Lagrangian discretization of variational problems in Wasserstein spaces

Clément Sarrazin

université
PARIS-SACLAY

Lagrangian discretization of variational problems in Wasserstein spaces

Discrétisation Lagrangienne de problèmes variationnels dans des espaces de Wasserstein

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 574, École Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat: Mathématiques
Graduate School : GS Mathématiques, Référent : Faculté des sciences d'Orsay

Thèse préparée dans l' unité de recherche Université Paris-Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, 91405, Orsay, France,
sous la direction de Quentin MÉRIGOT, Professeur
et la co-direction de Filippo SANTAMBROGIO, Professeur

Thèse soutenue à Paris-Saclay, le 04 Janvier 2022, par

Clément SARRAZIN

Composition du jury

Yann Brenier Directeur de Recherche CNRS, Université Paris-Saclay	Président
Jérémie Bigot Professeur, Université de Bordeaux	Rapporteur
David Bourne Associate Professor, Heriot-Watt University	Rapporteur
Quentin Mérigot Professeur, Université Paris-Saclay	Directeur de thèse
Filippo Santambrogio Professeur, Université Claude Bernard	Directeur de thèse
Irène Waldspurger Chargée de Recherche CNRS, Université Paris-Dauphine	Examinatrice

THESE DE DOCTORAT

NNT : 2022UPASM001

Contents

Acknowledgments I

Abstract III

Notations V

Introduction 1

1 Optimal transport and discretization of congestion penalizations 13
1.1 Optimal transport and the semi-discrete setting 13
1.2 Moreau envelope in the Wasserstein space 23
1.3 Computations in the semi-discrete setting 34

2 Discretization of Variational Mean Field Games with Congestion 47
2.1 Variational mean field games in Lagrangian setting 50
2.2 Space discretization in $\mathcal{P}(\Gamma)$ 56
2.3 The fully discrete problem 65
2.4 Numerics 69

3 Optimal quantization of measures 75
3.1 The optimal quantization problem 75
3.2 Lloyd's algorithm for optimal uniform quantization 79
3.3 Numerical results 92
3.4 Applications to the initialization of some particle flows 98

4 Lagrangian critical measures 107
4.1 Critical points for the quantization energy 107
4.2 Lagrangian critical measures for W_2^2 112

A Measure theory 125
A.1 Narrow convergence of measures 126
A.2 Absolutely continuous measures and the disintegration theorem . 127

A.3 The Hausdorff measure for lower dimensional sets 128

B Convex analysis **129**
B.1 Convex duality: 129
B.2 Different kinds of convexity 132

Acknowledgements

Je souhaiterais commencer par remercier mes amis Mathilde et Raphaël pour leur soutien, non seulement pendant ces trois années mais en fait depuis notre rencontre durant mon année d'agreg. Bien entendu, cette section de ma thèse ne constitue en aucun cas un classement des personnes citées par ordre d'importance ou d'appréciation. Cela étant dit, quand je pense aux personnes grâce à qui j'ai pu surmonter les années difficiles que je viens de passer, c'est à vous que je pense en premier et sans hésitation.

Je remercie, bien sûr, mes directeurs de thèse, Quentin et Filippo, pour avoir accepté de travailler avec moi ces trois années de thèse, ainsi que pour leur aide et nombreux conseils. Merci de n'avoir jamais montré de frustration dans mes moments les moins glorieux, et merci en particulier Filippo pour ton constant enthousiasme et tes suggestions pertinentes lors de nos discussions.

Thank you David Bourne and Jérémie Bigot for accepting to report on my manuscript and for the stimulating remarks you made on its content.

Merci à mes amis pour leur présence dans mes moments les plus sombres et pour les nombreuses (et au combien nécessaires) distractions pendant trois années de lourd travail. Merci aux "amis de l'agreg", Mathieu et Véronique, Didier et Julie, Paul et Marc, Camille et Isabelle, mais également Guillaume, Samra ainsi que Batoul pour tant de moments passés ensemble (et également un peu d'aide en mathématiques). Merci aux "amis de la colloc", Thomas, Antonin, Pierre-Louis, Régis, Tanguy et Camille, Quentin (pas les mêmes que plus haut) pour m'avoir immédiatement accueilli dans votre groupe et pour les nombreuses soirées, excursions et "tan saos" qui sont subséquemment venues agrémenter mes années de thèse (et beaucoup d'aide en ML). Merci enfin aux "amis-collègues", Nhi, Guillaume et Hugo pour avoir été des cobureaux en or, Elise, Etienne, Jules, Nicolas et Gwenaël pour nos nombreuses discussions (mathématiques et autres) et vos encouragements!

Je tiens également à remercier (c'est presque fini, promis) mes enseignants de sciences qui ont su encourager ma passion pour les mathématiques et la dé-

marche scientifique. Merci enfin à mes collègues du laboratoire d'Orsay, ainsi bien sûr qu'à nos fantastiques (ment patientes) gestionnaires d'équipe, Valérie, Ophélie et Estelle, pour leurs nombreux conseils et leur aide (en mathématiques comme pour l'administratif) ces trois années.

Je terminerai cette section en remerciant ma famille pour leurs encouragements, et leur aide (quand ils le pouvaient, aider un thésard n'est pas toujours aisé). Merci Papa, Maman, Adèle, Véronique, Denis, Camille et Delphine, vous avez contribué à cette thèse bien plus que vous ne le savez (et également un peu à ma cuisine)!

Abstract

In this thesis, we study the discretization of variational problems, via semi-discrete optimal transport methods. Although these techniques yield much simpler expressions for their solutions, they also introduce non-convex terms in what were convex problems before discretization.

The main model is that of a crowd, which minimizes a global energy accumulated during its motion. This evolution is modeled by a variational mean field game problem, which energy features a term penalizing highly congested area. We approximate the solutions using the trajectories of a finite set of particles, however, such finite crowds are not admissible for this problem, due to the congestion term. In order to circumvent this issue, we consider a similar energy, however one where the congestion is penalized by a regularized version of the previous term. Our main result is then the convergence of these discrete minimizers towards a solution of the mean field game problem. However this statement requires the selection of an appropriate sequence of regularization parameters, which cannot be determined just yet. This is only partially an impediment, as the discretization used is very robust, and even for relatively large values of these parameters, the numerical experiments showcase a satisfying behavior regarding the modeled phenomenon.

The regularization of the congestion term is defined as a Moreau envelope, using the 2-Wasserstein distance. These expressions introduce a non-convexity in the discretized problem which could make numerically computing their solutions difficult. In order to better understand these issues, we gather insightful information regarding the structure of critical points for the simpler but related non-convex problem of uniform optimal quantization. These critical configurations and in particular, their limits as measures when the discretization becomes more and more precise, remain a vastly unexplored part of these optimization problems.

Notations

We make use of the following notations throughout this document:

- Ω is a compact smooth subset of $\mathbb{R}^d$ (a domain) which has non-empty interior.
- $\lesssim$ is an inequality which is true up to a multiplicative (positive) constant which does not depend on the significant parameters in the quantities that are compared (which should be obvious from the context).
- $\int(X)$ is the interior of the set X.
- if $1 \leq p < \infty$, $\mathbb{L}^p(X)$ is the set of classes (up to the equivalence relationship of being equal Lebesgue almost-everywhere) of real functions whose p-th power is Lebesgue-integrable on X. If $p = \infty$ it is the set of such classes of functions which are (essentially) bounded on X.
- Γ will always be the Polish space $\mathcal{C}^0([0;T],\mathbb{R}^d)$ endowed with the uniform convergence norm and T, d can be fixed for the entirety of the document.
- $\mathcal{M}(X)$ is the space of finite (signed) Radon measures on the Polish space X.
- $\mathcal{M}_+(X)$ is the subspace of all positive measures in $\mathcal{M}(X)$.
- $\mathcal{P}(X)$ is the subspace of probability measures in $\mathcal{M}(X)$ (def: Appendix A).
- δ_x is the Dirac measure at $x \in X$, which gives mass 1 to $\{x\}$ and 0 to any set not containing x.
- $\mathrm{spt}(\rho)$ is the support of the measure ρ.
- $T\#\rho$ is the push-forward of the measure ρ along the Borel map T.
- $\rho \perp \mu$ means that ρ and μ are mutually singular (def: Appendix A).

- $\mu \ll \rho$ means that μ is absolutely continuous with respect to ρ (def: Appendix A).
- $\mathcal{H}^k$ is the k-dimensional Hausdorff measure (on $\mathbb{R}^d$) (def: Definition 13).
- For a set $A \subset \mathbb{R}^d$, $|A|$ is the Lebesgue measure of the set. For a real number a, $|a|$ is still the usual absolute value of a.
- $\mathcal{P}_N(X)$ is the space of uniform discrete probability measures on X with at most N points in their support (def: Appendix A).
- $\Pi(\rho, \mu)$ is the set of transport plans from ρ to μ (def: Definition 1).
- $\mathcal{C}^c(X, Y)$ is the set of admissible dual potentials for an optimal transport with cost c (def: Section 1.1, (1.4)).
- $\mathrm{dom}(F)$ is the domain of the convex function F (def: Definition 14)
- F^* is the Legendre transform of the convex function F (def: Definition 16).
- $\partial F(x)$ is the subgradient of the convex function F at x (def: Definition 15). On the other hand, $\partial^+ F(x)$ is the supergradient of the concave function F at x.
- χ_C is the convex indicator function of the convex set C (def: Appendix B).
- ϕ^c is the c-transform of the continuous, bounded function ϕ (def: Definition 2). ϕ^{cc} is the bi-transform $(\phi^c)^c$ of ϕ.
- $\mu_n \xrightarrow[n\to\infty]{} \mu$ means that the sequence $(\mu_n)_{n\in\mathbb{N}}$ narrowly converges towards μ (def: Appendix A).
- $\mathcal{P}_p(X)$ is the set of probability measures with finite p-th order moment on X (def: Definition 3).
- W_p is the p-Wasserstein distance on $\mathcal{P}_p(X)$ (X is clear from the context).
- $\mathrm{Lag}_i(Y, \Phi)$ is the i-th Laguerre cell associated to the point cloud $Y = (y_1, \ldots, y_N)$ and the weights $\Phi = (\phi_1, \ldots, \phi_N)$ (def: Definition 6).
- $\mathrm{Lag}_i(Y)$ is the i-th optimal Laguerre cell for the optimal transport between μ uniform supported on the point cloud Y and a reference probability measure ρ. This Laguerre cell is also written Lag_i when there can be no confusion on the point cloud considered.

- $B(Y) = (b_1(Y), \dots, b_N(Y))$ is the point cloud obtained by taking the ρ-barycenters (in the same order) of each optimal Laguerre cell $\mathrm{Lag}_i(Y)$ (in the optimal transport towards ρ).
- $\mathcal{AC}([0;T], X)$ is the set of absolutely continuous curves $[0;T]$ valued in X (def: Definition 7).
- $\mathbb{H}^s$ is the Sobolev space $\mathbb{H}^s([0;T], \mathbb{R}^d)$, again, for a T and d that can be fixed at the beginning of the document.
- $W_{\mathbb{H}^s}$ is the 2-Wasserstein distance on $\mathcal{P}_2(\mathbb{H}^s)$ associated with the Sobolev norm.
- $\Gamma_\delta^{\mathrm{lin}}$ is the subspace of Γ constituted of piecewise affine fonctions on the $([i\delta; (i+1)\delta])_i$ of $[0;T]$.

Introduction

Le point de départ de cette thèse est la description du mouvement d'une foule en utilisant un nombre fini d'individus. Ce mouvement est effectué de manière à éviter autant que possible l'attroupement d'individus (la congestion) en un même endroit de l'espace. Un bon exemple d'un tel mouvement est donné par l'évacuation d'une salle, dans laquelle un incendie (ou autre péril) s'est déclaré. Les individus ont alors tous un but, par exemple atteindre l'issue de secours, mais une ruée générale vers cet objectif peut engendrer un blocage, lorsqu'un trop grand nombre d'individu tente de passer au même endroit afin de sortir plus vite. Ce genre de congestion s'interprète mathématiquement très facilement lorsque la population est tellement nombreuse que sa répartition peut être assimilée à celle d'une densité sur le domaine d'évolution (en l'occurence, la salle). Dans notre exemple d'incendie, pour ce genre de foules "infinie", la contrainte se traduit par le fait que la densité μ de la foule ne doit jamais dépasser 1. En revanche, dans le cas d'une foule discrète (et donc non-fictive), le modèle mathématique cesse d'être adapté, et de manière plus générale, exprimer la contrainte de non-congestion est un problème n'acceptant pas de solution immédiate. Mentionnons en particulier qu'il ne semble pas satisfaisant d'imposer seulement la condition que deux individus infinitésimaux ne se croisent jamais. Un modèle similaire fut proposé par Maury et al. [47] dans lequel les individus sont représentés par des sphères dures dont le rayon tend vers 0 et il fut déjà observé dans ce cas que les dynamiques observées à la limite ne correspondent pas à la contrainte de borne supérieure sur la densité.

Mouvement de foule avec congestion comme un problème variationnel Le mouvement de foule est décrit comme la minimisation d'une énergie durant le mouvement des individus dans un domaine. Un des termes de cette énergie somme toutes les contributions infinitésimales des endroits où la concentration de la population est trop importante, donnant une (trop) grande énergie aux population s'agglomérant trop durant le mouvement. Ce terme de "congestion" est habituellement accompagné d'un terme évaluant (en un sens) la distance des individus par rapport à un but, lequel peut changer au cours du

temps. Le mouvement résultant cherche à ménager la chèvre et le chou en essayant d'amener tous les individus vers leur but, tout en ne les concentrant pas trop dans une même zone. La construction du mouvement lui-même, et en particulier sa discrétisation en temps dans les exemples numériques, peut ensuite se faire de plusieurs manières: La façon la plus intuitive, peut-être, étant donné la nature du problème étudié, est d'effectuer une descente de gradient sur l'énergie (terme de "congestion" + terme d' "objectif") dans l'espace des mesures de probabilités. Ces flots gradients pour la métrique de Wasserstein ont été introduits par Jordan, Kinderlehrer et Otto ([34]) comme une réécriture d'équations d'advection-diffusion, et les applications de ces notions pour des mouvements de foules avec congestion sont étudiées dans Maury et al [46], et Mérigot et al [40]. Un autre point de vue, qui est celui que nous adoptons dans le chapitre 2, s'est développé avec l'introduction des jeux à champs moyens en 2006 (Lasry et Lions, [37], [38] et Caines, Huang et Malhamé, [33]), et, dans notre cas, de leur formulation variationnelle (voir Cardaliaguet, [14], Santambrogio, [64], et [8]). Ces jeux à champs moyens "non-congestionnés" sont étudiés dans Cardaliaguet, Mészáros et Santambrogio, [15]), et modélise l'évolution de la foule comme la minimisation d'une énergie, globale sur toute la durée du mouvement cette fois-ci, la population étant représentée par une densité de probabilité evoluant continument dans le temps. Remarquons ici que ces deux modèles sont loin d'être équivalents en général. Le modèle "flot gradient" favorisera un comportement peu prédictif pour sa population, abouttisant souvent à des comportement très égoistes à l'échelle des individus qui n'anticipent pas la trajectoire possible de leurs congénères. Par opposition, la foule modélisée par le jeu à champ moyen cherche un mouvement minimisant de façon globale à la fois en temps et espace, et la trajectoire totale des individus est donc planifiée en fonction de celle des autres individus, de manière à minimiser l'énergie accumulée par toute la foule (le mécontentement global). Le second modèle n'est pas nécessairement "meilleur" parce qu'il implémente cette capacité d'un individu à anticiper le mouvement des autres, dans le sens où il décrit le mouvement d'une foule suivant idéalement les consignes (d'évacuation dans notre métaphore initiale) et une bonne modélisation d'une foule "réaliste" se trouverait probablement à mi-chemin de ces deux modèles mathématiques.

Afin de permettre la résolution numérique de ces problèmes (qui sont posés, par essence, en dimension infinie), la distribution de la population doit être discrétisée. Dans la littérature (Achdou et Capuzzo-Dolcetta [1], Briceno-Arias, Kalise et Silva [13]), cela est souvent fait en approximant sa densité par une fonction plus simple définie par morceau, et pour ce faire, les équations décrivant l'évolution de la foule (en particulier la fameuse équation de continuité) doivent être adaptées afin de conserver la régularité du mouvement. Au lieu de cette

discrétisation "Eulérienne", gardant des mesures à densité en chaque temps, nous cherchons des approximations par des mesures discrètes en définissant un problème variationnel similaire au problème de jeux à champs moyen initial, mais posé sur des populations ayant un nombre fini d'individus. Cette discrétisation est qualifiée de "Lagrangienne" car elle se fait au niveau des individus de la population, et non de la proportion de ces individus en chaque point du domaine. L'équation de continuité n'a alors pas à être changée, et se simplifie même en un système d'EDO. L'énergie à minimiser en revanche doit être régularisée, afin d'être toujours bien définies pour des distributions de population discrètes (qui sont, pour les termes de congestion utilisés, infiniment congestionnées). Le but est ensuite de montrer que lorsque les paramètres de régularisation sont pris très petits et que la population compte un grand nombre d'individus, on retrouve, de façon approchée, la répartition de la population infinie minimisant le problème de jeux à champ moyen.

Transport optimal et approximation discrètes de mesures La théorie du transport optimal joue un rôle central dans cette thèse permettant à la fois de définir une notion de continuité pour les déplacements d'individus mentionné dans le paragraphe précédent, mais fournissant également une métrique selon laquelle juger de la qualité de nos approximations. Au coeur de cette théorie (voir Villani, [67], [68], Santambrogio, [63] et Peyré et Cuturi [56]), réside la notion de comparer deux mesures en terme de la quantité d'énergie (ou du coût) à dépenser pour déplacer la masse de l'une (un tas de sable dans la métaphore originelle de Monge) vers celle de l'autre (un trou à remplir). Cette comparaison "horizontale" contraste avec celles plus standards, "verticales", induite par une norme $\mathbb{L}^p$ sur les éventuelles densités de ces mesures, ou de variation totale, par le fait que la "distance" entre les mesures est calculée suivant les déplacements de matière dans l'espace ambiant, et non les différences de quantité de matière en chaque point, permettant une prise en compte plus fine de la géométrie du domaine. En particulier, elle permet de définir une notion de distance et de continuité qui sont bien adaptées à la modélisation de foules, comme décrite dans le paragraphe précédent (ces faits sont bien connus, depuis les écrits fondateurs de Lasry et Lions et Jordan, Kinderlehrer et Otto).

Dans notre cas, cette notion de distance nous permet également de définir une version régularisée de la pénalisation de congestion d'une mesure apparaissant dans nos problèmes de jeux à champs moyens, laquelle est bien définie, même dans les cas où la population étudiée n'est pas distribuée selon une mesure à densité. Essentiellement (et même si ce n'est pas rigoureusement vrai), cette nouvelle fonction pénalise maintenant la distance de transport optimal de la mesure discrète à une mesure peu congestionnée. En revanche, le nouveau

problème "discret" perd la convexité dont profitait son parent "continu", et une étude des points critiques possibles s'impose.

Notion Lagrangienne de point critique pour la quantification optimale. Le problème d'approximation du mouvement non congestionné d'une foule par celui d'une nombre fini d'individus possède de nombreuses similarités évidentes avec celui d'approximer (au sens de la distance de Wasserstein) une densité de probabilité par une distribution discrète et uniforme d'un nombre fini de particules. Par certains aspect ce dernier problème peut être vu comme une coupe à un temps fixé du premier, et il est raisonnable d'espérer tirer des informations utiles de l'étude de ces questions dans un cadre "plus simple". Le second problème est connu dans la littérature sous le nom de quantification optimale d'une mesure (uniforme dans notre cas), et a été étudié sous plusieurs formes (notamment, avec l'énoncé que nous considérons, par Balzer et al. [7], De Goes, Breeden, Ostromoukhov and Desbrun [17]). Nous nous concentrons sur l'approximation par minimisation de la distance de Wasserstein entre notre densité et une mesure discrète uniforme, les positions des masses de Dirac étant les inconnues de ce problème. Ce problème décrit, en un certain sens, une version statique en temps du problème de jeux à champs moyen étudiée dans le Chapitre 2 et fut introduit initialement par Balzer et al. comme une alternative faisant intervenir les cellules de Laguerre, communes en transport optimal semi-discret, au lieu des méthodes basées sur les cellules de Voronoï, alors communément utilisées. La fonctionnelle (en les positions des masses de Dirac) minimisée est alors non-convexe et exhibe des points critiques ne correspondant pas du tout à des échantillonages satisfaisant pour la mesure à densité sous-jacente. On notera cependant que l'expérience suggère qu'il est impossible de rester bloquer sur un de ces points critiques "dégénérés" lors de la minimisation.

Modèle et énoncés mathématiques: Nous commençons par étudier un problème de mouvement de foules représenté par une mesure de probabilité sur l'ensemble des trajectoires possibles, $\Gamma := \mathcal{C}^0([0;T],\mathbb{R}^d)$:

$$Q \in \mathcal{P}(\Gamma)$$

Nous considérons le problème de minimisation suivante, parmi les mesures de probabilités sur Γ:

$$(\mathcal{M}_{\mu^0}) : \inf \left\{ J(Q) \;\middle|\; Q \in \mathcal{P}(\Gamma) \text{ s.t. } e_0 \# Q = \mu^0 \right\}$$

où l'on minimise l'énergie globale:

$$J(Q) := \int_\Gamma L(\gamma') \mathrm{d}Q + \int_0^T F(e_t \# Q) dt + G(Q).$$

Le premier terme est un terme cinétique $L(\gamma') = \int_0^T L(\gamma'(t))dt$ pénalisant les vitesses trop grande le long des trajectoires. Le second terme, faisant intervenir F, pénalise des valeurs trop larges de la densité de la population au temps t, $e_t\#Q$(et interdit la présence de singularités dans ces mesures). G, enfin, est une fonction continue pour la convergence étroite de mesures dans $\mathcal{P}(\Gamma)$, donnant un but aux individus de la foule représentée par Q sous la forme d'une valeur minimale à atteindre par Q. Mentionnons par exemple le terme "potentiel" $G(Q) = \int_\Omega \phi \mathrm{d}e_T\#Q$ qui encouragera les individus à terminer leur trajectoires aux points de Ω où le potentiel ϕ est le plus bas.

Nous qualifions ces problèmes de Jeux à champs moyen variationnels, du fait que l'énergie J peut être vue comme une énergie minimisée par les joueurs d'un jeu à champs moyen mis sous forme variationnelle dans le sens où une mesure de probabilité $Q \in \mathcal{P}(\Gamma)$ minimisant $(\mathcal{M}_{\mu^0})$ induit une stratégie mixte d'équilibre pour les participant à un tel jeu. Notre but dans l'étude de ces problèmes était d'approcher leurs solutions par des mesures de probabilité discrètes sur Γ,

$$Q_N \in \mathcal{P}_N(\Gamma) := \left\{ \frac{1}{N}\sum_{i=1}^N \delta_{\gamma_i} \;\middle|\; (\gamma_1, \dots, \gamma_N) \in \Gamma^N \right\}$$

Pour ce faire, nous cherchons à construire Q_N comme un minimiseur pour une énergie similaire à J, mais optimisée sur l'espace discret $\mathcal{P}_N(\Gamma)$. La fonctionnelle J elle-même, ne convient en revanche pas pour cette construction, car le terme de congestion F pourrait (et dans la plupart des cas va) être $+\infty$ sur les mesures discrètes, ne nous donnant aucune information sur notre éventuelle proximité à un minimiseur de J. Nous nous intéressons donc au problème (discret en espace):

$$(\mathcal{M}_{N,\mu_N^0,\varepsilon_N}) : \inf \left\{ J_{\varepsilon_N}(Q_N) \mid Q_N \in \mathcal{P}_N(\Gamma),\ e_0\#Q_N = \mu_N^0 \right\},$$

où l'énergie J_{ε_N} a été redéfinie afin d'être finie pour des mesures discrètes:

$$J_{\varepsilon_N}(Q) := \int_\Gamma L(\gamma')dQ(\gamma) + \int_0^T F_{\varepsilon_N}(e_t\#Q)dt + G(Q)$$

et la mesure de départ est mainentant également une mesure discrète uniforme dans $\mathcal{P}_N(\mathbb{R}^d)$. La suite de paramètre $(\varepsilon_N)_{N\in\mathbb{N}}$ devra tendre vers 0 quand N tend vers l'infini afin d'obtenir une bonne approximation à la limite.

Le terme pénalisant la congestion, $F(e_t\#Q)$ a été remplacé par une version régularisée, que nous appelons par la suite l'enveloppe de Moreau de F (par analogie avec une définition similaire dans un espace de Hilbert):

$$\text{Si } \mu \in \mathcal{P}(\mathbb{R}^d),\ F_\varepsilon(\mu) := \inf_{\rho \in \mathcal{M}(\Omega)} \frac{W_2^2(\rho,\mu)}{2\varepsilon} + F(\rho)$$

Les propriétés de cette enveloppe sont étudiées dans le premier chapitre. Dans les cas que nous étudions, les mesures construites sont discrètes et F est donnée par une intégrale:

$$F(\rho) = \begin{cases} \int_\Omega f(\rho(x))dx & \text{if } \rho \ll dx \\ +\infty & \text{otherwise.} \end{cases}$$

Le calcul des valeurs de F_ε se fait alors en utilisant son expression duale, similaire à celle obtenue pour un problème de transport optimal:

$$F_\varepsilon(\mu) = \max_{\Phi \in \mathbb{R}^N} \sum_{i=1}^N \left[\frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} f^* \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) dx \right]$$

pour $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$, où $(\mathrm{Lag}_i(Y,\Phi))_i$ sont les fameuses cellules de Laguerre (utilisées notamment en transport optimal semi-discret, [55]):

$$\mathrm{Lag}_i(Y,\Phi) := \{x \in \Omega,\ c(x,y_i) - \phi_i \leq c(x,y_j) - \phi_j \text{ for } j = 1,\dots,N\}$$

L'expression maximisée admet une Hessienne (en $\phi_1,\dots,\phi_N$) qui est inversible sous une condition d'aire non nulle sur les cellules de Laguerre $\mathrm{Lag}_i(Y,\Phi)$, permettant d'approcher efficacement ses solutions via un algorithme de Newton.

