. *arXiv preprint arXiv:1901.06482*, 2019.

[234] Michael Lindsey and Yanir A. Rubinstein. Optimal transport via a Monge-Ampère optimization problem. *SIAM J. Math. Anal.*, 49(4):3073–3124, 2017.

[235] Grégoire Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.

[236] Grégoire Loeper and Francesca Rapetti. Numerical solution of the Monge-Ampère equation by a Newton's algorithm. *C. R. Math. Acad. Sci. Paris*, 340(4):319–324, 2005.

[237] Damiano Lombardi and Emmanuel Maitre. Eulerian models and algorithms for unbalanced optimal transport. *ESAIM Math. Model. Numer. Anal.*, 49(6):1717–1744, 2015.

[238] Fritz London. The general theory of molecular forces. *Transactions of the Faraday Society*, 33:8b–26, 1937.

[239] Ming Ma, Na Lei, Wei Chen, Kehua Su, and Xianfeng Gu. Robust surface registration using optimal mass transport and teichmüller mapping. *Graphical Models*, 90:13–23, 2017.

[240] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[241] Jan Maas, Martin Rumpf, Carola Schönlieb, and Stefan Simon. A generalized model for optimal transport of images including dissipation and density modulation. *ESAIM Math. Model. Numer. Anal.*, 49(6):1745–1769, 2015.

[242] John Michael L MacNeil, Dmitriy Morozov, Francesco Panerai, Dilworth Parkinson, Harold Barnard, and Daniela Ushizima. Distributed global digital volume correlation by optimal transport. In *2019 IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP)*, pages 14–19. IEEE, 2019.

[243] Francesc Malet and Paola Gori-Giorgi. Strong correlation in kohn-sham density functional theory. *Physical review letters*, 109(24):246402, 2012.

[244] Francesc Malet, André Mirtschink, Jonas C Cremon, Stephanie M Reimann, and Paola Gori-Giorgi. Kohn-sham density functional theory for quantum wires in arbitrary correlation regimes. *Physical Review B*, 87(11):115146, 2013.

[245] Francesc Malet, André Mirtschink, Klaas JH Giesbertz, and Paola Gori-Giorgi. Density functional theory for strongly-interacting electrons. In *Many-Electron Approaches in Physics, Chemistry and Mathematics*, pages 153–168. Springer, 2014.

[246] Francesc Malet, André Mirtschink, Klaas JH Giesbertz, Lucas O Wagner, and Paola Gori-Giorgi. Exchange–correlation functionals from the strong interaction limit of dft: applications to model chemical systems. *Physical Chemistry Chemical Physics*, 16(28):14551–14558, 2014.

[247] Hadrien De March and Pierre Henry-Labordere. Building arbitrage-free implied volatility: Sinkhorn's algorithm and variants. *Available at SSRN 3326486*, 2019.

[248] Simone Di Marino and Augusto Gerolin. An Optimal Transport Approach for the Schrödinger Bridge Problem and Convergence of Sinkhorn Algorithm. *J. Sci. Comput.*, 85(2):Paper No. 27, 2020.

[249] Bertrand Maury, Aude Roudneff-Chupin, and Filippo Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.

[250] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[251] Christian B Mendl and Lin Lin. Kantorovich dual solution for strictly correlated electrons in atoms and molecules. *Physical Review B*, 87(12):125106, 2013.

[252] Quentin Mérigot. A multiscale approach to optimal transport. In *Computer Graphics Forum*, volume 30, pages 1583–1592. Wiley Online Library, 2011.

[253] Quentin Mérigot and Édouard Oudet. Discrete optimal transport: complexity, geometry and applications. *Discrete Comput. Geom.*, 55(2):263–283, 2016.

[254] Liang Mi and José Bento. Multi-marginal optimal transport defines a generalized metric. *arXiv preprint arXiv:2001.11114*, 2020.

[255] André Mirtschink, Michael Seidl, and Paola Gori-Giorgi. Derivative discontinuity in the strong-interaction limit of density-functional theory. *Physical review letters*, 111(12):126402, 2013.

[256] André Mirtschink, CJ Umrigar, John D Morgan III, and Paola Gori-Giorgi. Energy density functionals from the strong-coupling limit applied to the anions of the he isoelectronic series. *The Journal of chemical physics*, 140(18):18A532, 2014.

[257] James Mitroy and Michael WJ Bromley. Higher-order C_n dispersion coefficients for hydrogen. *Physical Review A*, 71(3):032709, 2005.

[258] Abbas Moameni. Multi-marginal Monge-Kantorovich transport problems: a characterization of solutions. *C. R. Math. Acad. Sci. Paris*, 352(12):993–998, 2014.