Dans le Chapitre 2, nous montrons un résultat de convergence en un certain sens (Proposition 17, Chapter 2) du problème discret en espace $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, vers le problème continu $(\mathcal{M}_{\mu^0})$. Cette convergence est très similaire à la fameuse Γ-convergence, et implique en particulier que les minimiseurs pour le problème discret en espace $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$ convergent étroitement (à une sous-suite éventuelle près) vers un minimiseur du problème continu $(\mathcal{M}_{\mu^0})$, quand N tend vers l'infini. Ce résultat requiert cependant une condition sur la décroissance de ε_N vers 0. Plus précisément, cette décroissante doit se faire de façon à dominer à la fois la vitesse à laquelle μ_N^0 approxime μ^0 au temps initial, c'est-à-dire que:

$$W_2^2(\mu_N^0,\mu^0) = o_{N\to\infty}(\varepsilon_N)$$

et la vitesse à laquelle une solution $Q_{\min}$ de $(\mathcal{M}_{\mu^0})$ est approchée par une mesure de $\mathcal{P}_N(\mathbb{H}^s([0;T],\mathbb{R}^d))$, pour un certain espace $\mathbb{H}^s$ ad-hoc:

$$\text{Si } \tau_N := \min\left\{ W_2^2(\tilde{Q},Q_{\min}) \;\middle|\; \tilde{Q} \in \mathcal{P}_N(\mathbb{H}^s([0;T],\mathbb{R}^d)) \right\} \text{ alors, il faut que:}$$

$$\tau_N = o_{N\to\infty}(\varepsilon_N).$$

Sous ces deux conditions de domination, la convergence mentionnée plus haut a lieu.

Enfin, dans les simulations numériques, nous considérons également le problème complètement discrétisé, à la fois en espace et en temps:

$$(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N}) : \inf \left\{ J_{\delta_N,\varepsilon_N}(Q) \;\middle|\; Q \in \mathcal{P}_N(\Gamma_{\delta_N}^{\mathrm{lin}}),\ e_0 \# Q = \mu_N^0 \right\}$$

Ici, δ_N est un pas de temps, tendant vers o quand N tend vers l'infini et $\Gamma_{\delta_N}^{\mathrm{lin}}$ est l'ensemble des trajectoires affines par morceaux sur des intervalles de taille δ_N. L'énergie totalement discrète minimisée est:

$$J_{\delta_N,\varepsilon_N}(Q) := \int_\Gamma L(\gamma')dQ(\gamma) + \delta_N \sum_{i=1}^{M_N-1} F_{\varepsilon_N}\left(e_{i\delta_N}\#Q\right) + G(Q)$$

avec $M_N\delta_N = T$.

De la même façon que pour le problème discret en espace et continu en temps $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, nous montrons que les solutions du problème complètement discret $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$ convergent étroitement vers des solutions du problème continu $(\mathcal{M}_{\mu^0})$. Ce résultat requiert les mêmes hypothèses sur la décroissance de ε_N plus une condition similaire sur le pas de discrétisation en temps, δ_N.

Nous approchons les solutions du problème discrétisé $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$ par un algorithme de type quasi-Newton, limited-memory BFGS. Ce problème est de dimension finie en les positions des trajectoires à chaque pas de temps, et le gradient de l'expression minimisée peut se calculer explicitement pour certaines énergies potentielles G. Les valeurs de ε_N et δ_N appropriées ne sont malheureusement pas accessibles, à l'exception de certains cas très particuliers, mais les résultats numériques sont très satisfaisants même pour des valeurs arbitraires de ε_N, et les valeurs de T et δ_N peuvent être déterminées expérimentalement de manière à limiter .

Les fonctionnelles minimisées dans $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$ et $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$ ne sont pas convexe et l'algorithme L-BFGS risque d'aboutir à un point critique (une trajectoire) non minimisante. Dans l'espoir de bien choisir l'initialisation μ_N^0 et ε_N, afin d'éviter ces trajectoires dans l'optimisation, nous étudions le problème de l'approximation d'une mesure à densité par une mesure uniforme discrète:

$$\min \left\{ W_2^2(\rho,\mu) \;\middle|\; \mu = \frac{1}{N}\sum_{i=1}^N \delta_{y_i} \right\}. \tag{1}$$

Il s'agit évidemment d'un problème plus simple, cependant, il présente des difficultés similaires du fait de l'interraction entre la distance de Wasserstein et une dépendance en la position de masses de Dirac. Ce dernier problème, parfois appelé quantification optimale uniforme d'une mesure, a été étudié numériquement dans plusieurs publications (notablement [17], [69]) depuis sa première mention par Balzer et al. [7]. Le problème général de quantification optimale est plus ancien (nous référons le lecteur à [32] pour une présentation plus détaillée), et l'algorithme privilégié pour sa résolution est celui de Lloyd, consistant à partir d'un nuage de points et raffiner les positions dans cet échantillon en envoyant chaque point vers le barycentre de zones d'une partition de l'espace qui est ensuite recalculée pour le nouveau nuage de points. Un algorithme similaire est utilisé pour la quantification optimal, où les points du nuage sont projetés sur le barycentre de la cellule de Laguerre optimale (pour le transport optimal vers ρ) associée:

$$\begin{cases} Y^0 \in \Omega^N \\ Y^{k+1} = B(Y^k) \end{cases} \tag{2}$$

où $B(Y) = (b_1(Y), \dots, b_N(Y))$ est le nuage des barycentres des cellules de Laguerre optimales associées à Y:

$$b_i(Y) = N \int_{\mathrm{Lag}_i(Y,\Phi)} x \mathrm{d}\rho(x).$$

Les études numériques que nous mentionnons mettent en évidence deux faits amplement observés et commentés dans la littérature, à savoir que le nuage initial Y^0 doit être constitué de points bien espacés afin d'obtenir une bonne convergence d'une part, et qu'une fois cette selection faite, peu d'itérations de l'algorithme de Lloyd (voire même une seule) aboutissent à une mesure donnant une bonne erreur de quantification. Cependant, il semblerait qu'aucune étude théorique de ces affirmations n'ait été menée, et en particulier quantitative du taux de convergence (en le nombre de masses de Dirac, N) d'un minimiseur local de (1) vers ρ, ou plus généralement de la distance entre les points du nuage initial, afin d'éviter d'obtenir un point critique non-minimisant en sortie d'algorithme.

Nous fournissons dans le chapitre 3 des estimations explicites (et en particulier non-asymptotiques) pour la distance de Wasserstein entre la mesure uniforme obtenue après une étape de l'algorithme de Lloyd et la mesure à densité échantillonnée. En particulier, si les points de Y^0 sont espacés d'au moins $C.N^{-1/d}$ (en l'esprit, répartis sur un ensemble de dimension d), alors,

$$W_2^2\left(\rho, \frac{1}{N}\sum_{i=1}^{N} \delta_{b_i(Y)}\right) \leq K.N^{-\frac{1}{d}},$$

avec une constante explicite K ne dépendant que de d, C et Ω. Ces estimées impliquent en particulier l'observation empirique qu'en échantillonnant à partir de points suffisamment espacés, la mesure supportée sur les barycentres est proche de la mesure sous-jacente ρ au sens de W_2, lorsque le nombre de points tend vers l'infini.

Ce résultat peut être vu comme une variante de l'inégalité de Polyak-Łojasiewicz pour la minimisation de la fonction:

$$F_N : Y \in (R^d)^{\mathbb{N}} \mapsto W_2^2 \left(\rho, \frac{1}{N} \sum_{i=1}^{N} \delta_{y_i} \right),$$

avec cependant un terme divergeant lorsque le nuage de points n'est pas assez espacé:

$$F_N(Y) - C_{d,\Omega} \frac{1}{N} \left(\frac{1}{\varepsilon} \right)^{d-1} \leq N \left\| \nabla F_N(Y) \right\|^2$$

où ε est la distance minimale entre deux points du nuage Y. Similairement à une véritable inégalité de Polyak-Łojasiewicz, cette inégalité suggère que suivre le flot gradient de F_N nous approche d'un point critique, tant que l'on peut maintenir une distance suffisante entre les points obtenus à chaque itération. Nous garantissons cette distance minimale durant le mouvement par une inégalité de Gronwall (ou son analogue géométrique discrète en temps), ce qui nous permet d'étudier tout d'abord les mesures obtenues après quelques pas de descente de gradient pour F_N:

$$\begin{cases} Y^0 \in \Omega^N \\ Y^{k+1} = Y^k - N\tau_N \nabla F_N(Y^k) = Y^k + \tau_N (B(Y^k) - Y^k) \end{cases}$$

Pour un pas de temps $0 < \tau_N < 1$ et un nuage de points espacés de $C.N^{-1/d}$, nous déterminons k_N tel que après k_N pas de descente de gradient, le nuage obtenu est proche de ρ, au sens de Wasserstein:

$$W_2^2 \left(\rho, \frac{1}{N} \sum_{i=1}^{N} \delta_{y_i^{k_N}} \right) = O_{N \to \infty} \left(W_2^2 \left(\rho, \delta_{Y_N^0} \right)^{1 - \frac{1}{d}} . N^{\frac{-1}{d^2}} \right).$$

Le taux de convergence (en N) est bien pire dans ce cas qu'après une étape de l'algorithme de Lloyd, et ce résultat n'a pas d'intéret au-delà de présenter, dans un cas simple, cette non-dégénérescence de la structure du nuage de points, durant un tel mouvement de particules. En revanche, retournant au contexte des mouvements de foules non-congestionnés, il est possible d'évaluer les valeurs

des paramètres de discrétisation ε_N dans le cas du flot gradient étudié par Leclerc, Mérigot, Santambrogio et Stra. Cette étude est effectuée dans la Section 3.4 et nous permet de nous débarasser d'une hypothèse artificielle utilisée par les auteurs pour démontrer la convergence du flot discrétisé.

Cependant, les inégalité de type Gronwall, garantissant une distance minimale entre les individus (les masses de Dirac) durant le mouvement, utilisent beaucoup l'ordre 1 des équations différentielles vérifiées par les trajectoires, et s'adaptent mal à notre modèle (d'ordre 2 en temps) de jeu à champ moyen.

Le dernier chapitre de cette thèse se tourne vers les cas où les mesures de quantification discrètes ne convergent pas vers la densité échantillonnée ρ. Même dans ces cas très dégénérés, il est possible de mettre en évidence une grande régularité des mesures limites, héritée principalement de la symétrie des cellules de Laguerre (elle-même due aux conditions d'optimalités pour F_N). Nous démontrons tout d'abord que les limites de ces mesures discrètes sont solutions d'un problème de minimisation de distance de Wasserstein, sous certaines restrictions de répartition faisant intervenir le support de la mesure limite, μ_∞:

$$\min\{W_2^2(\rho,\mu) \mid \mu \in \mathcal{P}(\Omega),\ \mu(C) = \mu_\infty(C) \text{ pour toute composante connexe } C \text{ de } \mathrm{spt}(\mu_\infty)\}.$$

Nous nous intéressons ensuite aux mesures vérifiant une version continue de la condition de criticité "être le barycentre de sa cellule de Laguerre" (obtenue pour le support d'une mesure discrète optimale pour le problème de quantification uniforme). Cette nouvelle condition se définit en remplaçant les cellules (et la mesure échantillonnée sur celles-ci) par une désintégration de ρ par rapport au plan de transport entre μ et ρ: $(\rho_y)_{y\in\mathrm{spt}(\mu)}$. La condition de criticité devient alors:

$$\text{Pour } \mu\text{-presque tout } y \in \Omega,\ y = \int_\Omega x \mathrm{d}\rho_y(x)$$

et on dit que la mesure μ est critique Lagrangienne (pour $W_2^2(.,\mu)$) lorsque cette condition μ-presque partout est vérifiée.

Le résultat principal étudié dans ce chapitre est si, oui ou non, toute mesure critique Lagrangienne est supportée sur une union dénombrable de surfaces régulières de dimensions entières et est absolument continue sur chaque surface par rapport à la mesure de Hausdorff de même dimension. Une réponse positive est suggérée par l'expérience, ainsi que par un résultat de "boules extérieures" pour le support d'une telle mesure: En effet, pour μ-presque tout point $y \in \Omega$, et $x \in \mathrm{spt}(\rho_y)$, aucun point du support de μ ne peut être dans la

boule de diamètre $[x, y]$.

Cette propriété est à priori trop faible pour obtenir une régularité meilleure que la rectifiabilité pour le support de μ, et en particulier, elle ne donne pas d'informations sur le comportement de la mesure Lagrangienne critique sur le support. Un résultat pouvant donner de la régularité supplémentaire est un théorème prouvé par Alberti, donnant une structure C^2-rectifiable pour les parties du support sur lesquelles la dimension de $\mathrm{spt}(\rho_y)$ est constante. Ce résultat est valable à un ensemble de mesure de Hausdorff nulle près, et la conjecture se réduit donc uniquement à démontrer l'absolue continuité de μ par rapport à la mesure de Hausdorff de dimension adéquate sur chaque partie du support. Cette affirmation reste encore une conjectures pour les parties du support dont la dimension k serait comprise entre 2 et $d-1$ (une définition rigoureuse du sens de cette dimension est donnée dans ce chapitre). En revanche, nous donnons une démonstration complète de cette affirmation pour les parties de dimensions $k = 0, 1$ ou d et réduisons le résultat (pour toutes les dimensions possibles) à une propriété de continuité Lipschitz sur les variations de l'orientation du support de ρ_y, en fonction de y.

Chapter 1

Optimal transport and discretization of congestion penalizations

In this chapter, we briefly recall some well-known facts in the theory of optimal transport and, in particular, on the semi-discrete case. We then present some new results and proofs of properties for functions defined as Moreau envelope with respect to the 2-Wasserstein distance (this is Section 1.2 of this chapter).

1.1 Optimal transport and the semi-discrete setting

Optimal transport was introduced in 1781 by G. Monge in his *"Mémoire sur la théorie des déblais et des remblais"* [54]. It took the form of a very simple problem of moving a mass of sand from a location to another while minimizing the travelled distance.

With this formulation, each particle of sand located at x is transported to $T(x)$ with a cost proportional to $\|T(x) - (x)\|$ and the total cost to move a pile of sand S_0 to a location S_1 (a set of same volume) is proportional to

$$\int_{S_0} \|T(x) - x\| \, \mathrm{d}x$$

which Monge then minimizes among the transport maps $T : S_0 \to S_1$ under the condition that they are volume preserving.

Many issues, such as existence of such maps, of one which minimizes the total cost, lack of linearity in the solution T, ... make this problem ill-posed and Monge's solution is neither very satisfactory, nor general.

The problem was explored again by A.N. Tolstoï in *"Methods of finding the minimal kilometrage in cargo-transportation in space"* for the National Commissariat of Transportation of the Soviet Union in 1930 [66]. In this paper, the problem is considered in a discrete setting, as different cargos are transported between sources and destinations along the railway network of the Soviet Union. Although it seems doubtful that Tolstoï had a rigourous proof for it, his solution is actually an optimum and he did discover the now well-known heuristic of there being no negative cycles in the transport plan (see [56], proposition 3.4).

However, it is L. Kantorovich who is often considered to be the first mathematician to rigourously study the problem and make significant advances regarding its resolution in his paper *"On the translocation of masses"* for the Proceedings of the USSR Academy of science in 1942 [35]. His definition of the optimal transport cost is the one used today, defined as the minimal average cost of pairing between two general probability measures, in Definition 1. He also showed that, in some specific cases, the optimal transport cost can be computed from the now well-known *Kantorovich Potentials* (see Proposition 2).

Modern formulation of Optimal transport: Kantorovich's formulation can be developed in very general spaces and for very general costs. However, in this thesis we will often make assumptions to guarantee the existence of these Kantorovich Potentials (typically through compactness of the support of our measures), or to obtain optimal transport plans which are given by maps (by considering measures which admit densities with respect to Lebesgue measures and a cost given by the euclidean distance). We now recall a somewhat general setting in which transport problems can be studied and then develop the results for existence of potentials or transport maps which we shall use. All of the notions on Optimal transport as well as the corresponding proofs can be found in the books by C. Villani, [67] and [68], or in the one by F. Santambrogio [63]:

Definition 1. *Let X, Y be two Polish spaces and let $c : X \times Y \to \mathbb{R}$ be a lower semi-continuous function, bounded from below. Let $\rho \in \mathcal{P}(X)$, $\mu \in \mathcal{P}(Y)$,*

The optimal transport cost associated with c, between these measures is the value of the infimum:

$$I_c(\rho, \mu) = \inf_{\gamma \in \Pi(\rho,\mu)} \int_{X\times Y} c(x, y) \mathrm{d}\gamma(x, y) \tag{1.1}$$

and the set of transport plans between ρ and μ, $\Pi(\mu, \nu)$ is the set:

$$\Pi(\mu, \rho) = \{\gamma \in \mathcal{P}(X \times Y),\ \pi_X \# \gamma = \rho, \pi_Y \# \gamma = \mu\} \tag{1.2}$$

In the above Definition 1, lower semi-continuity of c is the property described in Definition 17, Appendix B and $\pi_X \# \gamma$ is the *push-forward* measure of γ along

the projection on the first coordinate (and idem for $\pi_Y \# \gamma$):

$$\pi_X : (x, y) \in X \times Y \to x \in X, \quad \pi_Y : (x, y) \in X \times Y \to y \in Y.$$

Push-forwards of measures through maps are defined in Appendix A, Definition 9.

It is important to note here that, under these assumptions on X and c, the infimum problem (1.1) admits solutions. This is shown using the direct method in Calculus of Variations, of taking a minimizing sequence, showing some compactness on it and concluding by lower semi-continuity of the minimized functional. We make ample use of this method whenever we need to show existence of solutions to such variationnal problems (see the proof of Proposition 5 for an example).

A major interest of Kantorovich's formulation is that Problem (1) is a convex problem which admits a dual formulation as a concave one, through the celebrated Fenchel-Rockafellar duality (see Appendix B Theorem 46, as well as [23]). Let us recall here (Appendix B.1) that the space $\mathcal{M}(X)$ of finite signed (Borel) measures can naturally be put in duality with the space of continuous bounded functions on X (although it is not the topological dual of $\mathcal{C}_b^0(X)$, except when X is compact). It is this duality which is exploited in Fenchel-Rockafellar theorem:

Theorem 1 (Kantorovich's duality)**.** *Assume that X and Y are compact spaces.*

$$I_c(\rho, \mu) = \sup_{\psi,\ \phi \in \mathcal{C}^c(X,Y)} \int_X \psi \mathrm{d}\rho + \int_Y \phi \mathrm{d}\mu \tag{1.3}$$

the supremum being taken over the space

$$\mathcal{C}^c(X, Y) = \{\psi,\ \phi \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y), \\ \forall (x, y) \in X \times Y\ \psi \oplus \phi(x, y) := \psi(x) + \phi(y) \leq c(x, y)\} \tag{1.4}$$

We recall now a proof for this duality result, as it can be obtained in a very similar fashion to that in which we obtain strong duality for our regularized congestion Problem (1.13):

Proof. We define the following two convex functions and one linear map on the product of these last vector spaces:

$$F : \psi, \phi \in \mathcal{C}^0(X) \times \mathcal{C}^0(Y) \mapsto \int_X \psi \mathrm{d}\rho + \int_Y \phi \mathrm{d}\mu$$

$$G : \phi \in \mathcal{C}^0(X \times Y) \mapsto \begin{cases} 0 & \text{if } \phi \leq c \\ +\infty & \text{otherwise.} \end{cases}$$

and

$$A : \psi, \phi \in \mathcal{C}^0(X) \times \mathcal{C}^0(Y) \mapsto \phi \oplus \psi \in \mathcal{C}^0(X \times Y)$$

Then, the dual formulation (1.3) can be simply written as:

$$-\inf \left\{ F(-(\psi, \phi)) + G(A(\psi, \phi)) \;\middle|\; \psi, \phi \in \mathcal{C}^0(X) \times \mathcal{C}^0(Y) \right\}$$

We are going to use Fenchel-Rockafellar theorem from Appendix B, Theorem 46, for the duality between $\mathcal{C}^0(X \times Y)$ and $\mathcal{M}(X \times Y)$. The primal problème (1.1) could be formulated as an inf-convolution:

$$I_c(\rho, \mu) = \inf_{\gamma \in \mathcal{M}(X \times Y)} \int_{X \times Y} c(x, y) \mathrm{d}\gamma(x, y) + \chi_{\Pi(\mu, \rho)}(\gamma)$$

with the definition of the convex indicator function of transport plans from ρ to μ, $\chi_{\Pi(\mu,\rho)}$, from Appendix B. However, both functions in the sum enjoy too few continuity in order for this to be done easily, and instead, we apply the theorem to the dual expression (or rather, its opposite in order to start from an infimum). F is linear continuous on $\mathcal{C}^0(X) \times \mathcal{C}^0(Y)$ and G is lower semi-continuous on $\mathcal{C}^0(X \times Y)$ and $\mathrm{dom}(G) \neq \emptyset$ therefore, the hypotheses of the theorem are trivially verified. We may now write:

$$\begin{aligned} &\inf \left\{ F(-(\psi, \phi)) + G(A(\psi, \phi)) \;\middle|\; \psi, \phi \in \mathcal{C}^0(X) \times \mathcal{C}^0(Y) \right\} \\ &\qquad = \sup \left\{ -F^*(A^*(\gamma)) - G^*(\gamma) \mid \gamma \in \mathcal{M}(X) \times \mathcal{M}(Y) \right\} \end{aligned}$$

(we recall that A^* acts on the dual of $\mathcal{C}^0(X \times Y)$ which is $\mathcal{M}(X \times Y)$). Now, for $\alpha, \beta \in \mathcal{M}(X) \times \mathcal{M}(Y)$ and $\gamma \in \mathcal{M}(X \times Y)$:

$$\begin{aligned} F^*(\alpha, \beta) =& \sup_{\phi, \psi} \int_X \psi \mathrm{d}(\alpha - \rho) + \int_X \phi \mathrm{d}(\beta - \mu) \\ =& \chi_{\{\rho\}}(\alpha) + \chi_{\{\mu\}}(\beta), \end{aligned}$$

$$A^*(\gamma) = (\pi_X \# \gamma, \pi_Y \# \gamma)$$

and

$$\begin{aligned} G^*(\gamma) =& \sup_{\phi \leq c} \int_{X \times Y} \phi \mathrm{d}\gamma \\ =& \begin{cases} \int_{X \times Y} c \mathrm{d}\gamma & \text{if } \gamma \in \mathcal{M}_+(X \times Y) \\ +\infty & \text{otherwise} \end{cases} \end{aligned} \tag{1.5}$$

since c was assumed lower semi-continuous.

Plugging these expressions in Fenchel-Rockafellar's theorem and multiplying each side by -1, we obtain:

$$\sup\left\{\int_X \psi \mathrm{d}\rho + \int_Y \phi \mathrm{d}\mu, \;\middle|\; \psi, \phi \in \mathcal{C}^0(X) \times \mathcal{C}^0(Y),\ \psi \oplus \phi \leq c\right\}$$
$$= \inf\left\{\int_{X\times Y} c\mathrm{d}\gamma \mid \gamma \in \mathcal{M}_+(X \times Y),\ \pi_X \# \gamma = \rho,\ \pi_Y \# \gamma = \mu\right\} \quad (1.6)$$

which states strong duality for Problem $I_c(\rho, \mu)$. Furthermore, this is a case where Fenchel-Rockafellar's theorem guarantees that the infimum (which is really a supremum in the theorem) is attained, which is another way of obtaining existence of an optimal transport plan in this case. □

We restrain ourselves to the case where X and Y are both compact sets, both for simplicity and, again to keep close to the duality result that we show later on. To obtain the same result for general Polish spaces X and Y, one simply restrains the problem to larger and larger compact sets inside X and Y and use stability of the transport plans and the fact that ρ and μ are Radon measures. This does require however that the cost be lower bounded by a separable expression (see [68], Theorem 5.10).

c-concave functions: In this paragraph, we assume, again, that X and Y are compact spaces and that c is continuous on $X \times Y$ (which is the case in most of our problems) . A useful trick, which we adapt later on, comes from noticing that any of the optimization variables, ψ for instance, can (and should) be taken as large as possible, since both ρ and μ are non-negative. This variable can then be eliminated from problem (1.3), using the *c-transform* of the other variable:

Definition 2. *Let $\phi \in \mathcal{C}^0(Y)$, we call* c-transform of ϕ *the function, in $\mathcal{C}^0(X)$:*

$$\phi^c := \inf_{y \in Y} c(., y) - \phi(y).$$

Similarly, and since no confusion should arise from this, for $\psi \in \mathcal{C}^0(X)$, we call c-transform of ψ the function in $\mathcal{C}^0(Y)$:

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

A function (in $\mathcal{C}^0(X)$ or $\mathcal{C}^0(Y)$) will be called c-concave when it is the c-transform of a (any) function.

Looking back at the definition of $\mathcal{C}^c$ and the fact that ρ and μ are positive, one can take the variable ψ as large as the c-transform of ϕ and replace the variable ϕ by $(\phi^c)^c \geq \phi$ as it yields a better competitor. Finally since $((\phi^c)^c)^c \equiv \phi^c$, we are left with the dual problem

$$\begin{aligned} I_c(\rho,\mu) &= \sup_{\phi \in \mathcal{C}_b^0(Y)} \int_X \phi^c \mathrm{d}\rho + \int_Y \phi^{cc} \mathrm{d}\mu \\ &= \sup_{\phi \text{ c-concave}} \int_Y \phi \mathrm{d}\mu + \int_X \phi^c \mathrm{d}\rho \end{aligned} \tag{1.7}$$

Before continuing any further, we wish to make a remark here:

The reason (at least one) for considering μ and ρ as measures on different sets is of course that the functions in the dual problem can then be taken in the continuous and bounded functions on the support of the corresponding measure, for instance for ϕ above, the support of μ. However, one has to be conscious of the fact that the notion of c-concave function depends on the sets X and Y considered. Therefore, although changing X to the support of ρ and Y to that of μ in either formulations of $I_c(\rho,\mu)$ does not change its value, it does change which Kantorovich potentials are c-concave and therefore the actual solutions of (1.7). This is a point which will have a small importance in particular in semi-discrete optimal transport, see below, and in the dual formulation of the regularized congestion F_ε at the end of this chapter.

More importantly, a c-concave function has the same continuity modulus as c, with respect to the corresponding variable. Assuming that X and Y are compact and c continuous, one can show, using the same direct method of calculus of variation and Arzelà-Arscoli theorem that there exists solutions for the dual Problem:

Proposition 2. *The dual formulation* (1.7) *admits solutions called* Kantorovich potentials *for the optimal transport. Furthermore, there always exists pairs of c-concave Kantorovich potentials (ϕ, ϕ^c) for the optimal transport.*

Wasserstein distances: A key feature of Optimal transport is the fact that it induces, in some cases, a distance metrizing the *narrow convergence* on the space of probability measures $\mathcal{P}(X)$ (see Definition 11, Appendix A):

Indeed, consider the optimal transport problem (1.1) with $X = Y$, a cost $c_p(x,y) = d_X(x,y)^p$ being a power of the distance on X and $p \geq 1$. Then $I_{c_p}(\rho,\mu)$ is finite if ρ and μ have finite $p-th$ order moment:

Definition 3. *A measure $\mu \in \mathcal{P}(X)$ has finite p-th order moment when there exists $x_0 \in X$, $\int_X d_X^p(x, x_0)\mathrm{d}\mu(x) < +\infty$. Any $x_0 \in X$ can be chosen to test this property.*

We denote $\mathcal{P}_p(X)$ the set of probability measures with finite p-th order moment on X.

Taking the appropriate p-root, one can define a distance on the set $\mathcal{P}_p(X)$ called the p-*Wasserstein distance*. Here we note that we will use later on the notation $\mathcal{P}_N(X)$, for the set of uniform discrete probability measures over N points. Since such uniform discrete measures instantly enjoys finite p-th order moments for any $p \geq 1$ no confusion should arise from these slightly awkward notations. Note also that we will be working in compact spaces in most cases, in which case probability measures (discrete or not) also enjoy finite moments of any order.

Definition 4 (Wasserstein/Monge-Kantorovich distance)**.** *Let $\mu,\ \rho \in \mathcal{P}(X)$ with finite p-th order moment. The p-*Wasserstein distance *between μ and ρ is defined as:*

$$W_p(\mu, \rho) = \sqrt[p]{I_{c_p}(\mu, \rho)}$$

with I_{c_p} defined by Definition 1.

As its name suggests, W_p is a distance on $\mathcal{P}_p(X)$. It generally defines a stronger notion of convergence than the narrow convergence from Definition 11. However, when X is compact, both topologies coincide:

Proposition 3. *Let $(\mu_n)_{n\in\mathbb{N}} \in (\mathcal{P}_p(X))^{\mathbb{N}}$ and $\mu \in \mathcal{P}_p(X)$. Let $x_0 \in X$.*

The two following points are equivalent:

(1) $\lim_{n\to\infty} W_p^p(\mu_n, \mu) = 0$.

(2) $\mu_n \xrightarrow[n\to\infty]{} \mu$ *and* $\lim_{n\to\infty} \int_X d^p(x_0, x)d\mu_n = \int_X d^p(x_0, x)d\mu$

In particular, if X is compact, $(1) \Leftrightarrow (2) \Leftrightarrow \left(\mu_n \xrightarrow[n\to\infty]{} \mu\right)$

"Brenier" transport: Let us finish this section by identifying specific cases in which the optimal transport plan is actually induced by a map $T : X \to Y$, which was the model Monge assumed in his "traité". What this means in terms of transport plan is the existence of a measurable map $T : X \to Y$ such that the optimal $\gamma \in \mathcal{P}(X \times Y)$ (for Problem (1.1)) is induced by T in the sense:

$$\gamma = (Id, T)\#\rho.$$

As we shall not use these results to the full extent of their generality, we assume for the remainder of this section that X and Y are compact subsets of $\mathbb{R}^d$. Generalizations to optimal transport on Riemannian manifolds can be found in McCann [49], and Gigli [30] characterized the class of measures for which such a transport map exists.

A very common assumption to obtain such transport is the following twist condition on the cost c, allowing for almost all $x \in X$ to find a unique target $y \in Y$ to be transported to:

Definition 5. *We will say that the cost c satisfies the* twist condition *when there exists open sets $\Omega_Y \supset Y$ and $\Omega_X \supset X$ such that:*

- *c is $\mathcal{C}^1(\Omega_X \times \Omega_Y)$.*
- *For any $x \in \Omega_X$, the map*

$$Y \in \Omega_Y \mapsto D_x c(x, y)$$

is one to one.

A first, very visual, example of such a transport "à la Monge" can be observed in what is usually referred to as semi-discrete optimal transport, and it causes a total collapse of the complexity of the optimization problem, into a finite dimensional one.

In this setting, one of the measures, ρ for instance, is absolutely continuous with respect to the Lebesgue measure, while the other, μ is discrete:

$$\rho \ll \mathrm{d}x, \qquad \mu = \sum_{i=1}^{N} \alpha_i \delta_{y_i}$$

for some point cloud $Y = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N$ and masses $\alpha_1, \dots, \alpha_N$ in $(\mathbb{R}^{*+})^N$, such that $\sum_{i=1}^N \alpha_i = 1$.

When in this semi-discrete context, we will make the abuse of notations of denoting the support of μ using a N-tuples instead of a set, since it is convenient to be able to quickly assign the masses α_i (and later on weights ϕ_i) to the corresponding points in $\mathrm{spt}(\mu)$.

The dual formulation (1.7) rewrites into a simpler finite dimensional concave problem involving a specific decomposition of the domain into Laguerre cells:

Definition 6 (Laguerre cells)**.** *Consider a point cloud $Y = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N$ and a set of reals $\Phi = (\phi_1, \dots, \phi_N) \in \mathbb{R}^N$. We call* Laguerre cells *associated to Y with weights Φ the covering of $\mathbb{R}^d$: For $i = 1 \dots N$,*

$$\mathrm{Lag}_i(Y, \Phi) = \{x \in \mathbb{R}^d,\ c(x, y_i) - \phi_i \leq c(x, y_j) - \phi_j \text{ for } j = 1, \dots, N\}$$

For a function $\phi : \mathrm{spt}(\mu) \to \mathbb{R}$, and weights $\Phi = (\phi(y_1), \dots, \phi(y_N))$, $\mathrm{Lag}_i(Y, \Phi)$ is the subset of X upon which the minimum defining ϕ^c is attained for $j = i$:

$$\phi^c := \min_j c(., y_j) - \phi(y_j) = c(., y_i) - \phi(y_i)$$

Notice also that the twist condition implies that the boundary of these Laguerre cells are Lebesgue-negligible, allowing us to rewrite Problem (1.7) as:

$$I_c(\rho, \mu) = \sup_{\Phi \in \mathbb{R}^N} \sum_{i=1}^{N} \alpha_i \phi_i + \int_{\mathrm{Lag}_i(Y,\Phi)} (c(x, y_i) - \phi_i) \mathrm{d}\rho(x) \tag{1.8}$$

This is now a finite dimensional concave problem, and Φ is an optimal bunch of weights iff $\rho(\mathrm{Lag}_i(Y, \Phi)) = \alpha_i$ for every i. Furthermore, it is known since Aurenhammer et al. [6] that an optimal tesselation always exists and is, in some sense, unique. These optimality conditions are very similar to those we will get in proposition 9 below in the setting of discretized congestion penalty. Provided Φ verifies these conditions, one can then write $I_c(\rho, \mu)$ as the separated sum:

$$I_c(\mu, \rho) = \sum_{i=1}^{N} \int_{\mathrm{Lag}_i(Y,\Phi)} c(x, y_i) d\rho(x)$$

We mention that problem (1.8) is the maximisation of a smooth (up to $\mathcal{C}^{2,\alpha}$ under some conditions on the domain and cost), concave function of the weights Φ. We refer the reader to the article by Kitagawa, Mérigot and Thibert [36] or the recent survey by Mérigot and Thibert [55] for a study of this problem and in particular, an overview of the methods for its numerical solution.

Remark 1. In Chapter 4, we deal with point clouds that can vary (and in particular grow larger) inside a compact domain Ω which also contains the support of ρ. Let us recall briefly how to obtain Kantorovich potentials (φ, φ^c) defined on the whole domain Ω, from a set of optimal weights Φ for Problem (1.8), such that for any i, $\varphi(y_i) = \phi_i$:

For such a set of weights, define the function $\psi := \inf_i c(., y_i) - \phi_i$ and $\varphi := \inf_{x \in \Omega} c(x, .) - \psi(x)$. For any $i = 1 \dots N$,

$$\varphi(y_i) = \inf_{x \in \Omega} \sup_j c(x, y_i) - c(x, y_j) + \phi_j \geq \phi_i$$

(taking $j = i$) and for $x \in \Omega$, if $x \in \mathrm{Lag}_i(Y, \Phi)$,

$$c(x, y_i) - c(x, y_j) + \phi_j \leq \phi_i$$

by definition. Therefore, since $\mathrm{Lag}_i(Y,\Phi)$ is non-empty for every i by optimality of Φ, $\varphi(y_i)=\phi_i$. On the other hand, for any $x\in\Omega$,

$$\varphi^c(x)=(\psi^c)^c(x)\geq\psi(x)$$

and,

$$I_c(\mu,\rho)=\sum_{i=1}^{N}\alpha_i\phi_i+\int_\Omega(\inf_i c(x,y_i)-\phi_i)\mathrm{d}\rho(x)\leq\int_\Omega\varphi\mathrm{d}\mu+\int_\Omega\varphi^c\mathrm{d}\rho\leq I_c(\mu,\rho)$$

and φ is a Kantorovich potential for the transport from μ to ρ.

Let us also note here that if the cost c is given by the squared euclidean norm: $c(x,y)=C\left\|x-y\right\|^2$, the Laguerre cells are convex polyhedra in $\mathbb{R}^d$ as they are the defined by the linear inequalities:

$$\mathrm{Lag}_i(Y,\Phi)=\left\{x\in\mathbb{R}^d,\ (2x-y_i-y_j)\cdot(y_j-y_i)\leq\frac{\phi_i-\phi_j}{C}\right\}$$

This is a very nice feature of this euclidean setting, and one of the reasons why we will use quadratic Wasserstein distances through most of this thesis. On the other hand, for the exponent $p=1$, $c(x,y)=\|x-y\|$, the Laguerre cells are not even always convex and their boundaries are made of arcs of hyperbolae (as they are defined by constant differences of distances to two points)!

In this semi-discrete setting, it is obvious that the optimal transport plan will be induced by the map: $T:x\in\mathrm{Lag}_i(Y,\Phi)\mapsto y_i$. However, the existence of this transport map T is much more a consequence of the twist condition on c and the absolute continuity of ρ than the discrete nature of $\mathrm{spt}(\mu)$. In fact, the following celebrated theorem, gives the same structure to the optimal transport plan, for more general measures.

Theorem 4 (Brenier [12], Gangbo-McCann [49])**.** *Let $\mu,\ \rho\in\mathcal{P}(\mathbb{R}^d)$. We assume that $\rho\ll\mathrm{d}x$ and that the cost c verifies the twist condition.*

Then there is a unique optimal transport plan for $I_c(\rho,\mu)$ and this transport plan is induced by a map,

$$\gamma=(Id,T)\#\rho.$$

This result with a cost given by the squared norm, $c(x,y)=C\left\|x-y\right\|^2$ was the original theorem, proven by Brenier, [12] and in this case, the optimal map T has an explicit form involving the gradient of a (any) c-concave Kantorovich potential, ϕ. Indeed, in this case, a function being c-concave is equivalent to it being

1-semiconcave (in the sense of Appendix B, Definition 18) and a c-concave Kantorovich potential is differentiable almost everywhere (for the Lebesgue measure and therefore for ρ). In that case, ϕ^c (with the notations of (1.7)) verifies ρ-almost everywhere $\nabla\phi^c(x) = \nabla_x c(x,y) = 2C(x-y)$ (in particular, Kantorovich potentials are uniquely defined, up to a constant, on the support of ρ). The transport map T of Brenier's theorem is simply given by

$$T(x) = x - \frac{\nabla\phi^c(x)}{2C}$$

which is the gradient of the (convex) function $\frac{\|.\|^2}{2} - \frac{\phi^c}{2C}$(see also [63], section 1.3.1 for a proof and more general study).

1.2 Moreau envelope in the Wasserstein space

In the following chapter, we study variational problems for which the minimized energy features a congestion penalization term, in the form of a convex, lower semi-continuous function F on $\mathcal{M}(\Omega)$, where Ω is a compact smooth domain of $\mathbb{R}^d$ in which the solutions of the problem evolve.

To define the discretization used in Chapter 2, we need to be able to evaluate F at discrete measures on Ω, which typically give us the value $+\infty$ and therefore little information on how congested these measures are. This is the case for instance when F has the following integral form, very commonly featured in these kinds of variational problems:

$$F(\rho) = \begin{cases} \int_\Omega f(\rho(x))dx & \text{if } \rho \ll dx \\ +\infty & \text{otherwise.} \end{cases} \tag{1.9}$$

Our solution is to replace this functional by a regularized version defined as a Moreau envelope (or Moreau-Yosida regularization) for the 2-Wasserstein distance (see Definition 4):

$$\text{For } \mu \in \mathcal{P}(\mathbb{R}^d),\ F_\varepsilon(\mu) := \inf_{\rho\in\mathcal{M}(\Omega)} \frac{W_2^2(\rho,\mu)}{2\varepsilon} + F(\rho) \tag{1.10}$$

To make expressions more concise, we will use from time to time the optimal transport cost associated with the cost $c_\varepsilon(x,y) = \frac{||x-y||^2}{2\varepsilon}$ (and more importantly the associated Kantorovich potentials $(\phi, \phi^{c_\varepsilon})$), instead of the standard squared norm defining W_2. Let us quickly notice the more concise expression for F_ε:

$$F_\varepsilon(\mu) = \inf_{\rho\in\mathcal{M}(\Omega)} I_{c_\varepsilon}(\rho,\mu) + F(\rho). \tag{1.11}$$

The transport cost $I_{c_\varepsilon}(.,\rho)$ (we recall Definition 1) is $+\infty$ outside of $\mathcal{P}(\mathbb{R}^d)$ while F is $+\infty$ outside of $\mathrm{dom}(F)$ and, in the end, the infimum in (1.10) is only taken on the intersection of those two sets. For this reason, we make the additional assumption that $\mathcal{P}(\Omega) \cap \mathrm{dom}(F) \neq \emptyset$ otherwise there is nothing to study.

This regularization is inspired by a similar one used to discretize the incompressibility constraint in the numerical approximation of solutions to Euler's equations by Q. Mérigot and J-M. Mirebeau, [50]. In this paper, the discrete measure μ is projected on the set of measures verifying the constraints, which corresponds to computing F_ε when F is a convex indicator function. Formulation (1.10) is more general in the sense that F can be softer and, for example, such expressions were also used to regularize the penalization $F(\rho) = \rho\log(\rho)$ by Leclerc, Mérigot, Santambrogio and Stra [40], in the context of Wasserstein gradient flows (we come back to these kinds of crowd motions at the end of Chapter 3). The name Moreau envelope comes directly from the same notion in a Hilbert space H. There, the Moreau envelope of a convex function g, with parameter ε, is given by a similar inf-convolution

$$g_\varepsilon(x) = \inf_{y\in H} \frac{\|x-y\|_H^2}{2\varepsilon} + g(y).$$

It has the advantages of being finite, and even differentiable, for any $x \in H$, upon some mild assumptions on g. Notice also that, assuming lower semi-continuity of g, $g_\varepsilon(x)$ has limit $g(x)$ as ε goes to 0, whereas the limit is $\inf g$ as ε goes to $+\infty$.

For the remainder of this section F will be a convex lower bounded function, l.s.c. for the narrow convergence on $\mathcal{M}(\mathbb{R}^d)$, with $\mathrm{dom}(F) \subset \mathcal{M}_+(\Omega)$. These are the same hypotheses as those made in the next chapter for the congestion penalty of variational mean field game problems. In particular, we do not assume that F has the integral form (1.14) for now:

Proposition 5. *For every $\varepsilon > 0$, the infimum defining F_ε is attained and F_ε is l.s.c on $\mathcal{P}(\mathbb{R}^d)$, for the narrow topology.*

Furthermore, $\lim_{\varepsilon\to 0} F_\varepsilon(\mu) = F(\mu)$ whereas, assuming μ has finite second order moment, we also have $\lim_{\varepsilon\to+\infty} F_\varepsilon(\mu) = \inf_{\rho\in\mathcal{P}(\Omega)} F(\rho)$.

We will sometimes call an optimal ρ in Problem (1.10) a *Moreau-Yosida projection of μ*.

Proof. To show existence of minimizers for Problem (1.10) let us consider a minimizing sequence $(\rho_n)_{n\in\mathbb{N}}$ for this problem, i.e. a sequence in $\mathcal{M}(\Omega)$ such that

$$\lim_{n\to+\infty} \frac{W_2^2(\rho_n,\mu)}{2\varepsilon} + F(\rho_n) = F_\varepsilon(\mu).$$

As we mentioned, $\rho_n \in \mathcal{P}(\Omega)$ for every n and Ω is compact, therefore, the sequence is tight. Using Prokhorov (or even, in this compact case, Banach-Alaoglu) theorem, one can extract a subsequence from $(\rho_n)_{n\in\mathbb{N}}$ which narrowly converges towards a measure $\rho \in \mathcal{M}(\Omega)$. Now, both F (by hypothesis) and $W_2^2(.,\mu)$ are lower semi-continuous functions for the narrow topology and therefore:

$$\frac{W_2^2(\rho,\mu)}{2\varepsilon} + F(\rho) \leq \liminf_{n\to+\infty} \frac{W_2^2(\rho_n,\mu)}{2\varepsilon} + F(\rho_n) = F_\varepsilon(\mu)$$

and ρ is a minimizer for (1.10).

Furthermore, going back to the definition of lower semi-continuity, let us take a sequence μ_n narrowly converging to μ_∞ in $\mathcal{P}(\Omega)$, and for every n, a measure $\rho_n \in \mathrm{dom}(F) \cap \mathcal{P}(\Omega)$ optimal for the problem defining $F_\varepsilon(\mu_n)$. We may assume that $F_\varepsilon(\mu_n)$ has a finite limit as $n \to \infty$. Using Prokhorov theorem again, we can extract a subsequence from $(\rho_n)_n$, narrowly converging towards a $\rho_\infty \in \mathrm{dom}(F)$. We extract the corresponding subsequence from μ_n and rename these new sequences, ρ_n and μ_n. Then,

$$\begin{aligned} F_\varepsilon(\mu_\infty) &\leq \frac{W_2^2(\rho_\infty,\mu_\infty)}{2\varepsilon} + F(\rho_\infty) \\ &\leq \liminf_n \frac{W_2^2(\rho_n,\mu_n)}{2\varepsilon} + \liminf_n F(\rho_n) \\ &\leq \liminf_n F_\varepsilon(\mu_n) = l \end{aligned}$$

the first inequality being by definition of $F_\varepsilon(\mu_\infty)$ and the second one, by lower semi-continuity for the narrow topology of the two terms. This means, by definition that F_ε is l.s.c. since l was the inf limit of the initial sequence $(\mu_n)_{n\in\mathbb{N}}$.

Finally, to obtain the limits, let us immediately observe that for $\varepsilon > \varepsilon'$,

$$\inf_{\rho\in\mathcal{M}(\Omega)} F(\rho) \leq F_\varepsilon(\mu) \leq F_{\varepsilon'}(\mu) \leq F(\mu)$$

As $\varepsilon \to 0$, we may assume that $(F_\varepsilon(\mu))_\varepsilon$ remains upper bounded, otherwise, its limit is $+\infty \leq F(\mu)$ and there is nothing to prove. Consider for any ε, ρ_ε a minimizer for Eq. (1.10).

$$\lim_{\varepsilon\to 0} W_2^2(\rho_\varepsilon,\mu) = \lim_{\varepsilon\to 0} 2\varepsilon(F_\varepsilon(\mu) - F(\rho_\varepsilon)) = 0$$

and using the lower semi-continuity of F,

$$F(\mu) \leq \liminf_{\varepsilon\to 0} \frac{W_2^2(\rho_\varepsilon,\mu)}{2\varepsilon} + F(\rho_\varepsilon) = F_\varepsilon(\mu)$$

and we have our limit.

As $\varepsilon \to +\infty$ on the other hand, let us take $\rho_\varepsilon \in \mathcal{P}(\Omega)$ such that

$$\lim_{\varepsilon\to+\infty} F(\rho_\varepsilon) = \inf_{\mathcal{P}(\Omega)} F$$

(in fact, this infimum is even attained but we do not need this).

$$F_\varepsilon(\mu) \leq \frac{W_2^2(\rho_\varepsilon,\mu)}{2\varepsilon} + F(\rho_\varepsilon)$$

and since any probability measure on Ω is at bounded distance from μ as soon as μ has finite second order moment,

$$\lim_{\varepsilon\to+\infty} F_\varepsilon(\mu) \leq \lim_{\varepsilon\to+\infty} \frac{W_2^2(\rho_\varepsilon,\mu)}{2\varepsilon} + \lim_{\varepsilon\to+\infty} F(\rho_\varepsilon) = \inf_{\mathcal{P}(\Omega)} F$$

□

Before diving into the computation of a dual formulation for F_ε, let us mention that the Moreau-Yosida projection is unique and behaves in a continuous way either at absolutely continuous measures or when F enjoys a little more convexity.

Proposition 6. *If either (1) $\mu \ll \mathrm{d}x$, (2) F is strictly convex or (3) F is convex along generalized geodesics (see Definition 20), then the optimal ρ in* (1.10) *is unique.*

Moreover, writing ρ^μ the Moreau-Yosida projection of a measure μ (unique in either of these three cases),

(1) $\mu \mapsto \rho^\mu$ is continuous (for the narrow topology) on $\mathbb{L}^1(\Omega) \cap \mathcal{P}(\Omega)$.

(2) If F is strictly convex, $\mu \mapsto \rho^\mu$ is continuous on the whole of $\mathcal{P}(\Omega)$.

(3) If F is convex along generalized geodesics and $\mathrm{dom}(F) \subset \mathbb{L}^1(\Omega)$, then $\mu \mapsto \rho^\mu$ is $\frac{1}{2}$-Hölder on any Wasserstein neighbourhood of a sub-level of F.

More precisely, for any L and $C > 0$, there exists $C' > 0$ such that if μ, ν are probability measures with $W_2(\mu, \{F \leq L\}) \leq C$ and also $W_2(\nu, \{F \leq L\}) \leq C$, then

$$W_2^2(\rho^\mu, \rho^\nu) \leq W_2^2(\mu,\nu) + C'.W_2(\mu,\nu)$$

Proof. As for uniqueness in (1) and (2), in both these cases the minimized functional $\frac{W_2^2(.,\mu)}{2\varepsilon} + F$ is strictly convex, respectively (1) on $\mathbb{L}^1(\Omega) \cap \mathcal{P}(\Omega)$ because so is $W_2^2(.,\mu)$ and (2) everywhere when F is strictly convex. Therefore in these cases, the optimal ρ is unique.

In the case (3) of F convex along generalized geodesics, uniqueness comes naturally from the proof of Hölder regularity and we prove both claims at once!

This proof is a straightforward adaptation of [61], p 50. Take μ, ν as above and ρ^μ, ρ^ν, minimizers for (1.10). Consider the following generalized geodesics $(\rho_t^\mu)_{t\in[0;1]}$, $(\rho_t^\nu)_{t\in[0;1]}$ such that for any $\phi \in \mathcal{C}^0(\Omega)$:

$$\int_\Omega \phi(x) d\rho_\mu^t(x) = \int_{\Omega^3} \phi((1-t)x_1 + tx_3) d\gamma_\mu^\mu(x_2, x_1) d\gamma_\mu^\nu(x_2, x_3)$$

$$\int_\Omega \phi(x) d\rho_\nu^t(x) = \int_{\Omega^3} \phi(tx_1 + (1-t)x_3) d\gamma_\nu^\mu(x_2, x_1) d\gamma_\nu^\nu(x_2, x_3)$$

and for $m, n \in \{\mu, \nu\}$, γ_n^m is the optimal transport plan from n to ρ^m. One of these "geodesics", ρ_t^μ for instance represents a path, from ρ^μ to ρ^ν in this case, followed from the point of view of one of the initial measures, μ in this instance.

As F is convex along generalized geodesics and $\frac{1}{2}W_2^2$ is 1-convex along generalized geodesics, one can write

$$\begin{aligned}\frac{W_2^2(\mu,\rho^\mu)}{2\varepsilon} + F(\rho^\mu) \leq& \frac{W_2^2(\mu,\rho_\mu^t)}{2\varepsilon} + F(\rho_\mu^t)\\ \leq& (1-t)\frac{W_2^2(\mu,\rho^\mu)}{2\varepsilon} + t\frac{W_2^2(\mu,\rho^\nu)}{2\varepsilon} - \frac{t(1-t)}{2\varepsilon}W_2^2(\rho^\mu,\rho^\nu)\\ &+ (1-t)F(\rho^\mu) + tF(\rho^\nu)\end{aligned}$$

and

$$\begin{aligned}\frac{W_2^2(\nu,\rho^\nu)}{2\varepsilon} + F(\rho^\nu) \leq& \frac{W_2^2(\nu,\rho_\nu^t)}{2\varepsilon} + F(\rho_\nu^t)\\ \leq& t\frac{W_2^2(\nu,\rho^\mu)}{2\varepsilon} + (1-t)\frac{W_2^2(\nu,\rho^\nu)}{2\varepsilon} - \frac{t(1-t)}{2\varepsilon}W_2^2(\rho^\mu,\rho^\nu)\\ &+ tF(\rho^\mu) + (1-t)F(\rho^\nu)\end{aligned}$$

and summing and rearranging the terms:

$$\begin{aligned}\frac{t(1-t)}{\varepsilon}W_2^2(\rho^\mu,\rho^\nu) \leq t\Bigg(&\frac{W_2^2(\mu,\rho^\nu)}{2\varepsilon} + \frac{W_2^2(\nu,\rho^\mu)}{2\varepsilon}\\ &- \frac{W_2^2(\mu,\rho^\mu)}{2\varepsilon} - \frac{W_2^2(\nu,\rho^\nu)}{2\varepsilon}\Bigg)\end{aligned}$$

so that, dividing by t and letting $t \to 0$

$$2W_2^2(\rho^\mu,\rho^\nu) \leq W_2^2(\mu,\rho^\nu) + W_2^2(\nu,\rho^\mu) - W_2^2(\mu,\rho^\mu) - W_2^2(\nu,\rho^\nu) \qquad (1.12)$$

The end of the proof comes from the triangular inequality:

$$W_2^2(\mu, \rho^\nu) \leq W_2^2(\mu, \nu) + W_2^2(\nu, \rho^\nu) + 2W_2(\mu, \nu)W_2(\nu, \rho^\nu)$$

and similarly, switching the roles of μ and ν.

By assumption,

$$W_2^2(\nu, \rho^\nu) \leq W_2^2(\nu, \{F \leq L\}) + 2\varepsilon(L - \inf F) =: K$$

and, again, the same is true replacing ν with μ. Injecting these four inequalities in (1.12) (and dividing by 2),

$$W_2^2(\rho^\mu, \rho^\nu) \leq W_2^2(\mu, \nu) + 2\sqrt{K}.W_2(\mu, \nu)$$

Now, taking $\mu = \nu$, we immediately get that $\rho^\mu = \rho^\nu$ and Problem (1.10) has a unique solution. For general μ, ν this simply states the Hölder continuity of the claim. □

Regarding the last point of the previous proposition, we emphasize the fact that our generalized geodesics are meant for the 2-Wasserstein distance. For other powers, the p-Wasserstein distance fails to verify the strong convexity along these generalized geodesics.