[259] Abbas Moameni and Brendan Pass. Solutions to multi-marginal optimal transport problems concentrated on several graphs. *ESAIM Control Optim. Calc. Var.*, 23(2):551–567, 2017.

[260] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*, pages 666–704, 1781.

[261] John D Morgan III and Barry Simon. Behavior of molecular potential energy curves for large nuclear separations. *International journal of quantum chemistry*, 17(6):1143–1166, 1980.

[262] Luca Nenna. *Numerical methods for multi-marginal optimal transportation*. PhD thesis, PSL Research University, 2016.

[263] Luca Nenna and Brendan Pass. Variational problems involving unequal dimensional optimal transport. *J. Math. Pures Appl. (9)*, 139:83–108, 2020.

[264] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. 2001.

[265] Vitali D Ovsiannikov and J Mitroy. Regular approach for generating van der Waals C_s coefficients to arbitrary orders. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 39(1):159, 2005.

[266] Gilles Pagès. *Numerical Probability*. Springer, 2018.

[267] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. Optimal transport with proximal splitting. *SIAM J. Imaging Sci.*, 7(1):212–238, 2014.

[268] Robert G Parr. Density functional theory of atoms and molecules. In *Horizons of quantum chemistry*, pages 5–15. Springer, 1980.

[269] Brendan Pass. Uniqueness and Monge solutions in the multimarginal optimal transportation problem. *SIAM J. Math. Anal.*, 43(6):2758–2775, 2011.

[270] Brendan Pass. On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calc. Var. Partial Differential Equations*, 43(3-4):529–536, 2012.

[271] Brendan Pass. Optimal transportation with infinitely many marginals. *J. Funct. Anal.*, 264(4):947–963, 2013.

[272] Brendan Pass. Remarks on the semi-classical Hohenberg-Kohn functional. *Nonlinearity*, 26(9):2731–2744, 2013.

[273] Brendan Pass. Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions. *Discrete Contin. Dyn. Syst.*, 34(4):1623–1639, 2014.

[274] Brendan Pass. Multi-marginal optimal transport: theory and applications. *ESAIM Math. Model. Numer. Anal.*, 49(6):1771–1790, 2015.

[275] Linus Pauling and J. Y. Beach. The van der Waals interaction of hydrogen atoms. *Physical Review*, 47(9):686–692, May 1935.

[276] Gabriel Peyré. Entropic approximation of Wasserstein gradient flows. *SIAM J. Imaging Sci.*, 8(4):2323–2351, 2015.

[277] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[278] Federico Piazzon, Alvise Sommariva, and Marco Vianello. Caratheodory-tchakaloff subsampling. *Dolomites Research Notes on Approximation*, 10(1), 2017.

[279] Elijah Polak. *Optimization: algorithms and consistent approximations*, volume 124. Springer Science & Business Media, 1997.

[280] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 258(3):677–690, 2008.

[281] Aldo Pratelli. On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation. *Ann. Inst. H. Poincaré Probab. Statist.*, 43(1):1–13, 2007.

[282] Julien Rabin and Nicolas Papadakis. Convex color image segmentation with optimal transport distances. In *Scale space and variational methods in computer vision*, volume 9087 of *Lecture Notes in Comput. Sci.*, pages 256–269. Springer, Cham, 2015.

[283] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

[284] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. II*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Applications.

[285] Anthony Ralston and Philip Rabinowitz. *A first course in numerical analysis*. Courier Corporation, 2001.

[286] M Reed and B Simon. *Methods of Modern Mathematical Physics. Vol. 4. Operator Analysis*. Academic Press, New York, 1979.

[287] R Tyrrell Rockafellar. *Convex analysis*. Number 28. Princeton university press, 1970.

[288] Charles M Roth, Brian L Neal, and Abraham M Lenhoff. Van der Waals interactions involving proteins. *Biophysical Journal*, 70(2):977–987, 1996.

[289] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, 40(2):99–121, 2000.

[290] Giovanni Samaey, Tony Lelièvre, and Vincent Legat. A numerical closure approach for kinetic models of polymeric fluids: exploring closure relations for fene dumbbells. *Computers & fluids*, 43(1):119–133, 2011.

[291] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, pages 99–102, 2015.

[292] Filippo Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bulletin of Mathematical Sciences*, 7(1):87–154, 2017.

[293] Filippo Santambrogio. Crowd motion and evolution PDEs under density constraints. In *SMAI 2017—8^{e} Biennale Française des Mathématiques Appliquées et Industrielles*, volume 64 of *ESAIM Proc. Surveys*, pages 137–157. EDP Sci., Les Ulis, 2018.