Being defined as the minimum of a convex function (both $W_2^2(., \mu)$ and F are convex on $\mathcal{M}(\Omega)$), F_ε can be rewritten as the supremum of a concave dual problem using once again Fenchel-Rockafellar duality. We make here two hypotheses on our congestion penalizing function F to ensure that this dual problem has solutions. Let us point out the fact that, same as in the case of optimal transport, Fenchel-Rockafellar theorem, as stated in the appendix, cannot be used on the primal formulation of the problem, as both terms are only l.s.c. with respect to the narrow convergence. We therefore start from the dual formulation (1.13) and work our way back to the primal one:

Proposition 7. *For any $\mu \in \mathcal{P}(\mathbb{R}^d)$,*

$$F_\varepsilon(\mu) = \sup_{\varphi \ c_\varepsilon - concave} \int_{\mathbb{R}^d} \varphi d\mu - F^*(-\varphi^{c_\varepsilon}). \tag{1.13}$$

This supremum is attained at φ if and only if for any ρ optimal for the primal problem, ie $\rho \in \mathrm{argmin} W_2^2(., \mu) + F$,

- *$(\varphi, \varphi^{c_\varepsilon})$ is a pair of c_ε-concave Kantorovich potentials for the optimal transport from μ to ρ.*

- $(-\varphi^{c_\varepsilon}) \in \partial F(\rho)$ *(or, equivalently,* $\rho \in \partial F^*(-\varphi^{c_\varepsilon})$*).*

Assume furthermore that F *has non-empty subgradient at two measures* ρ^- *and* ρ^+ *such that* $\rho^-(\Omega) < 1$ *and* $\rho^+(\Omega) > 1$*, and that* μ *is supported on a compact set. Then, the supremum is attained.*

The definition of the Legendre transform F^* of F and of the c_ε-transform φ^{c_ε} of φ are given respectively in Appendix B and Section 1.1, Definition 2.

Proof. Take

$$G : \rho \in \mathcal{M}(\Omega) \mapsto \begin{cases} I_{c_\varepsilon}(\rho, \mu) & \text{if } \rho \in \mathcal{P}(\Omega) \\ +\infty & \text{otherwise} \end{cases}$$

For $\rho \in \mathcal{M}(\Omega)$, $G(\rho) = \sup_{\psi \in \mathcal{C}^0(\Omega)} \int_\Omega \psi d\rho + \int_{\mathbb{R}^d} \psi^{c_\varepsilon} d\mu = \bar{G}^*(\rho)$, the Legendre transform of the convex continuous function

$$\bar{G} : \psi \in \mathcal{C}^0(\Omega) \mapsto -\int_\Omega \psi^{c_\varepsilon} d\mu.$$

Therefore (this is Fenchel-Moreau's theorem Theorem 45 in Appendix B), $G^* = \bar{G}$.

We now have G^* a convex continuous function, and F^* which is convex l.s.c and not $+\infty$ everywhere. Applying Fenchel-Rockafellar duality theorem to the following infimum problem:

$$\inf_{\varphi \in \mathcal{C}^0(\Omega)} G^*(\varphi) + F^*(-\varphi),$$

we can write:

$$\begin{aligned} \inf_{\varphi \in \mathcal{C}^0(\Omega)} G^*(\varphi) + F^*(-\varphi) &= \max_{\rho \in \mathcal{M}(\Omega)} -G(\rho) - F(\rho) \\ &= -\min_{\rho \in \mathcal{M}(\Omega)} G(\rho) + F(\rho) \\ &= -\min_{\rho \in \mathcal{P}(\Omega)} \frac{W_2^2(\rho, \mu)}{2\varepsilon} + F(\rho) \end{aligned}$$

But, this inf problem also rewrites:

$$\begin{aligned} \inf_{\varphi \in \mathcal{C}^0(\Omega)} G^*(\varphi) + F^*(-\varphi) &= -\sup_{\varphi \in \mathcal{C}^0(\Omega)} \int_\Omega \varphi^{c_\varepsilon}(x) d\mu(x) - F^*(-\varphi) \\ &= -\sup_{\varphi \ c_\varepsilon\text{-concave}} \int_\Omega \varphi^{c_\varepsilon}(x) d\mu(x) - F^*(-\varphi) \end{aligned}$$

The reason why optimization can be done on c_ε is similar to that for optimal transport. Indeed, for any $\phi \in \mathcal{C}^0(\Omega)$, $F^*(-\varphi) \geq F^*(-(\varphi^{c_\varepsilon})^{c_\varepsilon})$ since $\varphi \leq (\varphi^{c_\varepsilon})^{c_\varepsilon}$ and $\partial F^* \subset \mathrm{dom}(F)$ is composed of positive measures only, by hypothesis (F^* is increasing). Therefore, using the fact that $\varphi^{c_\varepsilon c_\varepsilon c_\varepsilon} \equiv \varphi^{c_\varepsilon}$, our supremum is, in the end:

$$\begin{aligned} - \sup_{\varphi \in \mathcal{C}^0(\Omega)} \int_\Omega \varphi^{c_\varepsilon}(x) d\mu(x) - F^*(-\varphi^{c_\varepsilon c_\varepsilon}) \\ = - \sup_{\varphi\ c_\varepsilon\text{-concave}} \int_\Omega \varphi^{c_\varepsilon}(x) d\mu(x) - F^*(-\varphi). \end{aligned}$$

Finally, a change of variable $\varphi \mapsto \varphi^{c_\varepsilon}$ (which brings us to a supremum on c_ε-concave functions on $\mathrm{spt}(\mu)$) we obtain the primal and dual problems that we claimed:

$$\min_{\rho \in \mathcal{P}(\Omega)} \frac{W_2^2(\rho, \mu)}{2\varepsilon} + F(\rho) = \sup_{\varphi\ c_\varepsilon\text{-concave}} \int_\Omega \varphi(x) d\mu(x) - F^*(-\varphi^{c_\varepsilon})$$

Optimality conditions for both problems are straightforward. Indeed, for every φ and ρ, both admissible for their respective problem, we have

$$\int_\Omega \varphi d\mu + \int_\Omega \varphi^{c_\varepsilon} d\rho \leq \frac{W_2^2(\mu, \rho)}{2\varepsilon}$$

with equality if and only if $(\varphi, \varphi^{c_\varepsilon})$ are Kantorovich potentials for the transport from μ to ρ, and,

$$F^*(-\varphi^{c_\varepsilon}) + F(\rho) \geq - \int_\Omega \varphi^{c_\varepsilon}(x) d\rho(x)$$

with equality iff $\rho \in \partial F^*(-\varphi^{c_\varepsilon})$. Summing up these inequalities, and canceling the opposite terms, we get exactly

$$F(\rho) + \frac{W_2^2(\rho; \mu)}{2\varepsilon} \geq \int_\Omega \varphi(x) d\mu(x) - F^*(-\varphi^{c_\varepsilon})$$

with equality if and only if $(\varphi, \varphi^{c_\varepsilon})$ are Kantorovich potentials for the transport from μ to ρ and, $\rho \in \partial F^*(-\varphi^{c_\varepsilon})$.

Now, for the existence part of the proposition, let K be the compact support of μ. Following the standard method in calculus of variations, consider a maximizing sequence of c_ε-concave functions for the dual problem, $(\varphi_n)_n$. These functions all have the same Lipschitz constant as c_ε, on the compact set K. Let L be this common Lipschitz constant, and for $n \in \mathbb{N}$ $M_n = \max_K \varphi_n$, so that for

any $y \in K$, $M_n - L\mathrm{diam}(K) \leq \varphi_n(y) \leq M_n$. Setting $C = \max_{x\in K,\ y\in\Omega} c_\varepsilon(x,y)$, this gives us the bounds, for any $y \in \Omega$ and $n \in \mathbb{N}$,

$$-M_n \leq \varphi_n^{c_\varepsilon}(y) \leq C - M_n + L.\mathrm{diam}(K) = A - M_n,$$

the constant A depending only on ε, Ω and the discrete measure μ.

Assume now that M_n diverges towards $+\infty$ as $n \to \infty$ (this is equivalent to $(\varphi_n)_n$ not uniformly bounded from above on K). Since there exists φ^+ such that $\varphi^+ \in \partial F(\rho^+)$ with ρ^+ of mass strictly more than 1, or equivalently, $\rho^+ \in \partial F^*(\varphi^+)$, we can write:

$$\begin{aligned}\int_\Omega \varphi_n(x)d\mu(x) - F^*(-\varphi_n^{c_\varepsilon}) &\leq M_n - F^*(\varphi^+) - \int_\Omega (-\varphi_n^{c_\varepsilon} - \varphi^+)d\rho^+ \\ &\leq M_n(1-\rho^+(\Omega)) + A\rho^+(\Omega) \\ &\quad - F^*(\varphi^+) + \int_\Omega \varphi^+ d\rho^+\end{aligned}$$

and that last part diverges to $-\infty$ as $n \to \infty$, which is absurd since $(\varphi_n)_n$ is a maximizing sequence. Similarly, if M_n diverges towards $-\infty$ as $N \to \infty$, the fact that F has a non-empty subgradient at a measure of mass strictly less than 1 gives us again that φ_n cannot be a maximizing sequence.

Therefore, $(\varphi_n)_n$ is uniformly bounded and, using Arzela-Ascoli theorem, we can extract from it a subsequence that converges uniformly on K, as $n \to \infty$. By upper semi-continuity of the functions in the dual problem, this limit is a maximizer. □

Remark 2. The hypotheses on F are very natural ones considering our congestion terms have the integral form (1.14) in our numerical simulations. However, they are not the sharpest ones, even when F has this form, to obtain existence. Indeed, consider a domain with Lebesgue measure 1, and F of the form $F(\rho) = \int_\Omega f(\rho(x))dx$ with $f = \chi_{\{1\}}$ (only a density equal to 1 almost everywhere is allowed). One can check that the dual problem admits solutions which are the classical Kantorovich potentials for the corresponding transport, since $F_\varepsilon = \frac{W_2^2(.,dx)}{2\varepsilon}$. However, F itself does not satisfy the assumptions in proposition 7, since it is only finite at the Lebesgue measure which is of mass exactly 1.

Allowing F to be very general can allow the use of more complex penalizations, such as $F(\rho) = \int_\Omega \frac{\|\nabla\rho\|^2}{\rho}$ which can appear when viewing second-order mean field games with entropy penalization as first order mean field games (this is done in particular in [28] or [16]). However in our numerical simulations, we

focus on the cases where F has the simpler integral form:

$$F(\rho) = \begin{cases} \int_\Omega f(\rho(x))dx & \text{if } \rho \ll dx \\ +\infty & \text{otherwise.} \end{cases} \tag{1.14}$$

mentioned earlier. From now on, and until the end of this chapter, we assume that F has this integral form and we make the following assumptions on the function f, which guarantee that F be convex, lower semi-continuous for the narrow topology (see for instance [63], proposition 7.7) and that $\mathrm{dom}(F) \subset \mathbb{R}^+$:

1. f is convex and lower semi-continuous (for the standard topology on $\mathbb{R}$).
2. $\mathrm{dom}(f) \subset \mathbb{R}^+$
3. f is superlinear, $\lim_{x\to+\infty} f(x)/x = +\infty$.

Our hypotheses on ∂F guaranteeing existence of solutions to the dual formulation are equivalent to the simpler assumption that $\frac{1}{|\Omega|}$ be in the interior of $\mathrm{dom}(f)$ as the following proposition justifies:

Lemma 1. *With our hypotheses on F, let $\varphi \in \mathcal{C}^0(\Omega)$, and $\rho \in \mathcal{M}(\Omega)$, then $\varphi \in \partial F(\rho)$ if and only if $\rho \ll dx$ and for a.e $x \in \Omega$, $\varphi(x) \in \partial f(\rho(x))$. In fact:*

$$F^* : \varphi \in \mathcal{C}^0(\Omega) \mapsto \int_\Omega f^*(\varphi(x))\mathrm{d}x$$

Proof. Let $\rho \in \mathcal{M}(\Omega)$, $\varphi \in \mathcal{C}^0(\Omega)$. Then, by definition of the Legendre transforms F^* and f^*:

$$\begin{aligned} F^*(\varphi) &= \sup_{\rho\in\mathcal{M}(\Omega)} \int_\Omega \varphi(x)\mathrm{d}\rho(x)x - F(\rho) \\ &= \sup_{\rho\ll\mathrm{d}x} \int_\Omega \varphi(x)\rho(x) - f(\rho(x))\mathrm{d}x \\ &= \int_\Omega f^*(\varphi(x))\mathrm{d}x \end{aligned}$$

On the other hand, $\varphi \in \partial F(\rho)$ if and only if $\rho \ll dx$ and

$$\int_\Omega f(\rho(x))dx + \int_\Omega f^*(\varphi(x))dx = \int_\Omega \varphi(x)\rho(x)dx \tag{1.15}$$

Now, for a.e. $x \in \Omega$, there holds $f(\rho(x)) + f^*(\varphi(x)) \geq \varphi(x)\rho(x)$, therefore (1.15) is equivalent to

$$f(\rho(x)) + f^*(\varphi(x)) = \varphi(x)\rho(x), \text{ for a.e. } x \text{ in } \Omega \tag{1.16}$$

which is itself equivalent to $\varphi(x) \in \partial f(\rho(x))$ almost everywhere on Ω. □

We now state the corresponding dual problem for this simpler form of F. Notice that the integral form of the function transforms the optimality condition $\rho \in \partial F(-\phi^{c_\varepsilon})$ into an equality almost everywhere, completely determining the optimal ρ for the primal problem in this case:

Proposition 8. *For $\mu \in \mathcal{P}(\mathbb{R}^d)$:*

$$F_\varepsilon(\mu) = \sup_{\phi \ c_\varepsilon - concave} \int_{\mathbb{R}^d} \phi \mathrm{d}\mu - \int_\Omega f^*(-\phi^{c_\varepsilon}(x))\mathrm{d}x \tag{1.17}$$

Assume that $\frac{1}{|\Omega|} \in \mathrm{int}(\mathrm{dom}(f))$ and that $\mu \in \mathcal{P}(\mathbb{R}^d)$ has compact support, then the supremum in (1.17) *is attained.*

Finally, we write $\mu = \mu^{\mathrm{ac}}\mathrm{d}x + \mu^{\mathrm{sing}}$ the decomposition of μ into absolutely continuous and singular parts, with $\mathrm{d}x \perp \mu^{\mathrm{sing}}$. Consider $\rho \in \mathcal{P}(\Omega)$ (resp ϕ c_ε-concave) optimal for the primal formulation for $F_\varepsilon(\mu)$, (1.10) *(resp for the dual formulation* (1.17)*). Then for a.e. $x \in \Omega$, exactly one of the following is true:*

- $\rho(x) = (f^*)'(-\phi^{c_\varepsilon}(x))$ *and* $\nabla\phi^{c_\varepsilon}(x) \neq 0$.
- $\rho(x) = \mu^{\mathrm{ac}}(x)$ *and* $\nabla\phi^{c_\varepsilon}(x) = 0$.

In particular, the optimal ρ is unique.

Remark 3. Uniqueness of ρ comes essentially from the (Lipschitz) regularity of ϕ^{c_ε}. As such regularity cannot be expected for ρ, the dual problem (1.17) could admit multiple solutions ϕ.

The form of the optimal ρ at the end of the previous proposition exactly states, in the general case, that the projection ρ saturates the constraint enforced by F. This is a result already oserved by A. R. Mészáros et al. in [20] as a consequence of A. Figalli [25]. In their case $f = \chi_{[0;1]}$ and the optimal ρ is the projection of μ on the set of constraint-abiding densities $\rho \leq \mathrm{d}x$. We note that the first possible expression for ρ simplifies into $\rho(x) = 1$ almost everywhere on $\{\nabla\phi^{c_\varepsilon} \neq 0\}$, which is the highest possible value permitted.

Proof. The dual formulation 1.17 and the existence result are direct consequences of Lemma 1 and the same results in Proposition 7. We will only make the remark that

$$\begin{aligned}
&\frac{1}{|\Omega|} \in \mathrm{int}(\mathrm{dom}(f)) \\
\Leftrightarrow &\exists \varepsilon > 0,\ \partial f\left(\frac{1}{|\Omega|} - \varepsilon\right) \neq \emptyset \text{ and } \partial f\left(\frac{1}{|\Omega|} + \varepsilon\right) \neq \emptyset \\
\Leftrightarrow &\exists \varepsilon > 0,\ \partial F\left(\left(\frac{1}{\Omega} - \varepsilon\right)\mathbb{1}_\Omega \mathrm{d}x\right) \neq \emptyset \text{ and } \partial F\left(\left(\frac{1}{\Omega} + \varepsilon\right)\mathbb{1}_\Omega \mathrm{d}x\right) \neq \emptyset
\end{aligned}$$

and this last line is the assumption on ∂F sufficient to have existence in Proposition 7, with the measures $\rho^+ = \left(\frac{1}{\Omega} + \varepsilon\right) \mathbb{1}_\Omega \mathrm{d}x$ and $\rho^- = \left(\frac{1}{\Omega} - \varepsilon\right) \mathbb{1}_\Omega \mathrm{d}x$.

Take ρ, ϕ optimal for the primal and dual problems. Using the optimality conditions from proposition 7, recall that $\phi, \phi^{c_\varepsilon}$ are c_ε-concave Kantorovich potentials for the transport from μ to ρ and for a.e. $x \in \Omega$, $\rho(x) \in \partial f^*(-\phi^{c_\varepsilon}(x))$ (again, this is the conclusion of Proposition 7 with the equivalent form of Lemma 1). In particular, the optimal transport from ρ to μ is given by the map defined Lebesgue-almost-everywhere $T = Id - \nabla\phi^{c_\varepsilon}$, using Brenier's theorem, and the fact that c_ε trivially verifies the twist condition, on $\mathbb{R}^d \times \mathbb{R}^d$ even. Furthermore, for $\lambda \in \mathbb{R}$, the set

$$A_\lambda := \{x \in \Omega,\ \phi^{c_\varepsilon}(x) = \lambda \text{ and } \nabla\phi^{c_\varepsilon}(x) \text{ exists and is not } 0\}$$

is Lebesgue negligible. This last claim follows from the co-area formula, as

$$\mathcal{H}^d(A_\lambda) = \int_{-\infty}^{+\infty} \int_{\phi^{c_\varepsilon}=t} \frac{\mathbb{1}_{A_\lambda}(x)}{\|\nabla\phi^{c_\varepsilon}(x)\|} \mathrm{d}\mathcal{H}^{d-1}(x)\mathrm{d}t = 0$$

since the integrand is nonzero if and only if $t = \lambda$. Therefore, for almost any $x \in \Omega$, we can assume that $\nabla\phi^{c_\varepsilon}(x)$ is well-defined and, if $\nabla\phi^{c_\varepsilon}(x) \neq 0$, $\partial f^*(-\phi^{c_\varepsilon}(x))$ is a singleton (since there only are a countable amount of non-differentiability points for f^*, Appendix B, Theorem 44). We can conclude in that case that

$$\rho(x) = (f^*)'(-\phi^{c_\varepsilon}(x)).$$

Otherwise, if $\nabla\phi^{c_\varepsilon}(x) = 0$, the optimal transport map T is simply Id at x, implying that $\rho \leq \mu^{\mathrm{ac}}\mathrm{d}x$ on $A = \{\nabla\phi^{c_\varepsilon} = 0\}$. However, $T = \nabla u$ for a convex function u, Lebesgue-almost everywhere on the set A and therefore μ^{ac}-almost everywhere on A, $T^{-1}(x) = \{\nabla u^*(x)\} = \{x\}$ and $\rho(x) = \mu^{\mathrm{ac}}(x)$. □

1.3 Computations in the semi-discrete setting

In this section, we show how one can compute the value of $F_\varepsilon(\mu)$ when μ is a discrete measure on $\mathbb{R}^d$, for instance, μ has the form $e_t\#Q$ for some $Q \in \mathcal{P}_N(\Gamma)$ (with the definitions of e_t and Γ from pages 51-52 of the next chapter). Reasoning in a very similar fashion as for semi-discrete optimal transport, we then rewrite Eq. (1.17) as a concave finite dimensional problem over $\mathbb{R}^N$. To solve this problem, a Newton-like algorithm seems appropriate, using the first and second order derivatives computed in Propositions 10 and 12 below. For convenience, we rewrite the Moreau envelope F_ε as a function of the positions of the

Dirac masses in a uniform discrete measure, instead of the measure itself: For $Y \in (\mathbb{R}^d)^N$,

$$\mathcal{F}_\varepsilon(Y) := F_\varepsilon\left(\frac{1}{N}\sum_{i=1}^N \delta_{y_i}\right)$$

Proposition 9. *Assume that* $\frac{1}{|\Omega|} \in \mathrm{int}(\mathrm{dom}(f))$.

Then, for $Y \in (\mathbb{R}^d)^N$,

$$\mathcal{F}_\varepsilon(Y) = \max_{\Phi\in\mathbb{R}^N} \sum_{i=1}^N \left[\frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} f^*\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right) dx\right] \tag{1.18}$$

with the definition of the Laguerre cells $\mathrm{Lag}_i(Y,\Phi)$ *from Section 1.1, Definition 6.*

A pair $\rho \in \mathcal{P}(\Omega)$ *and* $\Phi \in \mathbb{R}^N$ *are optimal for respectively the primal and dual problems defining* $\mathcal{F}_\varepsilon(Y)$ *if and only if the following conditions hold for every* $i = 1 \dots N$*:*

- ***(Area)*** $\int_{\mathrm{Lag}_i(Y,\Phi)} (f^*)'(\phi_i - c_\varepsilon(x,y_i))dx = \frac{1}{N}$
- ***(Density)*** $\rho(x) = (f^*)'(\phi_i - c_\varepsilon(x,y_i))$*, for a.e.* $x \in \mathrm{Lag}_i(Y,\Phi)$

Proof. For any $Y \in (\mathbb{R}^d)^N$,

$$\mathcal{F}_\varepsilon(Y) \le \sup_{\Phi\in\mathbb{R}^N} \sum_{i=1}^N \left[\frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} f^*(\phi_i - c_\varepsilon(x,y_i))dx\right]. \tag{1.19}$$

Indeed, let φ be solution of problem (1.13) for $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$. For any $x \in \Omega$,

$$\varphi^{c_\varepsilon}(x) = \inf_{y\in\mathbb{R}^d} c_\varepsilon(x,y) - \varphi(y) \le \inf_i c_\varepsilon(x,y_i) - \varphi(y_i)$$

and since f^* is non-decreasing, inequality (1.19) is straightforward.

For the other inequality, similar arguments as for proposition 7 guarantee that the supremum in (1.19) is indeed a maximum, and that $\Phi \in \mathbb{R}^N$ is optimal if and only if it satisfies ***(Area)***. Take such a Φ in $\mathbb{R}^N$ optimal, then for any $\rho' \in \mathcal{P}(\Omega)$,

$\rho' \ll dx$

$$\begin{aligned}\frac{W_2^2(\rho',\mu)}{2\varepsilon} + \int_\Omega f(\rho'(x))dx &\geq \sum_{i=1}^N \frac{\phi_i}{N} + \int_{\mathrm{Lag}_i(Y,\Phi)} (c_\varepsilon(x,y_i) - \phi_i)d\rho'(x) \\ &\qquad + \int_{\mathrm{Lag}_i(Y,\Phi)} f(\rho'(x))dx \\ &= \sum_{i=1}^N \frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} (\phi_i - c_\varepsilon(x,y_i))\rho'(x) \\ &\qquad - f(\rho'(x))dx \\ &\geq \sum_{i=1}^N \frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} f^*(\phi_i - c_\varepsilon(x,y_i))dx\end{aligned}$$

The first inequality comes from the dual formulation of optimal transport in the semi-discrete case, Eq. (1.8), and the second one from the definition of f^*. Taking the infimum in ρ' yields the reverse inequality.

To conclude, we notice as before that this is an equality if and only if:

(1) the optimal transport from ρ' to μ is given by the Laguerre cells $\mathrm{Lag}_i(Y,\Phi)$.

and

(2) $\rho(x)$ belongs to $\partial f^*(\phi_i - c_\varepsilon(x,y_i))$ for every $i = 1 \dots N$ and Lebesgue a.e. $x \in \mathrm{Lag}_i(Y,\Phi)$.

Since $\sup_i(\phi_i - c_\varepsilon(.,y_i))$ has nonzero gradient almost everywhere on Ω, we are in the first case of proposition 8 and we get similarly that (2) is equivalent to $\rho(x) = (f^*)'(\sup_i \phi_i - c_\varepsilon(x,y_i))$ for a.e. $x \in \Omega$ which is exactly ***(Density)***. But then, ***(Area)*** is exactly equivalent to (1) (see the optimality conditions in the semi-discrete setting in Section 1.1). This concludes the proof. □

As we mentioned, the expression maximized in (1.18) is differentiable in Φ (as well as in Y almost everywhere). We end this section by computing these derivatives, necessary in order to numerically approximate the Moreau-Yosida projections of discrete measures in $\mathcal{P}_N(\mathbb{R}^d)$, as well as solutions of the fully discrete mean field game problem $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$ of Chapter 2, Section 2.3. We denote by G_ε the maximized functional:

$$G_\varepsilon : (Y,\Phi) \in \mathbb{R}^N \times (\mathbb{R}^d)^N \mapsto \sum_{i=1}^N \left[\frac{\phi_i}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} f^*\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right) dx \right]$$

Proposition 10. *G_ε is $\mathcal{C}^1$, concave with respect to Φ and for $i = 1..N$, and $Y \in (\mathbb{R}^d)^N$ and $\Phi \in \mathbb{R}^N$, we have*

$$\frac{\partial}{\partial \phi_i} G_\varepsilon(Y, \Phi) = \frac{1}{N} - \int_{\mathrm{Lag}_i(Y,\Phi)} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x$$

Proof. The proof of this claim is a variation of the corresponding one for semi-discrete transport, by Aurenhammer, Hoffmann and Aronov [6]. Take, $Y \in (\mathbb{R}^d)^N$ and $\Phi,\ \Psi \in \mathbb{R}^N$:

$$\begin{aligned} G_\varepsilon(Y, \Phi) - G_\varepsilon(Y, \Psi) = \sum_{i=1}^N \Bigg[\frac{\phi_i - \psi_i}{N} - \Bigg(& \int_{\mathrm{Lag}_i(Y,\Phi)} f^* \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \\ & - \int_{\mathrm{Lag}_i(Y,\Psi)} f^* \left(\psi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \Bigg) \Bigg] \\ \leq \sum_{i=1}^N \Bigg[\frac{\phi_i - \psi_i}{N} - \Bigg(& \int_{\mathrm{Lag}_i(Y,\Psi)} f^* \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \\ & - f^* \left(\psi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \Bigg) \Bigg] \end{aligned}$$

and

$$\begin{aligned} G_\varepsilon(Y, \Phi) - G_\varepsilon(Y, \Psi) \leq \sum_{i=1}^N \Bigg[& \frac{\phi_i - \psi_i}{N} \\ & - \int_{\mathrm{Lag}_i(Y,\Psi)} (f^*)' \left(\psi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) (\phi_i - \psi_i) \mathrm{d}x \Bigg] \end{aligned}$$

The first inequality comes from the definition of $\mathrm{Lag}_i(Y, \Phi)$ and the increasing nature of f^* and the second one, from the convexity of f^* and the fact that

$$-\Psi^{c_\varepsilon} := \max_j \psi_j - c(., y_j)$$

has Lebesgue-negligible level sets. Therefore, G_ε is concave and the vector

$$g_\varepsilon(\Psi) := \left(\frac{1}{N} - \int_{\mathrm{Lag}_i(Y,\Psi)} (f^*)' \left(\psi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \right)_{i=1..N}$$

belongs to the supergradient $\partial^+ G_\varepsilon(Y, \Psi)$ for every Ψ (see Appendix B, Definition 15). The function $(f^*)'$ is continuous except on an at most countable set of

reals, and the level sets of $x \mapsto \psi_i - \frac{||x-y_i||^2}{2\varepsilon}$ are negligible, as are the boundaries of the Laguerre cells. By dominated convergence, $g_\varepsilon(\Psi)$ defines a continuous (w.r.t. Ψ) selection of supergradients for $G_\varepsilon(Y,.)$ and therefore, this last function is $\mathcal{C}^1$ for any Y, with the partial derivatives that we claimed. □

Remark 4. G_ε being a concave function in Φ, this proposition is another way to show the optimality condition we claimed earlier, namely, for $i = 1 \dots N$:

$$\int_{\mathrm{Lag}_i(Y,\Phi)} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x = \frac{1}{N}.$$

We also note that these expressions are very reminiscent of those obtained by Bourne, Wirth and Schmitzer in [11] for unbalanced semi-discrete optimal transport. Here, our congestion penalization plays the role of the mass discrepancy penalization between our discrete measure μ and the Lebesgue measure on Ω, which does not have mass 1 under the assumptions of proposition 9. This suggests that one could rewrite F_ε as an unbalanced transport problem between these measures.