[294] Louis-Philippe Saumier, Martial Agueh, and Boualem Khouider. An efficient numerical algorithm for the l 2 optimal transport problem with periodic densities. *The IMA Journal of Applied Mathematics*, 80(1):135–157, 2015.

[295] Morgan A Schmitz, Matthieu Heitz, Nicolas Bonneel, Fred Ngole, David Coeurjolly, Marco Cuturi, Gabriel Peyré, and Jean-Luc Starck. Wasserstein dictionary learning: Optimal transport-based unsupervised nonlinear dictionary learning. *SIAM Journal on Imaging Sciences*, 11(1):643–678, 2018.

[296] Bernhard Schmitzer. A sparse multiscale algorithm for dense optimal transport. *J. Math. Imaging Vision*, 56(2):238–259, 2016.

[297] Bernhard Schmitzer. Stabilized sparse scaling algorithms for entropy regularized transport problems. *SIAM J. Sci. Comput.*, 41(3):A1443–A1481, 2019.

[298] Bernhard Schmitzer, Klaus P Schäfers, and Benedikt Wirth. Dynamic cell imaging in pet with optimal transport regularization. *IEEE Transactions on Medical Imaging*, 39(5):1626–1635, 2019.

[299] Bernhard Schmitzer and Christoph Schnörr. A hierarchical approach to optimal transport. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 452–464. Springer, 2013.

[300] Bernhard Schmitzer and Christoph Schnörr. Globally optimal joint image segmentation and shape matching based on Wasserstein modes. *J. Math. Imaging Vision*, 52(3):436–458, 2015.

[301] L. Ridgway Scott. *Introduction to Automated Modeling with FEniCS*. Computational Modeling Initiative, 2018.

[302] Michael Seidl. Strong-interaction limit of density-functional theory. *Physical Review A*, 60(6):4387, 1999.

[303] Michael Seidl, Simone Di Marino, Augusto Gerolin, Luca Nenna, Klaas JH Giesbertz, and Paola Gori-Giorgi. The strictly-correlated electron functional for spherically symmetric systems revisited. *arXiv preprint arXiv:1702.05022*, 2017.

[304] Michael Seidl, Paola Gori-Giorgi, and Andreas Savin. Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Physical Review A*, 75(4):042511, 2007.

[305] Thibault Séjourné, Jean Feydy, François-Xavier Vialard, Alain Trouvé, and Gabriel Peyré. Sinkhorn divergences for unbalanced optimal transport. *arXiv preprint arXiv:1910.12958*, 2019.

[306] Meisam Sharify, Stéphane Gaubert, and Laura Grigori. Solution of the optimal assignment problem by diagonal scaling algorithms. *arXiv preprint arXiv:1104.3830v2*, 2013.

[307] Michael Shub. Some remarks on bezout's theorem and complexity theory. In *From topology to computation: proceedings of the smalefest*, pages 443–455. Springer, 1993.

[308] Michael Shub and Steve Smale. Complexity of bezout's theorem. *Collected Papers Of Stephen Smale, The (In 3 Volumes)-Volume 3*, 3:1349–1476, 2000.

[309] John C Slater and John G Kirkwood. The van der Waals forces in gases. *Physical Review*, 37(6):682, 1931.

[310] Justin Solomon, Fernando De Goes, Gabriel Peyré, Marco Cuturi, Adrian Butscher, Andy Nguyen, Tao Du, and Leonidas Guibas. Convolutional wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Transactions on Graphics (TOG)*, 34(4):1–11, 2015.

[311] Volker Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.

[312] Zhengyu Su, Yalin Wang, Rui Shi, Wei Zeng, Jian Sun, Feng Luo, and Xianfeng Gu. Optimal mass transport for shape matching and comparison. *IEEE transactions on pattern analysis and machine intelligence*, 37(11):2246–2259, 2015.

[313] Maria Tchernychova. *Carathéodory cubature measures*. PhD thesis, University of Oxford, 2016.

[314] Ajit J Thakkar. Higher dispersion coefficients: Accurate values for hydrogen atoms and simple estimates for other systems. *The Journal of Chemical Physics*, 89(4):2092–2098, 1988.

[315] Matthew Thorpe, Serim Park, Soheil Kolouri, Gustavo K. Rohde, and Dejan Slepčev. A transportation L^p distance for signal analysis. *J. Math. Imaging Vision*, 59(2):187–210, 2017.

[316] R J A Tough and A J Stone. Properties of the regular and irregular solid harmonics. *Journal of Physics A: Mathematical and General*, 10(8):1261–1269, aug 1977.

[317] Johannes Diderik Van der Waals. *On the continuity of the gaseous and liquid states. Edited and with an Introduction by J. S. Rowlinson.* Dover Phoenix Editions, 1988.

[318] Cédric Villani. *Topics in optimal transportation.* Number 58. American Mathematical Soc., 2003.

[319] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[320] Daniela Vögler. Kantorovich vs. monge: A numerical classification of extremal multi-marginal mass transports on finite state spaces. *arXiv preprint arXiv:1901.04568*, 2019.

[321] James Wan and Wadim Zudilin. Generating functions of Legendre polynomials: a tribute to Fred Brafman. *Journal of Approximation Theory*, 170:198–213, 2013.

[322] S. C. Wang. The mutual influence between the two atoms of hydrogen. *Physikalische Zeitschrift*, 28:663, 1927.

[323] Zong-Chao Yan and A Dalgarno. Third-order dispersion coefficients for H(1s)-H(1s) system. *Molecular Physics*, 96(5):863–865, 1999.

[324] Wei Zhang. Ergodic sdes on submanifolds and related numerical sampling schemes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 54(2):391–430, 2020.

[325] Grigorii M Zhislin. Discussion of the spectrum of Schrödinger operators for systems of many particles. *Trudy Moskovskogo Matematiceskogo Obscestva*, 9:81–120, 1960.

[326] Grigorii Moiseevich Zhislin. Finiteness of the discrete spectrum in the quantum n-particle problem. *Theoretical and Mathematical Physics*, 21(1):971–980, 1974.

Résumé : Le transport optimal (TO) a de nombreuses applications; mais son approximation numérique est complexe en pratique. Nous étudions une relaxation du TO pour laquelle les contraintes marginales sont remplacées par des contraintes de moments (TOCM), et montrons la convergence de ce dernier vers le problème OT. Le théorème de Tchakaloff nous permet de montrer qu'un minimiseur du problème TOCM est une mesure discrète chargeant un nombre fini de points, qui, pour les problèmes multimarginaux, est linéaire en le nombre de marginales, ce qui permet de contourner le fléau de la dimension. Cette méthode est aussi adaptée aux problèmes de TO martingale. Dans certains cas importants en pratique, nous obtenons des vitesses de convergence en $O(1/N)$ ou $O(1/N^2)$, où N est le nombre de moments, ce qui illustre leur rôle.

Nous présentons un algorithme, basé sur un processus de Langevin sur-amorti contraint, pour résoudre le problème TOCM. Nous prouvons que tout minimiseur local du problème TOCM en est un minimiseur global. Et illustrons l'algorithme sur des exemples de larges problèmes TOCM symétriques.

Dans la seconde partie de la thèse, nous étendons une méthode (E. Cancès et L.R. Scott, SIAM J. Math. Anal., 50, 2018, 381–410) pour calculer un nombre arbitraire de termes dans la série asymptotique de l'interaction de van der Waals entre deux atomes d'hydrogène. Ces termes sont obtenus en résolvant un ensemble d'EDP de Slater–Kirkwood modifiées. La précision de cette méthode est montrée par des exemples numériques et une comparaison avec d'autres méthodes issues de la littérature. Nous montrons aussi que les états de diffusion de l'atome d'hydrogène ont une contribution majeure au coefficient C_6 de la série de van der Waals.

Abstract: Optimal Transport (OT) problems arise in numerous applications. Numerical approximation of these problems is a practical challenging issue. We investigate a relaxation of OT problems when marginal constraints are replaced by some moment constraints (MCOT problem), and show the convergence of the latter towards the former. Using Tchakaloff's theorem, we show that the MCOT problem is achieved by a finite discrete measure. For multimarginal OT problems, the number of points weighted by this measure scales linearly with the number of marginal laws, which allows to bypass the curse of dimension. This method is also relevant for Martingale OT problems. In some fundamental cases, we get rates of convergence in $O(1/N)$ or $O(1/N^2)$ where N is the number of moments, which illustrates the role of the moment functions.

We design a numerical method, built upon constrained overdamped Langevin processes, to solve MCOT problems; and proved that any local minimizer to the MCOT problem is a global one. We provide numerical examples for large symmetrical multimarginal MCOT problems.

We extend a method (E. Cancès and L.R. Scott, SIAM J. Math. Anal., 50, 2018, 381–410) to compute more terms in the asymptotic expansion of the van der Waals attraction between two hydrogen atoms. These terms are obtained by solving a set of modified Slater–Kirkwood PDE's. The accuracy of the method is demonstrated by numerical simulations and comparison with other methods from the literature. We also show that the scattering states of the hydrogen atom (the ones associated with the continuous spectrum of the Hamiltonian) have a major contribution to the C_6 coefficient of the van der Waals expansion.