Proposition 11 (Gradient of F_ε)**.** *For $Y = (y_1, ..., y_N) \in (\mathbb{R}^d)^N$, let ρ_Y, Φ_Y be optimal for the primal and dual problem defining $\mathcal{F}_\varepsilon(Y)$. Finally, we denote $B(Y) = (b_1(Y), \dots, b_N(Y)$ the barycenters of the Laguerre cells, according to the probability measure (on this cell) $N.\rho_Y$:*

$$b_i(Y) = N \int_{\mathrm{Lag}_i(Y,\Phi_Y)} x d\rho_Y(x) \tag{1.20}$$

Then, $\mathcal{F}_\varepsilon \in \mathcal{C}^1((\mathbb{R}^d)^N \setminus \mathbb{D}_N)$, where we must remove the diagonal set:

$$\mathbb{D}_N = \left\{ Y \in (\mathbb{R}^d)^N \mid \exists i \neq j,\ y_i = y_j \right\} \tag{1.21}$$

For every $Y \notin \mathbb{D}_N$, we have

$$\frac{\partial}{\partial y_i} \mathcal{F}_\varepsilon(Y) = \frac{y_i - b_i(Y)}{N\varepsilon}$$

Proof. We show that $\tilde{\mathcal{F}}_\varepsilon : Y \mapsto \mathcal{F}_\varepsilon(Y) - \frac{1}{2N\varepsilon} \sum_{i=1}^N \|y_i\|^2$ is concave on $(\mathbb{R}^d)^N$ with the vector $-\frac{1}{N\varepsilon}(b_i(Y))_i$ in its supergradient at Y ($\mathcal{F}_\varepsilon$ is semi-concave in the sense of Appendix B, Definition 18). Then showing that this is a selection of supergradients which is continuous on $\mathbb{D}_N$, will, as previously, prove that $\tilde{\mathcal{F}}_\varepsilon$ and therefore $\mathcal{F}_\varepsilon$ is $\mathcal{C}^1$ on this set. The supergradient is still valid in configurations where several

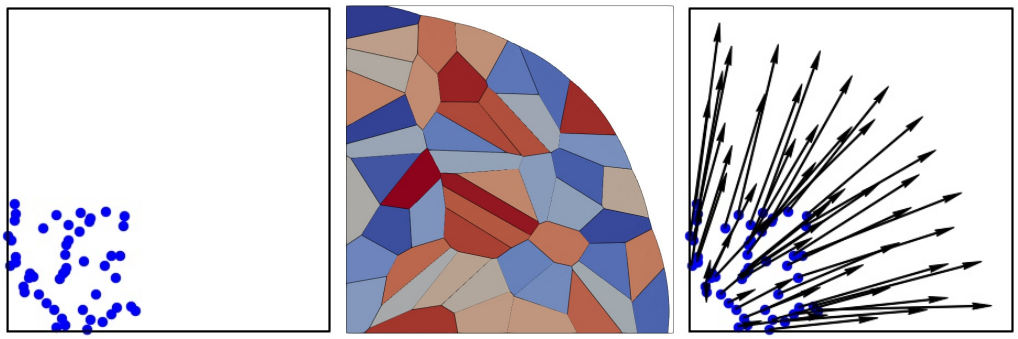


Figure 1.1: From left to right, (1) 50 points scattered in the bottom left corner of a 5×5 square, (2) the "charged" Laguerre cells obtained by intersecting the Laguerre cells with the support of the optimal ρ defining F_ε and (3) the vectors joining each point to the barycenter of its Laguerre cell. $\partial_{y_i} F_\varepsilon$ is colinear, opposite, to the corresponding vector $b(Y)_i - y_i$, for $i = 1 \dots N$.

points coincide, but continuity, and more generally, the fact that the supergradient is a singleton, does not hold for such points.

First, let us take $X, Y \in (\mathbb{R}^d)^N$,

$$\begin{aligned}
\tilde{\mathcal{F}}_\varepsilon(Y) \leq & \frac{1}{2\varepsilon} W_2^2 \left(\frac{1}{N} \sum_{i=1}^N \delta_{y_i}, \rho_X \right) + F(\rho_X) - \frac{1}{2N\varepsilon} \sum_{i=1}^N \|y_i\|^2 \\
\leq & \sum_{i=1}^N \int_{\mathrm{Lag}_i(X,\Phi_X)} \frac{1}{2\varepsilon} (||z||^2 - 2z \cdot y_i + ||y_i||^2) \rho_X(z) + f(\rho_X(z)) dz \\
& - \frac{1}{2N\varepsilon} \sum_{i=1}^N \|y_i\|^2 \\
\leq & \tilde{\mathcal{F}}_\varepsilon(X) - \sum_{i=1}^N \frac{1}{\varepsilon} \int_{\mathrm{Lag}_i(X,\Phi_X)} z d\rho_X(z) \cdot (y_i - x_i) \\
\leq & \tilde{\mathcal{F}}_\varepsilon(X) - \sum_{i=1}^N \frac{b_i(X)}{N\varepsilon} \cdot (y_i - x_i)
\end{aligned}$$

The second inequality might seem counter-intuitive, but it merely states that the Laguerre cells $\mathrm{Lag}_i(X, \Phi_X)$ may yield a greater transport cost than the optimal Laguerre cells (even though Φ is the variable in which the dual formulation is maximized).

That last inequality exactly means that $\tilde{\mathcal{F}}_\varepsilon$ is concave and that for every $X \in (\mathbb{R}^d)^N$,

$$\left(-\frac{1}{N\varepsilon} b_i(X)\right)_{i=1\dots N} \in \partial^+ \tilde{\mathcal{F}}_\varepsilon(X). \tag{1.22}$$

We now show that $B : Y \in (\mathbb{R}^d)^N \setminus \mathbb{D}_N \mapsto B(Y)$ is continuous and therefore, that (1.22) gives a continuous selection of super-gradients:

First, in order to make some parts of the proof more concise, we will write for $Y \in (\mathbb{R}^d)^N$, $\mu_Y = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$ and for $\Phi \in \mathbb{R}^N$, Φ^{c_ε} the continuous bounded function on Ω defined by:

$$\Phi^{c_\varepsilon}(x) = c_\varepsilon(x, y_i) - \phi_i \text{ on the Laguerre cell } \mathrm{Lag}_i(Y, \Phi).$$

Take $(Y^n)_{n\in\mathbb{N}}$ a sequence converging to Y in $(\mathbb{R}^d)^N \setminus \mathbb{D}_N$. Let for all n, ρ^n, Φ^n be optimal for the primal and dual problems defining $\mathcal{F}_\varepsilon(Y^n)$ respectively (ρ^n is uniquely defined for all n). The functions $(\Phi^n)^{c_\varepsilon}$ are c_ε-concave Kantorovich potentials in the transport from ρ^n to μ_{Y^n}, and, for almost every $x \in \Omega$, we have the equality $\rho^n(x) = (f^*)'\left(-(\Phi^n)^{c_\varepsilon}(x)\right)$.

By similar arguments as for Proposition 7, up to a subsequence, $(\Phi^n)_n$ converges towards a $\Phi \in \mathbb{R}^N$ (and therefore, $(\Phi^n)^{c_\varepsilon}$ uniformly converges on Ω towards Φ^{c_ε}). Using the ***(Density)*** condition from Proposition 9, $\rho^n = (f^*)'(-(\Phi^n)^{c_\varepsilon})$ converges almost everywhere (up to a subsequence) towards $\rho = (f^*)'(-\Phi^{c_\varepsilon})$. And since ρ^n is bounded in $\mathbb{L}^\infty(\Omega)$, the convergence is also a weak-* convergence in $\mathbb{L}^\infty(\Omega)$. In particular, Φ^{c_ε} is a c_ε-concave Kantorovich potential for the optimal transport from ρ to μ_Y, since μ_{Y^n} narrowly converges towards μ_Y and ρ^n towards ρ in their respective spaces. But then $\rho = \rho_Y$ is the unique minimizer for the primal problem. In particular, $b_i(Y) = N\int_{\mathrm{Lag}_i(Y,\Phi)} x\rho(x)dx$.

Finally, for all i and n,

$$\begin{aligned}
\|b_i(Y^n) - b_i(Y)\| &= N\left\|\int_{\mathrm{Lag}_i(Y^n,\Phi^n)} x\rho^n(x)dx - \int_{\mathrm{Lag}_i(Y,\Phi)} x\rho(x)dx\right\| \\
&\lesssim \left\|\int_{\mathrm{Lag}_i(Y^n,\Phi^n)} x\rho^n(x)dx - \int_{\mathrm{Lag}_i(Y,\Phi)} x\rho^n(x)dx\right\| \\
&\quad + \left\|\int_{\mathrm{Lag}_i(Y,\Phi)} x\rho^n(x)dx - \int_{\mathrm{Lag}_i(Y,\Phi)} x\rho(x)dx\right\|
\end{aligned}$$

Now, let T^n (resp T) be the Brenier map for the optimal transport from ρ^n (resp ρ_Y) to μ_{Y^n} (resp μ_Y), with cost c_ε. For n large enough,

$$\begin{aligned}
&\left\|\int_{\mathrm{Lag}_i(Y^n,\Phi^n)} x\rho^n(x)dx - \int_{\mathrm{Lag}_i(Y,\Phi)} x\rho^n(x)dx\right\| \\
&\qquad\qquad \lesssim \sum_{i\neq j} \rho^n\left((T^n = y_i^n \cap T = y_j)\bigcup(T^n = y_j^n \cap T = y_i)\right)
\end{aligned}$$

But since y_j^n converges to y_j for every j, this last term is bounded by $\rho^n(\|T^n - T\| \geq \omega_i)$ with $\omega_i = \frac{1}{2}\inf_{j\neq i}\|y_i - y_j\|\ (> 0)$ and we can write

$$\|b_i(Y^n) - b_i(Y)\| \lesssim \rho^n(\|T^n - T\| \geq \omega_i) + o_{n\to\infty}(1)$$

In a similar fashion as [68], Corollary 5.21, we show that the right-hand side vanishes,

$$\lim_{n\to\infty} \rho^n(\|T^n - T\| \geq \omega_i) = 0$$

which will immediately grant us the continuity of $Y \mapsto B(Y)$ on $\mathbb{D}_N$, and conclude our proof:

Take $\eta > 0$. Lusin theorem allows us to find a (compact) set A such that $\rho_y(\Omega\backslash A) \leq \eta$ and T coincides with a continuous function $\tilde{T}$ on A. Then also $\rho^n(\Omega\backslash A) \leq 2\eta$ for n large enough, by the weak-* convergence in $\mathbb{L}^\infty(\Omega)$ which we established earlier, and,

$$\begin{aligned}\rho^n(\left\|T^n - \tilde{T}\right\| < \omega_i) &\leq \rho^n((\|T^n - T\| < \omega_i) \cap A) + \rho^n(\Omega\backslash A)\\ &\leq \rho^n(\|T^n - T\| < \omega_i) + 2\eta\end{aligned}$$

Usual stability theorems, considering the fact that ρ^n and ρ are absolutely continuous with respect to Lebesgue, state that the optimal transport plan π^n from ρ^n to μ_{y_n} narrowly converges to the optimal transport plan π from ρ_y to μ_y.

Consider then the (open) set $\tilde{A} = \left\{(x,y),\ \left\|y - \tilde{T}(x)\right\| < \omega_i\right\}$. One has

$$\begin{aligned}1 - \eta \leq \rho(\left\|\tilde{T} - T\right\| < \omega_i) = \pi(\tilde{A})&\\ &\leq \liminf \pi^n(\tilde{A})\\ &\leq \liminf \rho^n\big(\left\|T^n - \tilde{T}\right\| < \omega_i\big)\\ &\leq \liminf \rho^n\big(\|T^n - T\| < \omega_i\big) + 2\eta\end{aligned}$$

and, this is true for any $\eta > 0$, hence, $\lim \rho_n(\|T^n - T\| < \omega_i) = 1$ and we have the limit we wanted. □

Note that with a (simpler) similar proof, one can show that, G_ε is $\mathcal{C}^1$ with respect to Y on $\mathbb{R}^d \setminus \mathbb{D}_N$, with the same partial derivatives:

For $Y \notin \mathbb{D}_N$ and $\Phi \in \mathbb{R}^N$,

$$\frac{\partial}{\partial y_i} G_\varepsilon(Y,\Phi) = \frac{1}{\varepsilon}\int_{\mathrm{Lag}_i(Y,\Phi)} (y_i - x)(f^*)'\left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon}\right)\mathrm{d}x \qquad (1.23)$$

we omit this proof here, due to this similarity, but we note that G_ε is in fact $\mathcal{C}^1$ on the whole $((\mathbb{R}^d)^N \setminus \mathbb{D}_N) \times \mathbb{R}^N$.

To obtain second order directional derivatives of G_ε w.r.t. Y and Φ, we rely on the approach by De Gournay, Kahn and Lebrat in [18] using their Lemma 1.1 and 1.2 p.9 with a function that depends non-linearly on the positions of the Dirac masses and the weights. In the following proposition and its proof, we denote $\mathrm{Lag}_{ij}(Y,\Phi)$ the intersection of the Laguerre cells $\mathrm{Lag}_i(Y,\Phi)$ and $\mathrm{Lag}_j(Y,\Phi)$:

$$\mathrm{Lag}_{ij}(Y,\Phi) = \mathrm{Lag}_i(Y,\Phi) \cap \mathrm{Lag}_j(Y,\Phi)$$

Proposition 12. *Assume that f^* is $\mathcal{C}^2$ on $\mathbb{R}$. Let $Y \in (\mathbb{R}^d)^N \setminus \mathbb{D}_N$, $\Phi \in \mathbb{R}^N$, then, for $1 \leq i \neq j \leq N$:*

$$\partial_{\phi_i}\partial_{\phi_j} G_\varepsilon(Y,\Phi) = \int_{\mathrm{Lag}_{ij}(Y,\Phi)} \frac{\varepsilon}{\|y_i - y_j\|} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}\mathcal{H}^{d-1}(x) \tag{1.24}$$

$$\begin{aligned} \partial_{\phi_i}\partial_{\phi_i} G_\varepsilon(Y,\Phi) = &- \int_{\mathrm{Lag}_i(Y,\Phi)} (f^*)'' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \\ &- \sum_{j \neq i} \int_{\mathrm{Lag}_{ij}(Y,\Phi)} \frac{\varepsilon}{\|y_i - y_j\|} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}\mathcal{H}^{d-1}(x) \end{aligned} \tag{1.25}$$

$$\begin{aligned} \partial_{\phi_i}\partial_{y_i} G_\varepsilon(Y,\Phi) = &\int_{\mathrm{Lag}_i(Y,\Phi)} \frac{(y_i - x)}{\varepsilon} (f^*)'' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \\ &+ \sum_{j \neq i} \int_{\mathrm{Lag}_{ij}(Y,\Phi)} \frac{y_i - x}{\|y_i - y_j\|} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}\mathcal{H}^{d-1}(x) \end{aligned} \tag{1.26}$$

$$\partial_{\phi_j}\partial_{y_i} G_\varepsilon(Y,\Phi) = - \int_{\mathrm{Lag}_{ij}(Y,\Phi)} \frac{(y_i - x)}{\|y_i - y_j\|} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}\mathcal{H}^{d-1}(x) \tag{1.27}$$

$$\partial_{y_j}\partial_{y_i} G_\varepsilon(Y,\Phi) = \int_{\mathrm{Lag}_{ij}(Y,\Phi)} \frac{(y_i - x)(y_j - x)^T}{\varepsilon \|y_i - y_j\|} (f^*)' \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}\mathcal{H}^{d-1}(x) \tag{1.28}$$

$$\partial_{y_i}\partial_{y_i}G_\varepsilon(Y,\Phi) = Id\int_{\mathrm{Lag}_i(Y,\Phi)}\frac{1}{\varepsilon}(f^*)'\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right)\mathrm{d}x$$
$$-\int_{\mathrm{Lag}_i(Y,\Phi)}\frac{(y_i-x)(y_i-x)^T}{\varepsilon^2}(f^*)''\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right)\mathrm{d}x$$
$$-\sum_{j\neq i}\int_{\mathrm{Lag}_{ij}(Y,\Phi)}\frac{(y_i-x)(y_i-x)^T}{\varepsilon\|y_i-y_j\|}(f^*)'\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right)\mathrm{d}\mathcal{H}^{d-1}(x) \quad (1.29)$$

Proof. As these expressions are all obtained in a very similar fashion, we only prove the one for $\partial_{y_i}\partial_{y_i}G_\varepsilon$ (it is, from our point of view, the most involved one). Let Y, Φ be as in the proposition, $v \in (\mathbb{R}^d)^N$, $\|v\| = 1$. We apply lemma 1.1 of [18] with the function:

$$a : (x,t) \in \Omega \times [-1;1] \mapsto (y_i + tv - x)(f^*)'\left(\phi_i - \frac{\|x-y_i-tv\|^2}{2\varepsilon}\right)$$

which is $\mathcal{C}^1$ on $\Omega \times \mathbb{R}$, bounded. Now (we use the same notations as De Gournay et al. in their paper):

$$\partial_{y_i}\partial_{y_i}\varepsilon G_\varepsilon(Y,\Phi)v = \partial_t\xi(0)$$
$$= -\sum_{j\neq i}\int_{\mathrm{Lag}_{ij}(Y,\Phi)}\frac{\partial_t u^{ij}(x,0)}{\|\nabla_x u^{ij}(x,0)\|}a(x,0)\mathrm{d}\mathcal{H}^{d-1}(x) + \int_{\mathrm{Lag}_i(Y,\Phi)}\partial_t a(x,0) \quad (1.30)$$

where we recall that

- $\xi(t) = \int_{\mathcal{L}_i(t)} a(x,t)\mathrm{d}x$,
- $\mathcal{L}_i(t) = \mathrm{Lag}_i(Y+tv,\Phi)$
- $u^{ij}(x,t) = \frac{\|x-y_i-tv\|^2}{2\varepsilon} - \phi_i - \left(\frac{\|x-y_j\|^2}{2\varepsilon} - \phi_j\right)$

Now, for x in Ω,

$$\partial_t u^{ij}(x,0) = \frac{1}{\varepsilon}v.(y_i - x) \text{ and } \nabla_x u^{ij}(x,0) = \frac{y_i - y_j}{\varepsilon}$$

and of course, for almost every $x \in \Omega$,

$$\begin{aligned}\partial_t a(x,0) =& v(f^*)'\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right)\\ &- \frac{1}{\varepsilon} v\cdot(y_i-x)(y_i-x)(f^*)''\left(\phi_i - \frac{\|x-y_i-tv\|^2}{2\varepsilon}\right)\\ =&\left(Id(f^*)'\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right)\right.\\ &\left.- \frac{1}{\varepsilon}(y_i-x)(y_i-x)^T(f^*)''\left(\phi_i - \frac{\|x-y_i-tv\|^2}{2\varepsilon}\right)\right)v\end{aligned}$$

and replacing these terms in (1.30) (and dividing by ε) yields exactly the right expression for $\partial_{y_i}\partial_{y_j} G_\varepsilon(Y,\Phi)$. □

Let us end this section with a few remarks:

Remark 5. The above formulas are actually valid when f^* is only $W^{2,1}_{loc}$ on $\mathbb{R}$. In this case, an (unpublished) result by Serrin (see [41]) states that $a \in W^{1,1}$, with a derivative given Lebesgue-almost everywhere by the chain rule. Lemma 1.1 of [18] is then still valid and our expressions are unchanged. However, in the case of the function

$$F : \rho \in \mathcal{M}(\Omega) \mapsto \begin{cases} 0 & \text{if } 0 \le \rho \le \mathrm{d}x \\ +\infty & \text{otherwise.}\end{cases}$$

featured in the numerical section of Chapter 2, $(f^*)' \equiv \mathbb{1}_{\mathbb{R}^+}$ is the Heaviside function, which does not satisfy even these weaker conditions.

Another point concerns the two first expressions, where we differentiate with respect to Φ twice. One can notice that the "Hessian" matrix $D^2_\Phi G_\varepsilon$, for fixed $Y \in \mathbb{D}_N$, is diagonally dominant (as the sums of integrals over the borders $\mathrm{Lag}_{ij}(Y,\Phi)$ of $\mathrm{Lag}_i(Y,\Phi)$ cancel each other). This means that this matrix is non-singular provided that the "residual diagonal terms" are non-zero:

$$\int_{\mathrm{Lag}_i(Y,\Phi)} (f^*)''\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right) \neq 0$$

In this case, the optimal Φ is unique, whereas, in the case of the dual formulation of optimal transport in the semi-discrete case, it was only unique up to a constant $\lambda(1,\dots,1)$. One can then use a Newton algorithm in order to approximate

the optimal weights Φ for the dual formulation of $\mathcal{F}_Y$ (instead of a damped Newton one, as for semi-discrete OT). This is used in particular in the Pysdot library for Python in order to estimate the value of $\mathcal{F}_\varepsilon$ and its gradient, and therefore in the numerical simulations of Chapter 2 (among other tricks from computational geometry to compute these kinds of integrals over Laguerre cells).

Chapter 2

Discretization of Variational Mean Field Games with Congestion

In this chapter we study the motions of a crowd which moves while trying to avoid overcrowded areas. The model we used is defined as the solution to a variational problem, which we discretize according to the trajectories of the members of the crowd (hence the Lagrangian discretization). Convergence results for these discretization can be obtained, and, solving numerically these discrete problems, one can observe behaviours for the finite crowds that are compatible with the congestion constraint that was modelized. These results and observations are gathered in [65].

When dealing with such motions, a well-adapted mathematical setting is to represent the crowd by a curve on the probability space $\mathcal{P}(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is the domain in which the motion takes place:

$$\mu \in \mathcal{C}^0([0;T], \mathcal{P}(\Omega))$$

where continuity is meant for the narrow convergence or equivalently, if we assume that Ω is compact, any Wasserstein distance on $\mathcal{P}(\Omega)$. Governing this motion is often the so-called *continuity equation*:

$$\partial_t \mu + \nabla_x.(\mu v) = 0 \tag{2.1}$$

where $v : [0;T]\times\Omega \to \mathbb{R}^d$ is a vector field representing the velocity of the particles charged by μ. Mathematically speaking, equation (2.1) is verified in the sense of distributions, meaning that for any compactly supported $\phi \in \mathcal{C}_c^\infty(]0;T[\times\Omega, \mathbb{R}^d)$,

$$\int_{[0;T]\times\mathbb{R}^d} (\partial_t \phi + \nabla_x \phi \cdot v) \mathrm{d}\mu(t,x)\mathrm{d}t = 0. \tag{2.2}$$

Note also that often (and in particular in our model), we want to fix the initial distribution of the crowd equal to a $\mu^0 \in \mathcal{P}(\mathbb{R}^d)$. In this case we say that the equation is verified in the sense of distribution with initial data μ^0 to express that:

$$\int_{[0;T]\times\Omega} (\partial_t \phi + \nabla_x \phi \cdot v) \mathrm{d}\mu(t,x)\mathrm{d}t = -\int_{\Omega} \phi(0,x)\mathrm{d}\mu^0(x) \tag{2.3}$$

for any test function $\phi \in \mathcal{C}_c^\infty([0;T[\times\Omega, \mathbb{R}^d)$, and the reader can note that the test functions ϕ can now be non-zero at $t = 0$. In particular, both these definitions imply Neumann boundary conditions $v(t,.)\mu(t,.) = 0$ on $\partial(\Omega)$. In this chapter, we will omit to precise in which sense μ solves (2.1), as it will always be in the sense of distributions with an initial data which will always be obvious.

The continuity equation expresses in particular conservation of the mass of μ, which translates, in the context of crowd motions, to no individual leaving the domain. It is also equivalent to a stronger continuity for the trajectory of the population, μ (we refer the reader to [4] or [63] for a proof):

Theorem 13. *Let $(\mu_t)_{t\in[0;T]}$ a family of measures in $\mathcal{P}_r(\mathbb{R}^d)$ and v in $\mathbb{L}^r(\mathrm{d}\mu_t\mathrm{d}t, \mathbb{R}^d \times [0;T])$, both verifying the continuity equation* (2.1)*.*

Then, $\mu : t \mapsto \mu_t$ is in $\mathcal{AC}([0;T], \mathcal{P}_r(\mathbb{R}^d))$.

On the other hand, if $\mu \in \mathcal{AC}([0;T], \mathcal{P}_r(\mathbb{R}^d))$, then, there exists a vector field $v \in \mathbb{L}^r(\mathrm{d}\mu(t)\mathrm{d}t, \mathbb{R}^d \times [0;T])$ such that (μ, v) solves the continuity equation.

Furthermore, the smallest (in $\mathbb{L}^r$ norm) vector field v introduced in the second part must verify for $\mathrm{d}t$-a.e $t \in [0;T]$, $\|v\|_{\mathbb{L}^r(d\mu(t))} = |\mu'|(t)$, the W_r-metric derivative of μ (see [4], Theorem 1.1.2).

In Theorem 13, the absolute continuity is, of course, meant with respect to the Wasserstein distance W_r on $\mathcal{P}_r$. Let us recall that absolutely continuous curves are continuous paths on a metric space X, which have integrable velocity in the following sense (we refer the reader to the monograph by Ambrosio, Gigli and Savare, [4]):

Definition 7. *Let $\mu : [0;T] \to X$, (X,d) a metric space, then μ is absolutely continuous, $\mu \in \mathcal{AC}([0;T], X)$ when there exists a function $g : [0;T] \to \mathbb{R}^+$ summable, such that for any $0 < t_0 < t_1 < T$,*

$$d(\mu(t_0), \mu(t_1))) \leq \int_{t_0}^{t_1} g(t)dt.$$

The smallest ($\mathrm{d}t$-almost everywhere) such function g is the metric derivative $|\mu'|$ mentioned in Theorem 13.

As we said, the specificity of our crowd motion comes from the fact that the population tries to avoid overcrowded (congested) areas during its movement. Let us mention two ways in which it has been enforced in a mathematical model, aside from the mean field game one which we use:

In Maury et al. [47], [46] one considers a granular model in which each individual q_i is represented by a ball $B(q_i, r_i)$. Then, if the desired velocity of individual i is U_i (for instance $\nabla V(q_i)$ for a potential $V : \Omega \to \mathbb{R}$), the actual velocity of the individual is given by the projection of $U := (U_1, \dots, U_N)$ onto the feasibility condition:

$$(\|q_i - q_j\| - (r_i + r_j))_+ = 0 \implies (U_i - U_j).(q_i - q_j) \geq 0$$

(or a generalisation of this condition in the case of weaker non-congestion conditions).

This model corresponds to the individuals having non-overlaping "zones" around them and having to adapt their trajectory to their proximity to others.

The corresponding "macroscopic model", also presented in [46] as well as [45], (although the authors insist on the fact that it is not obtained from the "microscopic" one above, as the limit model as the radii r_i vanish) is given by the evolution of a density $\rho(t)$ along the system of equations:

$$\begin{cases} \partial_t \rho + \nabla.(\rho u) = 0 \\ u = \Pi_{C_\rho}(U) \end{cases} \tag{2.4}$$

where the continuity equation is verified in the sense of (2.3), with a fixed initial data ρ^0. Here, the desired velocity at x, $U(x)$, is projected on the set of feasible velocity:

$$C_\rho = \left\{ v \in \mathbb{L}^2(\Omega, \mathbb{R}^2), \ \int_\Omega v.\nabla q \leq 0 \ \forall q \in \mathrm{H}^1_\rho \right\}$$

defined by duality with the set of pressures:

$$\mathrm{H}^1_\rho = \left\{ q \in \mathrm{H}^1(\Omega), \ q \geq 0 \text{ a.e. in } \Omega, \ q(x) = 0 \text{ a.e. on } \{\rho < 1\} \right\}.$$

When $U = \nabla V$ is given by the gradient of a potential, this last problem is a case of Wasserstein gradient flow (introduced in the pioneering work of Otto et al. [34] and detailed in [4]), where the gradient descent is done on the energy $\int_\Omega f(\rho(x)) + V(x)\rho(x)\mathrm{d}x$, where $f \equiv \chi_{[0;1]}$ is the convex indicator function of the non-congestion condition on the density $0 \leq \rho \leq 1$. An actual discretization of this problem and the one when $f(r) = r\log(r)$ (in the sense that it features a convergence result as the number of players goes to infinity), is given in Leclerc, Mérigot, Santambrogio and Stra, [40]. This discretization is done using the Moreau-Yosida regularization studied in the previous chapter, (1.10). The

convergence of the discrete measures μ_N, which is one in the sense of the uniform topology associated with the 2-Wasserstein distance, is proven under some assumptions and we present arguments to get rid of some of them in Section 3.4 of the next chapter.

Similarly to the mean field games we study below, the population represented by μ will evolve towards the objective represented by the potential V, without violating the constraint enforced by f. However, in this setting, the players lack the global planning aspect of the solutions to the variationnal problem and numerical computation can feature, in some cases, more selfish behaviours that can generate blockades (think of the narrow corridor in the non-convex example of Section 2.4).

2.1 Variational mean field games in Lagrangian setting

The model we use is inspired by mean field games, which were introduced by Lasry and Lions in [37, 38] and, independentely by Caines, Huang and Malhamé in [33]. This setting focuses on the limit, as its size grows to infinity, of a population which moves in order to minimize a global energy accumulated by its individuals. Under suitable conditions of symmetry on the terms in the energy, one can show that the infinite limit population is distributed according to a density with each infinitesimal player minimizing the same energy along its movement. For these kinds of motions, the global density moves according to a mean field system similar to (2.5) (with an additional second order term in cases where the movement features stochastic parts).

This is the point at which we pick up the model and let us start by considering a simple case where the members of the crowd (which we sometimes name "players", borrowing the term from the game theoretical aspect of this model) evolve in a domain Ω (which we will assume to be compact and smooth), while each trying to minimize an accumulated energy, which for this example we take equal to:

$$\int_0^T \left[\frac{\|x'(t)\|^2}{2} + g(\mu(t, x(t))) + V(x(t)) \right] dt + \Phi(x(T)).$$

Here, $x \in \mathrm{H}([0;T], \Omega)$ represents the trajectory of a player, x' its velocity whereas μ is the density of the population. Congestion, in this case, is penalized through the function g which is non-decreasing on $\mathbb{R}$. Passing through regions where μ is high will then cause the player to have a higher energy (which he wants to avoid). V and Φ are scalar functions on $\mathbb{R}^d$ which define what we will call the

potential energy. To have low energy, players have to pass through areas where V is low (and end their trajectory where Φ is). At equilibrium, each player is following an optimal path, avoiding crowded areas while passing where their potential energy is low. Optimal control theory then links the evolution of the population's density μ to that of the so-called *value function*:

$$\phi : (t_0, x_0) \mapsto \inf \left\{ \int_{t_0}^{T} \left[\frac{\|x'(t)\|^2}{2} + g(\mu(t, x(t))) + V(x(t)) \right] dt + \Phi(x(T)) \;\middle|\; x \in \mathbb{H}^1([t_0; T]),\ x(t_0) = x_0 \right\}$$

via a mean-field game system (with the appropriate boundary conditions):

$$\begin{cases} -\partial_t \phi + \frac{\|\nabla \phi\|^2}{2} = g(\mu) + V \\ \partial_t \mu - \nabla.(\mu \nabla \phi) = 0 \\ \mu(0, .) = \mu^0,\ \phi(T, .) = \Phi \end{cases} \tag{2.5}$$

The idea to obtain a variational form for this mean field game is to consider a primitive f for g, $f' = g$, in such a way that f is a convex function (g was non-decreasing). The first (Hamilton-Jacobi) equation in system (2.5) can then be interpreted as $-\partial_t \phi + \frac{\|\nabla \phi\|^2}{2} - V$ belonging to the subgradient of this convex function f at the values of the density μ_t. The system then states optimality conditions for the following minimization problem, with μ seen as a curve in $\mathcal{C}^0([0; T], \mathcal{P}(\Omega))$:

$$\inf_{\substack{\partial_t \mu + \nabla.(\mu v) = 0 \\ \mu(0) = \mu_0}} \int_0^T \int_\Omega \left[\frac{\|v(t,x)\|^2}{2} \mu(t,x) + f(\mu(t,x)) + V(x)\mu(t,x) \right] dx dt + \int_\Omega \Phi(x)\mu(T,x) dx \tag{2.6}$$

To be slightly more precise, this inf problem is nonconvex, but it can be made convex with the change of variable $w = \mu v$ and then (2.5) become optimality conditions for the pair μ, $w = \mu \nabla \phi$. Of course, solutions to (2.5) are often defined in a very weak sense (as viscosity solutions), and the previous discussion must be adapted in order to be followed rigourously. For an overview of this variational formulation and a rigourous statement and proof of this result, in particular under strong assumptions on f, we refer the reader to the course notes on mean field games by Cardaliaguet [14], the lecture notes by Santambrogio [64] as well

as the survey on variational mean field games by Santambrogio, Carlier and Benamou, [8].

Forsaking the PDE setup of (2.5), we will consider the following, more general, form of Problem (2.6):

$$\inf\{J(\mu, v) \mid \partial_t \mu + \nabla.(\mu v) = 0,\ \mu(0) = \mu_0\} \tag{2.7}$$

where

$$J(\mu, v) = \int_0^T \left[\int_\Omega L(v(t,x))\mu(t,x)dx + F(\mu(t,.)) \right] dt + G(\mu)$$

Now is as good a time as any to be a little more precise on what each term in this energy is and what assumptions we make on them:

- The first term in J, $\int_0^T \int_\Omega L(v(t,x))d\mu(t,x)dt$ is the one we call "kinetic" term and measures the cost of displacement of μ following the velocity field v. It can be interpreted as the crowd getting tired from maintaining a high velocity during its motion and obviously, a discontinuous move brings an infinite amount of fatigue, which from a mathematical standpoint, guarantees smooth trajectories. The standard assumptions, in mean field games, are that $L : \mathbb{R}^d \to \mathbb{R}$ is a convex continuous function on $\mathbb{R}^d$ which behaves like $\|.\|^r$ for some $r > 1$. More precisely, there exists $C > 0$,

$$\forall p \in \mathbb{R}^d,\ \frac{1}{rC}||p||^r - C \leq L(p) \leq \frac{C}{r}||p||^r + C \tag{2.8}$$

- In order to stay consistent with the congestion term in mean field game (2.6), we take F to be a convex function on $\mathcal{M}(\Omega)$ with domain in $\mathcal{M}_+(\Omega)$ (the space of positive finite measures on Ω). The point of the Moreau envelope, described in the previous chapter, is to define a congestion penalization even when F is finite only on a very restricted class of measures. For this reason, we do not need to assume a lot of regularity for F and we make the most basic assumptions in order to have minimizers for Problem (2.7). Specifically, we ask that F be lower semi-continuous for the narrow topology on $\mathcal{M}(\Omega)$ and lower bounded on this set. Upon giving F the integral form from Chapter 1, Section 1.2, (1.14), we recover the congestion term of Problem (2.6). Let us also mention that in the case:

$$F : \mu \in \mathcal{M}(\Omega) \mapsto \int_\Omega \chi_{[0;1]}(\mu(x))\mathrm{d}x = \begin{cases} 0 & \text{if } 0 \leq \mu \leq \mathrm{d}x \\ +\infty & \text{otherwise} \end{cases}$$

we recover the "hard" constraint studied by Maury et al., where crowds are given an infinite (unacceptable) energy when more than one player is at a given position (in a continuous sense). For more general F, this term models the impact, on J, of congestion in the population as described in the initial example. Finally, notice that in (1.14), F is finite only on measures which are absolutely continuous with respect to the Lebesgue measure. We will therefore need to smooth it up in order to compute an analog for discrete populations, which is the *raison d'être* of the Moreau envelope F_ε.

- The last term we call "potential" term and this is an abuse of notations from the case where G derives from an actual potential (V or Φ in the example above). It represents how far (not necessarily in the sense of a distance) the crowd is from a target "preferred" disposition towards which it will try to move. G is a continuous function over the set $\mathcal{AC}([0;T],\mathcal{P}_r(\mathbb{R}^d))$ of absolutely continuous curves with respect to the topology induced by W_r (which is stronger than the one induced by the narrow topology on $\mathcal{C}^0([0;T],\mathcal{P}(\mathbb{R}^d))$).

  Theorem 13 guarantees that any curve in $\mathcal{C}^0([0;T],\mathcal{P}(\mathbb{R}^d))$ with a finite kinetic term is in $\mathcal{AC}([0;T],\mathcal{P}_r(\mathbb{R}^d))$ and therefore that G is well-defined at that curve and more generally, continuous on the set of admissible curves. Therefore, we will not be smoothing G and, to guarantee existence of solutions to the variational mean field game problem, we only make the additional assumption that G is lower bounded. The role of G is to drive the players toward a goal (the trajectories with the lowest potential energy), and this last term often favours congestion, working in opposition to the congestion term F. In our numerical simulations, G will be given as the mean value of a potential across the population, at intermediate times (V in the introductory example above) and/or at final time (Φ in the example).

To approximate solutions to Problem (2.7), we wish to use a finite population of players, whose motion should describe "well enough" the global behaviour of the continuous crowd. Although we could continue using the space $\mathcal{C}([0;T],\mathcal{P}(\mathbb{R}^d))$ and simply replace $\mu(t)$ with a discrete uniform probability measure in $\mathcal{P}_N(\mathbb{R}^d)$ and adapt v, it is more convenient to recast the mean field game problem (2.7) into a minimization problem on the distribution of possible trajectories for the players, $Q \in \mathcal{P}(\mathcal{C}^0([0;T],\mathbb{R}^d))$. We recall here the notation $\Gamma := \mathcal{C}^0([0;T],\mathbb{R}^d)$ for this space of trajectories. Notice that curves in Γ do not have to remain inside Ω, however, since F is $+\infty$ at measures charging points outside Ω, the congestion term will also enforce this restriction in our continuous problem.

A curve in $\mathcal{C}^0([0;T],\mathcal{P}(\Omega))$ admissible for the variational mean field game (2.7)

can always be seen as a probability measure in $\mathcal{P}(\Gamma)$. This is stated by the following theorem by Lisini (see also Theorem 8.2.1 of [4]). This correspondence gives a very natural way of seeing the density of the population at time t, $\mu(t)$, as the distribution Q of the players' trajectories evaluated at time t, $e_t\#Q$, where e_t is the (continuous) operator of *evaluation at time* t on Γ, $e_t : \gamma \in \Gamma \mapsto \gamma(t)$:

Theorem 14 (Lisini, [42])**.** *Let $\mu \in \mathcal{C}([0;T];\mathcal{P}(\mathbb{R}^d))$, be solution (in the sense of distributions) of the continuity equation $\partial_t\mu + \nabla.(\mu v) = 0$, with a $\mathbb{L}^r(d\mu_t dt)$ velocity vector field v and $r > 1$. Then there exists a probability measure $Q \in \mathcal{P}(\Gamma)$ such that:*

1. *Q-almost every $\gamma \in \Gamma$ is in $\mathbb{W}^{1,r}([0;T];\mathbb{R}^d)$ and satisfies $\gamma'(t) = v(t,\gamma(t))$ for $\mathcal{L}^1$-almost every $t \in [0;T]$.*
2. *$\mu(t) = e_t\#Q$ for every $t \in [0;T]$.*

Conversely, any $Q \in \mathcal{P}(\Gamma)$ which satisfies

$$\int_\Gamma \|\gamma'\|^r_{\mathbb{L}^r}\, dQ(\gamma) < +\infty \tag{2.9}$$

induces an absolutely continuous curve in $\mathcal{AC}([0;T];\mathcal{P}_r(\mathbb{R}^d))$, solution to a continuity equation, via $\mu(t) = e_t\#Q$.

We can now define the continuous problem which we will be discretizing using the Moreau envelope from Chapter 1, (1.10). For the rest of this chapter, we fix an initial distribution of players, $\mu^0 \in \mathcal{P}(\Omega)$, admitting a density with respect to the Lebesgue measure on Ω. The variational problem we consider is the rewriting of (2.7) as a minimization problem over $\mathcal{P}(\Gamma)$, using the representation of Theorem 14:

$$(\mathcal{M}_{\mu^0}) : \inf\left\{J(Q) \;\middle|\; Q \in \mathcal{P}(\Gamma) \text{ s.t. } e_0\#Q = \mu^0\right\}$$
$$\text{with } J(Q) = \int_\Gamma \int_0^T L(\gamma'(t))dtdQ(\gamma) + \int_0^T F(e_t\#Q)dt + G(Q).$$

The kinetic and the congestion terms have simply been replaced by their corresponding equivalents on measures on Γ. L therefore stays the same convex continuous function on $\mathbb{R}^d$ verifying inequalities (2.8) and F is the same convex function, now penalizing $e_t\#Q$. Notice that in order for these two terms to be finite, Q must verify the converse implication in Theorem 14 and therefore induce an admissible pair (μ, v) for the Eulerian formulation (2.6), and conversely so. Let us mention here that we will be using the abuse of notation, for $\gamma \in \Gamma$,

$$L(\gamma') := \begin{cases} \int_0^T L(\gamma'(t))dt & \text{if } \gamma \in \mathbb{W}^{1,r}([0;T],\mathbb{R}^d) \\ +\infty & \text{otherwise} \end{cases}$$

in the future as no confusion should arise.

The potential term G in $(\mathcal{M}_{\mu^0})$ is a little more fuzzy: in problem (2.7) it was a continuous function on $\mathcal{AC}([0;T],\mathcal{P}_r(\mathbb{R}^d))$. As we will not revisit this Eulerian formulation, for Problem $(\mathcal{M}_{\mu^0})$, we will take G to be any continuous, lower bounded function on the subset of measures in $\mathcal{P}(\Gamma)$ verifying the finite kinetic energy condition (2.9) of Theorem 14, and $+\infty$ on the remainder of $\mathcal{P}(\Gamma)$ (for simplicity). Should the reader need to go back to the Eulerian setting, Theorem 14 followed by Theorem 13 can give a meaning to G as a continuous (lower bounded) function on $\mathcal{AC}([0;T],\mathcal{P}_r(\mathbb{R}^d))$ as well as equality of the minimal values of both problems.

We now briefly recall why $(\mathcal{M}_{\mu^0})$ admits minimizers. These minimizers then induce minimizers for the Eulerian formulation, using the correspondence of Theorem 14:

Proposition 15. *The functional J above is l.s.c (for the narrow convergence on $\mathcal{P}(\Gamma)$). The minimization problem $(\mathcal{M}_{\mu^0})$ admits solutions.*

In the rest of this paper, $Q_{\min}$ will always denote (any) one of these minimizers for Problem $(\mathcal{M}_{\mu^0})$.

Proof. We first prove that J is l.s.c by treating its terms separately. Theorem 4.5 of [31] and the bounds in (2.8) on L directly imply the lower-semicontinuity on $\mathcal{P}(\Gamma)$ of the kinetic term, $Q \mapsto \int_\Gamma L(\gamma')dQ(\gamma)$.

Take now $(Q_n)_{n\in\mathbb{N}}$ a sequence converging to Q_∞ in $\mathcal{P}(\Gamma)$. To prove the rest of our lower semi-continuity, we may assume that $J(Q_n)$ is upper bounded. By continuity of e_t on Γ, for every t, $e_t\#Q_n$ narrowly converges to $e_t\#Q_\infty$ as n goes to infinity. Thus, by lower semi-continuity of F and Fatou Lemma, we get as desired

$$\int_0^T F(e_t\#Q_\infty)dt \leq \int_0^T \liminf_{N\to\infty} F(e_t\#Q_N)dt \leq \liminf_{N\to\infty} \int_0^T F(e_t\#Q_N)dt.$$

Finally, from the bound on $(J(Q_n))_{n\in\mathbb{N}}$ and the fact that F and G be lower bounded, every Q_n and therefore also Q_∞ are contained in a sub-level set:

$$\left\{ Q \in \mathcal{P}(\Gamma) \;\middle|\; \int_\Gamma \|\gamma'\|_{\mathbb{L}^r}^r \,\mathrm{d}Q(\gamma) < C \right\}$$

for some $C > 0$. G is continuous on this set by assumption, and thus:

$$J(Q_\infty) \leq \liminf_{n\to+\infty} J(Q_n).$$

To show existence of minimizers for $(\mathcal{M}_{\mu^0})$, we follow, again, the standard method in calculus of variations. We first note that for any upper bound $C > 0$, the set

$$K_C := \{\gamma \in \Gamma \mid L(\gamma') \leq C,\ \gamma(0) \in \Omega\}$$

is compact. Indeed, Ω is itself compact and any curve in K_C is Hölder-continuous. Compactness of K_C immediately follows from Arzela-Ascoli's theorem. If we take a minimizing sequence $(Q_n)_n$ for $(\mathcal{M}_{\mu^0})$ we notice that it is tight since for any n and C,

$$Q_n(\Gamma \backslash K_C) < \frac{J(Q_n) - A}{C}$$

for $A = T \inf F + \inf G$.

Using Prokhorov's theorem, we can extract from it a sequence converging, for the narrow topology, to a $Q_\infty \in \mathcal{P}(\Gamma)$ (in particular, $e_0 \# Q_\infty = \mu^0$). But then, by lower-semi-continuity, Q_∞ is a minimizer for our problem since

$$J(Q_\infty) \leq \liminf_{n\to\infty} J(Q_n) = \inf \left\{ J(Q) \ \mid \ Q \in \mathcal{P}(\Gamma) \text{ s.t. } e_0 \# Q = \mu^0 \right\}. \quad \square$$

2.2 Space discretization in $\mathcal{P}(\Gamma)$

Our discretization is Lagrangian in the sense that it aims at approximating the solution $Q_{\min}$ by discrete measures $Q_N \in \mathcal{P}_N(\Gamma)$ following the point of view of Theorem 14. Not having too many informations on $Q_{\min}$, a good way to do so is to look for these discrete measures as minimizers of a variational problem similar to $(\mathcal{M}_{\mu^0})$. For such measures, we can prescribe the initial distribution $e_0 \# Q_N = \mu_N^0 \in \mathcal{P}_N(\mathbb{R}^d)$ and continuity of e_0 tells us that it will have to approximate μ^0 in some sense. Similarly, the kinetic and potential terms are well-defined on "reasonable" discrete measures (meaning supported on $\mathbb{W}^{1,r}([0;T],\mathbb{R}^d)$ curves for L). However, for such discrete distributions of trajectories, the value of the congestion term F could be $+\infty$ (and will be in the specific cases we consider later on), giving us no information as to how close to an optimal uncongested motion we could be!

To avoid this problem, we replace the congestion term by the regularized version of it, F_ε, defined in the previous chapter, (1.10). Proposition 5 tells us that, as $\varepsilon \to 0$, F_ε behaves as F, penalizing congested measures, or ones that are not supported in Ω (since $F(\rho) < +\infty$ implies $\mathrm{spt}(\rho) \subset \Omega$). On the other hand, for any $\varepsilon > 0$, F_ε is finite at discrete measures, penalizing them less and less as ε becomes larger. Since we wish for a regime in the middle (congestion being penalized, but by a finite value), we expect to have to let ε go to zero, but not too fast, in order to have convergence, in some sense, to a minimizer $Q_{\min}$.

The corresponding discretization (or rather regularization) of our energy J is then straightforward, as the kinetic and potential terms are well-defined in this discrete case:

$$(\mathcal{M}_{N,\mu_N^0,\varepsilon}) : \inf \left\{ J_\varepsilon(Q) \mid Q \in \mathcal{P}_N(\Gamma),\ e_0 \# Q = \mu_N^0 \right\}$$
$$\text{where } J_\varepsilon(Q) := \int_\Gamma L(\gamma') dQ(\gamma) + \int_0^T F_\varepsilon(e_t \# Q) dt + G(Q)$$

Immediately, using similar arguments as for Problem $(\mathcal{M}_{\mu^0})$, we have existence of minimizers for these discrete problems:

Proposition 16. *For every $N \in \mathbb{N}^*$, $\varepsilon > 0$, J_ε is l.s.c for the narrow convergence and for every $\mu_N^0 \in \mathcal{P}_N(\Omega)$, the infimum in $(\mathcal{M}_{N,\mu_N^0,\varepsilon})$ is attained.*

As we mentioned, we expect minimizers for Problem $(\mathcal{M}_{N,\mu_N^0,\varepsilon})$ to converge, in some sense, to a minimizer of $(\mathcal{M}_{\mu^0})$ as $N \to \infty$ and $\varepsilon \to 0$. This is the case, but only provided ε is related to N in such a way that it does not vanish too quickly as $N \to \infty$. The kind of convergence we get is very much in the spirit of Γ-convergence and is stated in Proposition 17 below. Note however that the result stated in ***(Upper bound)*** is weaker than the usual Γ-limsup one.

The proof of this proposition uses a quantization argument for a solution $Q_{\min}$ of $(\mathcal{M}_{\mu^0})$ (see Chapter 3, (3.1)), using measures supported on a Sobolev space $\mathbb{H}^s$. From standard Sobolev inclusions, we can find $\frac{1}{2} < s \leq 1$ such that

$$\mathbb{W}^{1,r}([0;T],\mathbb{R}^d) \hookrightarrow \mathbb{H}^s([0;T],\mathbb{R}^d) \hookrightarrow \mathcal{C}([0;T],\mathbb{R}^d)$$

These injections are compact (recall that r is the exponent in the definition of the Lagrangian L). From now on, we will denote $\mathbb{H}^s([0;T],\mathbb{R}^d)$ by $\mathbb{H}^s$ and the 2-Wasserstein distance associated with the Sobolev norm on $\mathcal{P}_2(\mathbb{H}^s)$ by $W_{\mathbb{H}^s}$. In particular, $Q_{\min}$ is supported on $\mathbb{H}^s$ and we will take our quantization measures supported in this same space.

The reason behind this choice is the following: to prove the (Upper bound) part of proposition 17 below, we need to approximate $Q_{\min}$ using discrete probabilities which have lower kinetic energy. Although the approximation can be done by quantization measures according to most Wasserstein distances, taking one associated with a Hilbert norm (on the $\mathbb{H}^s$ Sobolev space), yields quantization measures supported on suitable barycenters, in some sense. L being a convex function, this gives us measures with a lower kinetic energy than $Q_{\min}$, which will be useful in the proof of the (Upper bound) claim below:

Proposition 17. *Let $(\varepsilon_N)_N$ be a positive sequence vanishing at infinity and assume that μ_N^0 narrowly converges towards μ^0 in $\mathcal{P}(\mathbb{R}^d)$.*

- ***(Lower bound)*** *Let $(Q_N)_N$ narrowly converge to Q_∞ in $\mathcal{P}(\Gamma)$. Then, we have*

$$J(Q_\infty) \leq \liminf_{N\to\infty} J_{\varepsilon_N}(Q_N).$$

For $N \in \mathbb{N}$, let

$$\tau_N := \inf\left\{ W^2_{\mathbb{H}^s}(\tilde{Q}_N, Q_{\min}) \;\middle|\; \tilde{Q}_N \in \mathcal{P}_N(\mathbb{H}^s) \right\} \tag{2.10}$$

be the optimal N-point quantization error for $Q_{\min}$ in $\mathcal{P}(\mathbb{H}^s)$.

- ***(Upper bound)*** *Assume that $\tau_N = o_{N\to\infty}(\varepsilon_N)$ and $W_2^2(\mu_N^0, \mu^0) = o_{N\to\infty}(\varepsilon_N)$. Then for any sequence $(Q_N)_N$ where Q_N is a minimizer, respectively for $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$,*

$$\limsup_{N\to\infty} J_{\varepsilon_N}(Q_N) \leq J(Q_{\min}).$$

The proof of ***(Lower bound)*** is a straightforward proof of Γ-liminf inequality:

Proof of proposition 17 (Lower bound). Take Q_N and Q_∞ as in the proposition. For every $t \in [0;T]$ and every N, define ρ_N^t as a minimizer in the problem defining $F_{\varepsilon_N}(e_t\#Q_N)$.

One can assume that $J_{\varepsilon_N}(Q_N)$ is bounded from above. Therefore, there exists $C > 0$ such that $\int_0^T W_2^2(e_t\#Q_N, \rho_N^t)dt \leq C\varepsilon_N$ for every N, since F and G are also bounded from below. Up to extracting a subsequence, we can assume that for almost all $t \in [0;T]$, ρ_N^t narrowly converges, as N goes to infinity, towards $e_t\#Q_\infty$. Using Fatou lemma, we get

$$\int_0^T F(e_t\#Q_\infty)dt \leq \int_0^T \liminf_{N\to\infty} F_{\varepsilon_N}(e_t\#Q_N)dt \leq \liminf_{N\to\infty} \int_0^T F_{\varepsilon_N}(e_t\#Q_N)dt$$

The lagrangian part (as well as G of course) is l.s.c for the narrow convergence, and we can write:

$$\begin{aligned}\int_\Gamma L(\gamma')dQ_\infty(\gamma) + \int_0^T F(e_t\#Q_\infty)dt + G(Q_\infty) \leq & \liminf_{N\to\infty} \int_\Gamma L(\gamma')dQ_N(\gamma) \\ & + \liminf_{N\to\infty} \int_0^T F_{\varepsilon_N}(e_t\#Q_N)dt \\ & + \liminf_{N\to\infty} G(Q_N) \\ \leq & \liminf_{N\to\infty} J_{\varepsilon_N}(Q_N)\end{aligned}$$

which is our claim. □

To prove the ***(Upper bound)***) inequality, we first need a lemma on the minimizers of the problem defining τ_N:

Lemma 2. *For every $N \in \mathbb{N}^*$, there exists $\tilde{Q}_N \in \mathcal{P}(\mathbb{H}^s)$ optimal for* (2.10)*, $\tau_N = W^2_{\mathbb{H}^s}(\tilde{Q}_N, Q_{min})$ and,*

$$\int_\Gamma L(\gamma')d\tilde{Q}_N(\gamma) \leq \int_\Gamma L(\gamma')dQ_{\min}(\gamma).$$

Furthermore, $\tau_N \to 0$ as N goes to infinity and in particular, $\tilde{Q}_N$ narrowly converges towards $Q_{\min}$ in $\mathcal{P}(\Gamma)$, as $N \to \infty$.

Proof. $W^2_{\mathbb{H}^s}(., Q_{\min})$ is l.s.c for the narrow convergence on $\mathcal{P}_N(\Gamma)$, from the lower semi-continuity of the $\mathbb{H}^s$ norm with respect to the uniform norm on Γ. Take a minimizing sequence $(Q_n)_n$ for our problem. We can choose Q_n to have lower kinetic energy than $Q_{\min}$:

To see this, fix $n \in \mathbb{N}$, and set $Q_n = \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{\gamma}^i}$ and $P = \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{\gamma}^i} \times Q^i_{\min}$ an optimal transport plan from Q_n to $Q_{\min}$ (in particular, $Q^i_{\min} \in \mathcal{P}(\Gamma)$ for every i).

We construct a competitor to Q_n for the infimum problem (2.10), supported on the barycenters of the measures $Q^i_{\min}$ (which play the role of the Laguerre cells from semi-discrete optimal transport, Definition 6):

For $i = 1 \dots N$, set

$$\eta^i = \int_\Gamma \gamma dQ^i_{\min}(\gamma).$$

Each η^i is a minimizer of the convex functional $\int_\Gamma ||. - \gamma||^2_{\mathbb{H}^s} dQ^i_{\min}(\gamma)$ over $\mathbb{H}^s$. Indeed, this functional is differentiable on $\mathbb{H}^s$, with gradient

$$2\int_\Gamma (. - \gamma)dQ^i_{\min}(\gamma)$$

which vanishes at η^i. Therefore,

$$\begin{aligned} W^2_{\mathbb{H}^s}\left(\frac{1}{N}\sum_{i=1}^N \delta_{\eta^i}, Q_{\min}\right) &\leq \frac{1}{N}\sum_{i=1}^N \int_\Gamma ||\eta^i - \gamma||^2_{\mathbb{H}^s} dQ^i_{\min}(\gamma) \\ &\leq \frac{1}{N}\sum_{i=1}^N \int_\Gamma ||\tilde{\gamma}^i - \gamma||^2_{\mathbb{H}^s} dQ^i_{\min}(\gamma) = W^2_{\mathbb{H}^s}(Q_n, Q_{\min}) \end{aligned}$$

and we can assume that Q_n is supported on the barycenters η_i. But, then, by convexity of L, Q_n has lower kinetic energy than $Q_{\min}$: $\int_\Gamma L(\gamma')dQ_n(\gamma) \leq \int_\Gamma L(\gamma')dQ_{\min}(\gamma)$. Similarly to proposition 15, we can conclude that $(Q_n)_n$ is tight and, up to a subsequence, it narrowly converges towards a minimizer $\tilde{Q}_N$ of $W^2_{\mathbb{H}^s}(., Q_{\min})$ over $\mathcal{P}_N(\Gamma)$, which verifies

$$\int_\Gamma L(\gamma')d\tilde{Q}_N(\gamma) \leq \int_\Gamma L(\gamma')dQ_{\min}(\gamma)$$

To show that τ_N vanishes at infinity, it is sufficient to show that there exists $(Q_N)_N$, such that, for every N, $Q_N \in \mathcal{P}_N(\mathbb{H}^s)$, Q_N narrowly converges towards $Q_{\min}$ in $\mathcal{P}(\mathbb{H}^s)$, and $\int_\Gamma ||\gamma||^2_{\mathbb{H}^s} dQ_N(\gamma)$ converges towards $\int_\Gamma ||\gamma||^2_{\mathbb{H}^s} dQ_{\min}(\gamma)$, as N goes to infinity. This can be done, for instance, as in Theorem 2.13 of [9], by sampling trajectories in the support of $Q_{\min}$ and using a law of large numbers. Finally, since $s > \frac{1}{2}$, $\mathbb{H}^s$ is continuously injected in Γ, and we have the narrow convergence in $\mathcal{P}(\Gamma)$ (for the uniform norm, this time). □

We will also need the following lemma to compare the "global" Wasserstein distance $W_{\mathbb{H}^s}$ to the local one at time t:

Lemma 3. *Let $Q, Q' \in \mathcal{P}(\mathbb{H}^s)$. Then, there exists $C > 0$, such that for every $t \in [0;T]$,*

$$W_2(e_t\#Q, e_t\#Q') \leq C W_{\mathbb{H}^s}(Q, Q')$$

Proof. Since $s > \frac{1}{2}$, Sobolev injections give the existence of $C > 0$ such that $||.||_\infty \leq C||.||_{\mathbb{H}^s}$ on $\mathbb{H}^s$. Take P, an optimal transport plan from Q to Q' with the cost given by $||.||^2_{\mathbb{H}^s}$, and $t \in [0;T]$. Then, $(e_t, e_t)\#P$ is a transport plan from $e_t\#Q$ to $e_t\#Q'$, and we can write:

$$\begin{aligned} W_2^2(e_t\#Q, e_t\#Q') &\leq \int_{\Gamma\times\Gamma} ||\gamma^1(t) - \gamma^2(t)||^2 dP(\gamma^1, \gamma^2) \\ &\leq C^2 \int_{\Gamma\times\Gamma} ||\gamma^1 - \gamma^2||^2_{\mathbb{H}^s} dP(\gamma^1, \gamma^2). \end{aligned}$$ □

Proof of proposition 17 (Upper bound): Set $N \in \mathbb{N}^*$ and $\tilde{Q}_N$ and τ_N as in Lemma 2. As is, $\tilde{Q}_N$ is not necessarily admissible since it may not satisfy $e_0\#\tilde{Q}_N = \mu_N^0$. However, since they are discrete measures with the same amount of Diracs and the same masses, we can simply translate the curves in $spt(\tilde{Q}_N)$ in order for it to be admissible for $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, using vectors that are constant in time. This new measure, which we denote by $\tilde{Q}_{\mu_N^0,N}$ is admissible for $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, has the same kinetic energy as $\tilde{Q}_N$ and satisfies $W^2_{\mathbb{H}^s}(\tilde{Q}_{\mu_N^0,N}, \tilde{Q}_N) = W_2^2(\mu_N^0, e_0\#\tilde{Q}_N)$.

Now, if Q_N is a minimizer for $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, then

$$\begin{aligned} J_{\varepsilon_N}(Q_N) \leq J_{\varepsilon_N}(\tilde{Q}_{\mu_N^0,N}) &\leq \int_\Gamma L(\gamma')d\tilde{Q}_N(\gamma) + G(\tilde{Q}_{\mu_N^0,N}) \\ &\quad + \int_0^T \frac{W_2^2(e_t\#\tilde{Q}_{\mu_N^0,N}, e_t\#Q_{\min})}{2\varepsilon_N} + F(e_t\#Q_{\min})dt \\ &\leq \int_\Gamma L(\gamma')dQ_{\min}(\gamma) + G(\tilde{Q}_{\mu_N^0,N}) + T\frac{W_2^2(e_0\#\tilde{Q}_N, \mu_N^0)}{\varepsilon_N} \\ &\quad + \int_0^T \frac{W_2^2(e_t\#\tilde{Q}_N, e_t\#Q_{\min})}{\varepsilon_N} + F(e_t\#Q_{\min})dt \end{aligned} \tag{2.11}$$

Lemma 3 gives us the bound:

$$\int_0^T W_2^2\left(e_t\#\tilde{Q}_N, e_t\#Q_{\min}\right)dt \lesssim W_{\mathbb{H}^s}^2\left(\tilde{Q}_N, Q_{\min}\right) = \tau_N.$$

And, by convexity of the transport cost,

$$W_2^2(e_0\#\tilde{Q}_N, \mu_N^0) \leq 2\left(W_2^2(e_0\#\tilde{Q}_N, \mu^0) + W_2^2(\mu^0, \mu_N^0)\right) \lesssim \left(\tau_N + W_2^2(\mu^0, \mu_N^0)\right),$$

and $\tilde{Q}_{\mu_N^0,N}$ narrowly converges to $Q_{\min}$ in $\mathcal{P}(\Gamma)$. If we chose $(\varepsilon_N)_{N\in\mathbb{N}}$ and $(\mu_N^0)_{N\in\mathbb{N}}$ such that τ_N and $W_2^2(\mu^0, \mu_N^0)$ are negligible compared to ε_N, as $N \to \infty$, then taking the limsup in inequalities (2.11) gives us

$$\limsup_{N\to\infty} J_{\varepsilon_N}(Q_N) \leq J(Q_{\min}). \quad \square$$

Corollary 17.1. *With the same notations and assumptions on $(\varepsilon_N)_{N\in\mathbb{N}}$ and $(\mu_N^0)_{N\in\mathbb{N}}$ as in proposition 17 **(Upper bound)**, Q_N narrowly converges, up to a subsequence, towards a minimizer of J. In particular, if $(\mathcal{M}_{\mu^0})$ has a unique minimizer $Q_{\min}$, then any such sequence $(Q_N)_{N\in\mathbb{N}}$ narrowly converges toward $Q_{\min}$ (this time without extraction).*

Proof. Similarly to Γ-convergence, this is a direct consequence of propostion 17. By the ***(Upper bound)*** property, up to a subsequence, $J_{\varepsilon_N}(Q_N)$ converges towards a limit $l \leq \min J$. Then as before, $\int_\Gamma L dQ_N$ is bounded in N and $(\mu_N^0)_N$ is tight in $\mathcal{P}(\mathbb{R}^d)$, therefore, $(Q_N)_N$ is tight, in $\mathcal{P}(\Gamma)$. Let us extract from it a subsequence converging towards $Q_\infty \in \mathcal{P}(\Gamma)$. Then $e_0\#Q_\infty = \mu^0$, and by the ***(Lower bound)*** property, $J(Q_\infty) \leq l \leq \min J$ hence, Q_∞ is a minimizer of $(\mathcal{M}_{\mu^0})$. $\square$

The proper sequence ε_N of parameters (or rather their precise behaviour as $N \to \infty$), remains beyond our reach even in the simpler convex situation presented in section 2.4. However, one can obtain an upper bound on the vanishing rate of τ_N by bounding the box-dimension, in $\mathbb{H}^s$ of the support of the minimizing measure $Q_{\min}$. This bound is given by the following correspondence between optimal quantization and optimal covering of a set. We refer the reader to [50] proposition 4.2 for a proof, as well as [29] for more details on the subject of vector quantization:

Proposition 18. *For a metric space (X, d_X), take $Q \in \mathcal{P}(X)$ supported on $\Sigma \subset X$. Define the optimal quantization error of Q,*

$$\tau_N = \min \left\{ W^2_{d_X,2}(Q, \tilde{Q}) \;\middle|\; \tilde{Q} \in \mathcal{P}_N(X) \right\}$$

as in proposition 17, and the optimal covering radius of Σ as the quantity (where, d_H is the Hausdorf distance between subsets of X),

$$r_N = \inf \left\{ d_H(\Sigma, P) \mid P \subset X,\ |P| \leq N \right\}.$$

Then, assuming $r_N = O_{N\to\infty}\left(N^{-\frac{1}{D}}\right)$ one gets:

$$\tau_N = \begin{cases} O_{N\to\infty}\left(N^{-1}\right) & \text{if } D < 2 \\ O_{N\to\infty}\left(N^{-1} \ln N\right) & \text{if } D = 2 \\ O_{N\to\infty}\left(N^{-\frac{2}{D}}\right) & \text{if } D > 2 \end{cases} \tag{2.12}$$

The constant D in this proposition is called the box-dimension or Minkowski dimension of the set Σ.

An initial point to make is that Proposition 18 with $X = \mathbb{R}^d$ guarantees that we can choose μ_N^0 in such a way that, for instance for $d \geq 3$, $W_2^2(\mu_N^0, \mu^0)$ be $O_{N\to\infty}(N^{-2/d})$. This is an information to take into account when choosing ε_N (although it is likely to be redundant with the one given by the growth of τ_N, defined in proposition 17).

Let us give one positive result, when the velocity field associated with the measure $Q_{\min}$ is Lipshitz-continuous (a similar conclusion was reached by Mérigot and Mirebeau in [50]):

Proposition 19. *Define $Q_{\min}$ as in Section 2.1 and take (μ, v) solution to the continuity equation, induced by $Q_{\min}$ according to Theorem 14. We assume that v is Lipshitz-continuous with respect to x, uniformly in t.*

Then, $\mathrm{spt}(Q_{\min})$ is of box-dimension at most d, where $\Omega \subset \mathbb{R}^d$.

Proof. With these notations, Theorem 14 tells us that

$$\mathrm{spt}(Q_{\min}) \subset A := \{\gamma \in \mathbb{W}^{1,r}([0;T],\mathbb{R}^d) \mid \text{for a.e. } t \in [0;T],\ \gamma(t) \in \Omega \text{ and } \gamma'(t) = v(t,\gamma(t))\} \quad (2.13)$$

In particular, $\mathrm{spt}(Q_{\min})$ is bounded in $\mathbb{W}^{1,r}([0;T],\mathbb{R}^d)$ and therefore in the larger space $\mathbb{H}^s([0;T],\mathbb{R}^d)$, since Ω is compact. By Cauchy-Lipshitz theorem (v is Lipshitz-continuous with respect to x, uniformly in t), the function $e_0 : \gamma \mapsto \gamma(0)$ is bi-lipschitz continuous on A, for the $\mathbb{W}^{1,r}$ norm and therefore, for the $\mathbb{H}^s$ norm. This exactly means that a covering of $\mathrm{spt}(Q_{\min})$ by balls corresponds to a covering of the corresponding subset of Ω by as many balls and conversely, therefore,

$$r_N(\mathrm{spt}(Q_{\min})) = O_{N\to\infty}(N^{-\frac{1}{d}}) \quad \square$$

This proposition would allow us to chose ε_N dominating the corresponding bound for τ_N according to (2.12), depending on the dimension of the domain Ω. However, it is worthwhile to mention that such regularity is much higher than the one recently obtained for solutions of congested mean field games (and in particular variational ones). On this topic, let us cite a recent result by Santambrogio and Lavenant [39], in the already restrictive case of Ω convex, a quadratic Lagrangian $L = \frac{\|.\|^2}{2}$, a potential term G given by the integral of $\mathbb{H}^1$ potentials (V and Φ in the introductory example) and a strong congestion enforced by $F = \chi_{\rho\leq \mathrm{d}x}$ (see the numerical Section 2.4, case of a convex domain, (2.17)). Under these assumptions, one can claim Hölder continuity of the value function ϕ appearing in the mean field system (2.5), which then only translate, through $v = -\nabla_x\phi$ into slightly better than $\mathbb{L}^2$ integrability (w.r.t. t and x) for the velocity field v.

We finish this section by mentioning a stronger convergence result, in the cases where F has the integral form 1.14 with a function f strongly convex on $\mathbb{R}$. We recall the notation for the Moreau-Yosida projections of Q_N at various times, introduced in the proof of Proposition 17, ***(Lower bound)***:

$$\rho_N : t \in [0;T] \mapsto \mathrm{argmin}_\rho \frac{W_2^2(\rho, e_t\#Q_N)}{2\varepsilon_N} + \int_\Omega f(\rho(x))dx$$

associated with a sequence $(Q_N)_N$ of minimizers for problem $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, in this section. Let us assume that problem $(\mathcal{M}_{\mu^0})$ has a unique solution, $Q_{\min}$. We take the appropriate values for the parameters ε_N and μ_N^0 such that (up to a subsequence), Q_N narrowly converges to $Q_{\min}$ and $J_{\varepsilon_N}(Q_N)$ converges to $J(Q_{\min})$ (see 17).

Lemma 4. *With these notations,*

$$\lim_{N\to\infty}\int_0^T\int_\Omega f(\rho_N(t)(x))dx = \int_0^T\int_\Omega f(e_t\#Q_{\min}(x))dx$$

Proof. From lower semi-continuity of F, we already have,

$$\int_{[0;T]\times\Omega} f(e_t\#Q_{\min}(x))dxdt \leq \liminf_{N\to\infty}\int_{[0;T]\times\Omega} f(\rho_N(t)(x))dxdt$$

But, on the other hand,

$$\begin{aligned}\int_{[0;T]\times\Omega} f(\rho_N(t)(x))dxdt \leq &\int_{[0;T]\times\Gamma} L(\gamma')d(Q_{\min}-Q_N)(\gamma)\\ &+\int_{[0;T]\times\Omega} f(e_t\#Q_{\min}(x))dxdt\\ &+G(Q_{\min})-G(Q_N)+o_{N\to\infty}(1)\end{aligned}$$

and taking the limsup as $N\to\infty$, we obtain

$$\limsup_{N\to\infty}\int_{[0;T]\times\Omega} f(\rho_N(t)(x))dxdt \leq \int_{[0;T]\times\Omega} f(e_t\#Q_{\min}(x))dxdt$$

and our lemma. □

Strong convergence immediately follows at least in two cases:

Proposition 20. *Under the assumptions on ε_N, μ_N^0, Q_N and $Q_{\min}$ listed above,*

- *If f **is strongly convex**, then ρ_N strongly converges in $\mathbb{L}^2([0;T]\times\Omega)$ (as a function of t and x), towards $\rho_{\min}:(t,x)\in[0;T]\times\Omega\mapsto e_t\#Q_{\min}(x)$.*
- *If $f:\rho\mapsto|\rho|^m$, $m\geq 2$ **is a power**, then a similar strong convergence is true, this time in $\mathbb{L}^m([0;T]\times\Omega)$.*

Proof. Let us first handle the case where f is a power. In this case, lemma 4 guarantees that $\|\rho_N\|_{\mathbb{L}^m}$ converges to $\|\rho_{\min}\|_{\mathbb{L}^m}$. Since ρ_N already narrowly converges towards $\rho_{\min}$, this convergence is also a weak convergence in duality with $\mathbb{L}^{m'}$ where $m'=\frac{m}{m-1}$. But, from the convergence of the norms, this implies strong convergence in $\mathbb{L}^m$ using the Radon-Riesz property.

If f is strongly convex, one can claim for any $N \in \mathbb{N}^*$ and almost any $t \in [0;T]$ and $x \in \Omega$,

$$\frac{1}{8}\left\|\rho_N(x) - \rho_{\min}(x)\right\|^2 \leq \frac{1}{2}f(\rho_N(t)(x)) + \frac{1}{2}f(\rho_{\min}(t)(x)) - f\left(\frac{1}{2}(\rho_N(t)(x) + \rho_{\min}(t)(x))\right).$$

Integrating the right-hand side in t and x, and taking the inf-limit as $N \to \infty$, one would obtain a negative value, from Lemma 4 and the lower semi-continuity of F (remember that ρ_N narrowly converges to $\rho_{\min}$ from our previous ***Lower bound*** properties). Looking at the integral of the left-hand side, this exactly states the strong-$\mathbb{L}^2$ convergence that we claimed. □

More generally, if there exists functions j and j^* on $\mathbb{R}$ and a constant $C > 0$ such that for p and ρ in $\mathbb{R}$,

$$f(\rho) + f^*(p) \geq p \cdot \rho + C|j(\rho) - j^*(p)|^2,$$

we get strong convergence of the functionals $j(\rho_N)$ towards $j(\rho_{\min})$ in $\mathbb{L}^2$, provided some invertibility on j, which implies strict convexity for f. This is a common assumption to show regularity results via duality on the solutions of a convex problem (see for instance [58]). This convergence could bring no information, and $j = j^* \equiv 0$ are actually always suitable for any convex function f. However, in the case of f strongly convex, $j(\rho) = \rho$ with $j^*(p) = (f^*)'(p)$ are suitable and we recover the first case of proposition 20. Similarly, $j(\rho) = \rho.|\rho|^{m/2-1}$ and $j^*(p) = p.|p|^{m'/2-1}$ are suitable in the situation $f \equiv |.|^m$. Again, we recover the $\mathbb{L}^m$ convergence claimed above. We highlighted the two cases of Proposition 20, where we recover the values (as a density) of $e_t\#Q_{\min}$ in an almost-everywhere way.

2.3 The fully discrete problem

We now use a uniform time discretization $0,\ \delta, ...,\ M\delta = T$ to compute a fully discretized version, with respect to space and time, of problem $(\mathcal{M}_{\mu^0})$. Rather than writing heavy formulae for a new global energy, we will change the subset of $\mathcal{P}(\Gamma)$ upon which the minimization is done, allowing for an energy almost identical to J_ε. The Lagrangian and potential parts will remain the same as in J and J_ε, whereas the congestion term will be approximated by a Riemann sum. This is mainly done in order to simplify computations, and any time-discretization of curves in Γ which allows $\mathbb{W}^{1,r}$ bounds of the sort of (2.14) and (2.15) should also work here.

We perform our optimization on the space of functions in Γ which are affine on each interval $[i\delta, (i+1)\delta]$ with $i = 0 \dots M-1$, denoted $\Gamma_\delta^{\text{lin}}$. Our fully discrete problem is then:

$$(\mathcal{M}_{N,\mu_N^0,\delta,\varepsilon}) : \inf \left\{ J_{\delta,\varepsilon}(Q) \;\middle|\; Q \in \mathcal{P}_N(\Gamma_\delta^{\text{lin}}) \text{ s.t. } e_0 \# Q = \mu_N^0 \right\}$$
$$\text{with } J_{\delta,\varepsilon}(Q) =:= \int_\Gamma L(\gamma') dQ(\gamma) + \delta \sum_{i=1}^{M-1} F_\varepsilon \left(e_{i\delta} \# Q \right) + G(Q)$$

Similarly to $(\mathcal{M}_{N,\mu_N^0,\varepsilon})$ and $(\mathcal{M}_{\mu^0})$, we have existence of minimizers for any value of the parameters, and we omit the demonstration as it would be almost identical:

Proposition 21. *For every $N \in \mathbb{N}^*$, $\delta,\ \varepsilon > 0$, $J_{\delta,\varepsilon}$ is l.s.c for the narrow convergence and for every $\mu_N^0 \in \mathcal{P}_N(\Omega)$, the infimum in $(\mathcal{M}_{N,\mu_N^0,\delta,\varepsilon})$ is attained.*

What is more interesting is a similar convergence result to the one in proposition 17, with an additional constraint on the parameters of the time discretizations, δ_N:

Proposition 22. *Assume that $(\delta_N)_N$, $(\varepsilon_N)_N$ converge to 0, and that μ_N^0 narrowly converges towards μ^0 in $\mathcal{P}(\mathbb{R}^d)$ as $N \to \infty$:*

- ***(Lower bound)** Let $(Q_N)_N$ narrowly converge to Q_∞ in $\mathcal{P}(\Gamma)$. We then have*

$$J(Q_\infty) \leq \liminf_{N\to\infty} J_{\delta_N,\varepsilon_N}(Q_N).$$

- ***(Upper bound)** Under the same assumptions as in proposition 17, and also assuming that $(\delta_N)^{2/r'} = o(\varepsilon_N)$ where $r' = \frac{r}{r-1}$ is the dual exponent for r, introduced in* (2.8)*. Then for every sequence $(Q_N)_N$, with Q_N a minimizer respectively for $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$, we have*

$$\limsup_{N\to\infty} J_{\delta_N,\varepsilon_N}(Q_N) \leq J(Q_{\min})$$

.

Proof of Proposition 22 (Lower bound): We can assume that $J_{\delta_N,\varepsilon_N}(Q_N)$ is bounded from above uniformly in N. In particular, Q_N is supported in $\mathbb{W}^{1,r}$ for every N. Then, as before,

$$\int_\Gamma L(\gamma') dQ_\infty(\gamma) \leq \liminf_{N\to\infty} \int_\Gamma L(\gamma') dQ_N(\gamma).$$

For every $\gamma \in \mathbb{W}^{1,r}$, and $i\delta_N \leq t \leq (i+1)\delta_N$,

$$||\gamma(t) - \gamma(i\delta_N)||^2 \leq \delta_N^{2/r'} \left(\int_0^T ||\gamma'(u)||^r du \right)^{2/r} \tag{2.14}$$

and, integrating this inequality along Q_N, we get $W_2^2(e_{i\delta_N}\#Q_N, e_t\#Q_N) \leq C\delta_N^{2/r'}$ for every t in $]i\delta_N; (i+1)\delta_N]$, since $\int_\Gamma L(\gamma')dQ_N(\gamma)$ is bounded. In particular, for every t, $e_{\lfloor t/\delta_N \rfloor \delta_N}\#Q_N$ narrowly converges towards $e_t\#Q_\infty$. Then, by Fatou lemma,

$$\begin{aligned}\int_0^T F(e_t\#Q_\infty)dt &\leq \liminf_{N\to\infty} \sum_{i=0}^{M_N-1} \int_{i\delta_N}^{(i+1)\delta_N} F_{\varepsilon_N}(e_{i\delta_N}\#Q_N)dt \\ &\leq \liminf_{N\to\infty} \delta_N \sum_{i=1}^{M_N-1} F_{\varepsilon_N}(e_{i\delta_N}\#Q_N)\end{aligned}$$

and that last term is exactly the congestion term in $J_{\delta_N,\varepsilon_N}$. Finally, continuity of G gives us our ***(Lower bound)*** inequality. □

Proof of Proposition 22 (Upper bound). We momentarily fix $N \in \mathbb{N}^*$. Take $\tilde{Q}_N$ and $\tilde{Q}_{\mu_N^0,N}$ as in lemma 2 and the proof of proposition 17 and define the piecewise affine interpolation operator, $T_{\delta_N}^{\text{lin}} : \gamma \in \Gamma \mapsto \gamma_{\delta_N}^{\text{lin}}$ where, for t in $[i\delta_N; (i+1)\delta_N]$, and $\gamma \in \Gamma$, $\gamma_{\delta_N}^{\text{lin}}(t) = \gamma(i\delta_N) + \frac{\gamma((i+1)\delta_N)-\gamma(i\delta_N)}{\delta_N}(t - i\delta_N)$. The measure $\tilde{Q}_{N,\delta_N}^{\text{lin}} = T_{\delta_N}^{\text{lin}}\#\tilde{Q}_{\mu_N^0,N}$ will take the role of competitor for the problem $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$, role that $\tilde{Q}_{\mu_N^0,N}$ had for problem $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$.

Then, convexity of L gives us for every N, the inequalities

$$\int_\Gamma L(\gamma')d\tilde{Q}_{N,\delta_N}^{\text{lin}}(\gamma) \leq \int_\Gamma L(\gamma')d\tilde{Q}_N(\gamma) \leq \int_\Gamma L(\gamma')dQ_{\min}(\gamma).$$

For F_{ε_N}, we have, as previously,

$$\begin{aligned}\delta_N \sum_{i=1}^{N} F_{\varepsilon_N}(e_{i\delta_N}\#\tilde{Q}^{\mathrm{lin}}_{N,\delta_N})dt \leq & \sum_{i=1}^{N}\int_{i\delta_N}^{(i+1)\delta_N} \frac{W_2^2(e_{i\delta_N}\#\tilde{Q}_{\mu_N^0,N}, e_t\#Q_{\min})}{2\varepsilon_N} \\ & + F(e_t\#Q_{\min})dt \\ \leq & \sum_{i=1}^{N}\int_{i\delta_N}^{(i+1)\delta_N} \frac{W_2^2(e_{i\delta_N}\#\tilde{Q}_N, e_t\#\tilde{Q}_N)}{\varepsilon_N} \\ & + \frac{W_2^2(e_t\#\tilde{Q}_N, e_t\#Q_{\min})}{\varepsilon_N}dt \\ & + T\frac{W_2^2(\mu_N^0, e_0\#\tilde{Q}_N)}{\varepsilon_N} + \int_0^T F(e_t\#Q_{\min})dt \\ \leq & C\left[\frac{W_2^2(\mu^0,\mu_N^0)}{\varepsilon_N} + \frac{\delta_N^{2/r'}}{\varepsilon_N} + \frac{\tau_N}{\varepsilon_N}\right] + \int_0^T F(e_t\#Q_{\min})dt\end{aligned}$$

Finally, for $\gamma \in \mathbb{W}^{1,r}$ and $i\delta_N < t \leq (i+1)\delta_N$

$$\begin{aligned}\left\|\gamma(t) - \gamma(i\delta_N) - \frac{\gamma((i+1)\delta_N) - \gamma(i\delta_N)}{\delta_N}(t - i\delta_N)\right\| \leq & \int_{i\delta_N}^{t} ||\gamma'(u)||du \\ & + \frac{t - i\delta_N}{\delta_N}\int_{i\delta_N}^{(i+1)\delta_N} ||\gamma'(u)||du \\ & \leq 2\delta_N^{1/r'}\left(\int_0^T ||\gamma'(u)||^r du\right)^{1/r}\end{aligned} \tag{2.15}$$

and, integrating along $\tilde{Q}_N$, $\lim_{N\to\infty} W^1_{\mathbb{L}^\infty}(\tilde{Q}^{\mathrm{lin}}_{N,\delta_N}, \tilde{Q}_N) = 0$, therefore, by continuity of G on Γ, $\lim_{N\to\infty} G(\tilde{Q}^{\mathrm{lin}}_{N,\delta_N}) = G(Q_{\min})$. To conclude, we observe, as earlier that

$$\begin{aligned}J_{\delta_N,\varepsilon_N}(Q_N) \leq J_{\delta_N,\varepsilon_N}(\tilde{Q}^{\mathrm{lin}}_{N,\delta_N}) \leq J(Q) + C\frac{\delta_N^{2/r'} + W_2^2(\mu^0,\mu_N^0) + \tau_N}{\varepsilon_N} \\ + G(\tilde{Q}^{\mathrm{lin}}_{N,\delta_N}) - G(Q)\end{aligned} \tag{2.16}$$

and, as soon as $(\delta_N)_N$ is taken such that $\delta_N^{2/r'} = o_{N\to\infty}(\varepsilon_N)$ along with the same growth for the other parameters as in proposition 17, one can conclude

$$\limsup_{N\to\infty} J_{\delta_N,\varepsilon_N}(Q_N) \leq J(Q) \quad \square$$

As previously, minimizers of $J_{\delta_N,\varepsilon_N}$ narrowly converge to minimizers of J, under these assumptions on ε_N, δ_N and μ_N^0.

2.4 Numerics

In this section, we approximate solutions to Problem $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$. This problem is a non-convex one, if only due to the non-convexity of the space of optimization. To obtain approximately minimizing discrete trajectories, we use a Low-memory BFGS algorithm on the energy $J_{\delta_N,\varepsilon_N}$, replacing the measures $Q \in \Gamma^{\text{lin}}_{\delta_N}$ by their expressions in term of the positions of each individual at each time (except for the positions at $t = 0$ which are fixed). As we mentioned in the previous chapter, the gradient of this scalar function in high dimension is computed using the expression in Proposition 11. There is no guarantee, *à priori*, that the trajectories we obtain via quasi-Newton algorithm are close to an actual minimum one, however, the estimates on optimal quantization from the next chapter, and especially the results on gradient flows from Lemma 7 and Lemma 9 allow us to be hopeful that these trajectories are indeed close to minimizing the fully discrete problem $(\mathcal{M}_{N,\mu_N^0,\delta_N,\varepsilon_N})$, provided gradient descent is initialized on (constant) trajectories not too close to each other, meaning with positions aligned on a regular grid.

The following trajectories have all been computed using an energy J with the same Lagrangian, given by the squared norm: $L : x \mapsto \frac{1}{2}\|x\|^2$, or equivalently, $L(\gamma') = \int_0^T \frac{\|\gamma'(t)\|^2}{2} dt$ with our abuse of notations. The congestion term F and the potential energy G have to be somewhat tailored to each domain. However, F will always enforce the "hard" congestion constraint $\rho \leq 1$ in some sense, and G always penalize the distance (euclidean for the convex domain, derived from an Eikonal equation for the non-convex one) to a set of target points. The various integrals over Laguerre cells are computed using the Pysdot[1] library for Python. This library also allows the approximation of the optimal weights in the dual formulation of Proposition 9 using a Newton algorithm on the maximized convex function,

$$G_\varepsilon : (Y, \Phi) \in (\mathbb{R}^d)^N \times \mathbb{R}^N \mapsto \sum_{i=1}^N \left[\frac{\phi_i}{N} - \int_{\text{Lag}_i(Y,\Phi)} f^* \left(\phi_i - \frac{\|x - y_i\|^2}{2\varepsilon} \right) \mathrm{d}x \right]$$

which is concave and smooth (although we did not obtain the second order regularity needed to guarantee total convergence). Solutions to the eikonal equation featured in the second example were approximated using the code available on J.-M. Mirebeau's Github[2]. Finally, the code used to obtain Fig. 2.1 and Fig. 2.2 can

[1] https://github.com/sd-ot/pysdot

[2] https://github.com/Mirebeau/HamiltonFastMarching.git, see [52], [53], [22]

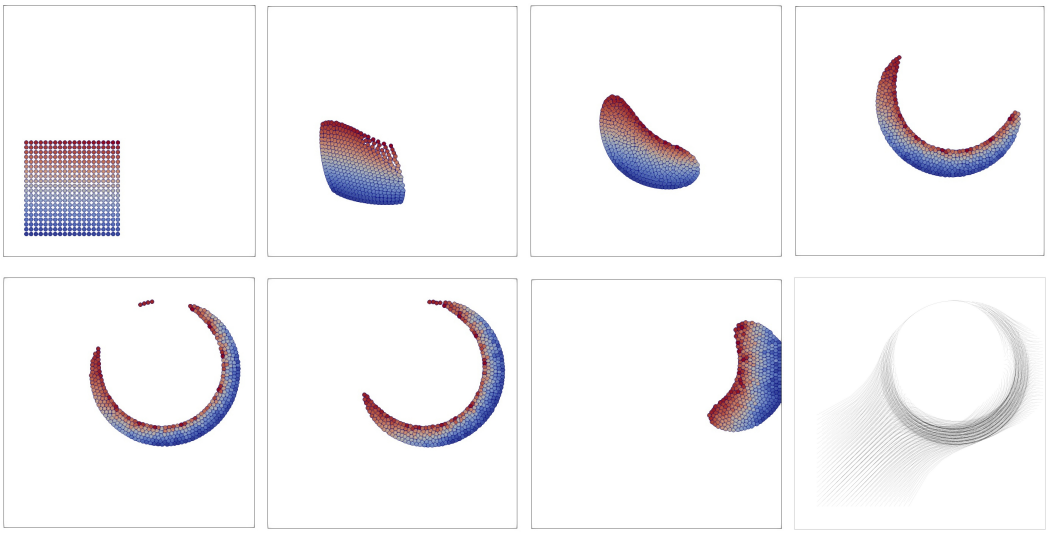


Figure 2.1: On the first 7 images (top left to bottom right), the evolution of the "charged" Laguerre cells (intersected with the support of $\rho_N(t)$) at several time steps for 400 particles in the convex domain $[-1;10]^2$. The final picture (bottom right) represents the full trajectories of the particles.

be found on the author's Github[3].

Evolution in a convex domain: In this case, our players evolve in a convex domain $\Omega = [-1;10]^2$. We use the "strong" congestion penalization,

$$F : \rho \in \mathcal{M}(\Omega) \mapsto \int_\Omega \chi_{[0;1]}(\rho(x))dx = \begin{cases} 0 & \text{if } 0 \leq \rho \leq dx \\ +\infty & \text{otherwise} \end{cases} \tag{2.17}$$

with $\chi_{[0;1]}$ being the convex indicator function of $[0;1]$. Admissible population trajectories for the continuous problem $(\mathcal{M}_{\mu^0})$ cannot have a density higher than 1 at almost any time or position.

The conclusions of, Proposition 9 apply in this case, provided $|\Omega| > 1$ to guarantee existence of dual solutions. With $f \equiv \chi_{[0;1]}$, $f^* \equiv \max(.,0)$ is the positive part function on $\mathbb{R}$ and $(f^*)' \equiv \mathbb{1}_{\mathbb{R}^+}$ (almost everywhere). For $y \in (\mathbb{R}^d)^N$, the associated optimal density in $\mathcal{F}_\varepsilon(y)$ is given on $\mathrm{Lag}_i(y,\phi)$ by the ***(Density)*** condition:

$$\rho(x) = \mathbb{1}_{\mathbb{R}^+}\left(\phi_i - \frac{\|x-y_i\|^2}{2\varepsilon}\right) = \begin{cases} 1 & \text{if } ||x-y_i||^2 \leq 2\varepsilon\phi_i \\ 0 & \text{otherwise} \end{cases}$$

[3] https://github.com/CSarrazin-prog/Congested_MFG.git

and the charged Laguerre cells (intersected with the support of ρ) are the intersection of the actual Laguerre cells, with the respective balls $B(y_i, \sqrt{2\varepsilon\phi_i})$. Although proposition 20 does not apply in this case, one can expect these cells to give a good idea of the support of the limit measure $e_t\#Q_{\min}$, and we have highlighted them on the pictures instead of the actual player's positions for this reason. The "charged" Laguerre cell, $\mathrm{Lag}_i(Y,\Phi) \cap \mathrm{spt}(\rho)$ defines a zone around y_i where non-congestion prevents other points to be. These zones correspond to a more flexible version of the "hard" balls around the particles used by Maury et al. [47] to represent the non-congestion constraint in this case.

Finally, the potential term is defined as the integral of actual potentials V and Φ:

$$G : Q \in \mathcal{P}(\Gamma) \mapsto \int_\Gamma \int_0^T V(\gamma(t))dt + \Phi(\gamma(T))dQ(\gamma)$$

with $V(x) = (\|x-(6,6)\|^2 - 9)^2$ and $\Phi(x) = \|x-(11,6)\|^2$. V penalizes the players when moving away from the circle $\mathcal{S}^1((6,6),3)$ in the course of their trajectory, while Φ gives high energy to trajectories ending too far away from the point $(11,6)$. We ran the optimization for a population of 400 players, each of mass 1/40 (for a total mass of 10, in order to have visible charged Laguerre cells), starting aligned on a regular grid on the square $[0;4] \times [0;4]$ (top left image of Fig. 2.1). These pictures where obtained for values of the discretization parameters $\varepsilon = 0.01$, $\delta = 1/64$ and $T = 15$.

In this case, the hypothesis of a Lipschitz velocity field for the continuous solution of $(\mathcal{M}_{\mu^0})$ could not hold as, in the experiment, we observe some (but not many) players around the circle from the other side, in order to avoid waiting. However this seems to be the only point of splitting for our optimal trajectories, which suggests that $spt(Q_{\min})$ should still be of dimension 2. In such a case, any sequence ε_N dominating $\ln(N)/N$ should be suitable to obtain the convergence of Proposition 22.

Evolution in a non-convex domain: This second example features a crowd motion in a non-convex domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ made of two "rooms", $\Omega_1 = [0;8]^2$ and $\Omega_3 = [11;19] \times [0;8]$ connected by a narrow corridor, $\Omega_2 = [8;11] \times [3.5;4.5]$.

Here, we had to adapt our congestion term, as the Newton algorithm, computing the optimal Laguerre cells, did not always converge for Dirac masses located too far away from Ω (which was the case for some particles not using the corridor before convergence was reached). To make this optimization easier we

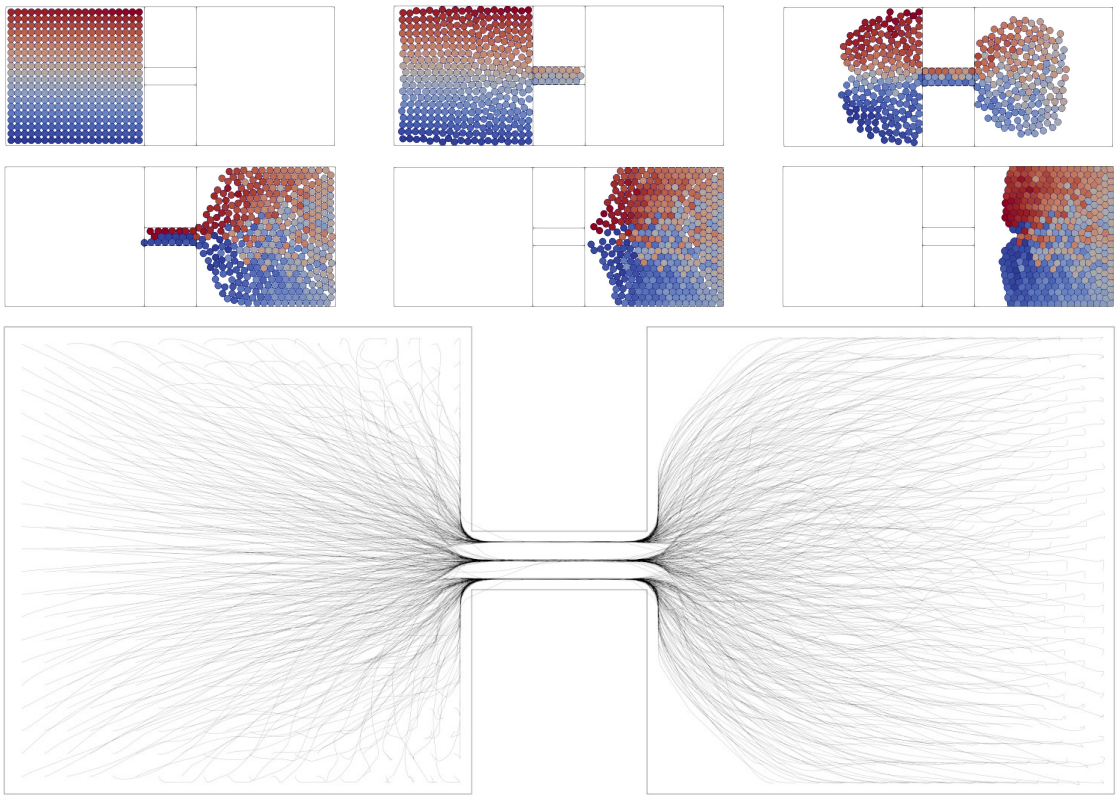


Figure 2.2: On the top six images are represented the positions and “charged" Laguerre cells of 400 particles moving in Ω. The bottom picture shows the trajectories of all the particles.

fixed a small maximum density $0 < m \ll 1$ for the area outside Ω but inside its convex envelope $\mathrm{conv}(\Omega)$, and 1 inside Ω. This results in the congestion penalization:

$$F : \rho \in \mathcal{M}(\mathrm{conv}(\Omega)) \mapsto \int_{\mathrm{conv}(\Omega)} f(x, \rho(x))dx$$

where $f : (x, \rho) \in \mathrm{conv}(\Omega) \times \mathbb{R} \mapsto \begin{cases} 0 & \text{if } 0 \leq \rho \leq 1 \text{ and } x \in \Omega \\ 0 & \text{if } 0 \leq \rho \leq m \text{ and } x \in \mathrm{conv}(\Omega) \backslash \Omega \\ +\infty & \text{otherwise} \end{cases}$

Although this isn't quite the framework of Propositions 9 to 11, these can be easily adapted to this form of congestion. The optimal ρ for positions Y of the

players is then given by:

$$\text{For a.e. } x \in \Omega,\ \rho(x) = \begin{cases} 1 & \text{if } x \in \mathrm{Lag}_i(Y,\Phi) \cap \mathcal{B}(y_i, \sqrt{2\varepsilon\phi_i}) \cap \Omega \\ m & \text{if } x \in \mathrm{Lag}_i(Y,\Phi) \cap \mathcal{B}(y_i, \sqrt{2\varepsilon\phi_i}) \cap (\mathrm{conv}(\Omega)\backslash\Omega) \\ 0 & \text{otherwise} \end{cases}$$

The support of the Moreau projection will still be an intersection of balls with the Laguerre cells, but the value of the optimal density ρ will not be 1 everywhere on this support. Instead, $\rho(x) = m$ a.e. on $\mathrm{conv}(\Omega)\backslash\Omega$, giving us larger charged Laguerre cells for the points passing near the border (or outside) of Ω. For low values of the outside density m, only very few particles can fit outside the corridor, and we should recover the strong penalization of the convex example.

Finally, due again to the non-convexity of Ω, penalizing trajectories that do not end in the right-side room using the euclidean distance did not yield satisfying results (or even convergence for the L-BFGS algorithm). We therefore define our potential term using a solution to an Eikonal equation:

$$G : Q \in \mathcal{P}(\Gamma) \mapsto \int_\Gamma \Phi(\gamma(T)) dQ(\gamma)$$

with Φ being solution of the Eikonal equation on $\mathrm{conv}(\Omega)$:

$$\begin{cases} \|\nabla\Phi(x)\| = 1 & \text{on } \Omega. \\ \|\nabla\Phi(x)\| = v & \text{outside.} \\ \Phi(18,1) = \Phi(18,7) = 0 \end{cases}$$

and v being a small value of the velocity, outside the corridor. Such a potential term penalizes heavily trajectories which end far away from the closest point between (18,1) and (18,7), while also discouraging trajectories that do, but while moving outside of Ω. Notice that, unlike the one made for F, this prescription $\|\nabla\Phi(x)\| = v$ outside Ω is dictated by the theory since our discrete trajectories could pass outside the corridor and we do not regularize G. We therefore need to chose a potential term which is continuous at least at trajectories moving though $\mathrm{conv}(\Omega)$. However, even for our value $v = 0.1$, trajectories leaving Ω were, in the end, mostly rejected by the optimization, provided the maximum time T is large enough for them to wait their turn and use the corridor. In the end, we chose to take a fairly strong congestion penalization outside the corridor, with $m = 10^{-3}$, while putting a much weaker penalization on the speed in this "forbidden zone", via the Eikonal equation with a value $v = 0.1$. This resulted in no particles crossing the borders of Ω, even though very few trajectories cut

the corners of the corridor.

To obtain the motion of Fig. 2.2, we ran the optimization for 400 particles, each of mass 1/8, starting on a regular grid over the first square Ω_1. The trajectories on these images were obtained for a slightly higher value $\varepsilon = 0.1$, as players need to be flexible enough to move through the narrow corridor. We then chose a time step of $\delta_N = 1/2^8$ and a maximum time $T = 600$ to avoid certain phenomena of "teleportation" from one room to the other. Notice that we recover a well-known phenomenon where "channels" naturally form in the corridor, in order to optimize the flow of the players in this very congested segment of the motion.

Chapter 3

Optimal quantization of measures

In this chapter, we study the simpler, but not unrelated problem of approximating a probability with density, using a discrete uniform one. The convergence estimates for one step of the Lloyd algorithm are gathered in [51]. We then present an application of these estimates, in the context of uncongested crowd motions.

3.1 The optimal quantization problem

For a measure $\rho \ll \mathrm{d}x$ in $\mathcal{P}(\mathbb{R}^d)$, we will call N-points optimal uniform quantization of ρ any solution of the minimization problem:

$$\inf \left\{ \frac{W_2^2(\mu, \rho)}{2} \mid \mu \in \mathcal{P}_N(\mathbb{R}^d) \right\} \tag{3.1}$$

This problem can also be linked to these of Wasserstein Generative Adversarial Networks (we refer the reader to Arjovsky, Chintala and Bottou [5]) and perhaps more closely Wasserstein regression (see Genevay, Peyré and Cuturi [27]). The model studied in this chapter is simpler in the sense that the discrete quantization μ can be supported on any point cloud in $\mathbb{R}^d$, whereas Wasserstein GAN and regression often construct their measure μ_θ as the push-forward of a "simple" reference measure (not ρ necessarily) through a map T_θ depending on a parameter (e.g. a computed by a neural net with weights θ). In this section, we are only interested in the non-convexity of the Lagrangian discretization and not its interaction with the generative process, hence the simpler setting.

Let us immediately state existence of solutions as well as necessary (but not sufficient, as we will see) conditions for optimality:

Proposition 23. *Let $\rho \ll \mathrm{d}x$ in $\mathcal{P}(\Omega)$. Then Problem* (3.1) *admits solutions. Furthermore, if $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$ is such a solution supported on a point cloud Y, then for*

$i = 1 \dots N$,

$$y_i = N \int_{\mathrm{Lag}_i(Y,\Phi)} x d\rho(x)$$

where $(\mathrm{Lag}_i(Y,\Phi))_{i=1\dots N}$ *are the optimal Laguerre cells for the optimal transport from* ρ *to* μ*.*

Note in particular that provided Ω is convex and $\rho > 0$ on Ω, then an optimal μ is supported on Ω and in fact, $y_i \in \mathrm{Lag}_i(Y,\Phi)$ for every i (where Φ is optimal for the transport between ρ and μ.

The proof of this proposition could easily be done using the standard method in calculus of variations and a reasoning similar to that of Lemma 2. Instead, we reformulate this problem as a finite dimensional one, in the spirit of the underlying semi-discrete optimal transport:

We define the optimal quantization error function:

$$\mathcal{F}_N : Y \in (\mathbb{R}^d)^N \mapsto \frac{W_2^2(\frac{1}{N}\sum_{i=1}^N \delta_{y_i}, \rho)}{2} \tag{3.2}$$

Then, $\mathcal{F}_N$ enjoys the following properties:

Proposition 24. *The function* $\mathcal{F}_N$ *is* $\frac{1}{N}$*–semiconcave (see Definition 18) on* $(\mathbb{R}^d)^N$ *and is of class* $\mathcal{C}^1$ *on* $(\mathbb{R}^d)^N \setminus \mathbb{D}_N$*. Here, the singular "diagonal" set* $\mathbb{D}_N = \{Y \in (\mathbb{R}^d)^N \mid \exists i \neq j y_i = y_j\}$ *of point clouds of cardinal strictly less than N.*

In addition, for any $Y \notin \mathbb{D}_N$ *one has*

$$\nabla \mathcal{F}_N(Y) = \frac{1}{N}(Y - B_N(Y)) \tag{3.3}$$

where $B_N(Y) = (b_1(Y), \dots, b_N(Y))$ *and* $b_i(Y)$ *is the* ρ*-barycenter of the* i*th optimal Laguerre cell,*

$$b_i(Y) = N \int_{\mathrm{Lag}_i(Y)} x \mathrm{d}\rho(x).$$

The proof of semiconcavity and regularity is almost identical to that of Proposition 11 and can be found, for instance, in Proposition 21 from [50]. Note also that $\mathcal{F}_N$ is coercive and l.s.c and therefore, Proposition 23 is immediate.

Remark 6. Let us note here that (3.1) is not the widely used definition of optimal quantization of measures and usually, the masses of the Dirac are also variables of the optimization. The problem is then:

$$\inf \left\{ \frac{W_2^2(\mu,\rho)}{2} \;\middle|\; \mu = \sum_{i=1}^N \alpha_i \delta_{y_i}, \alpha \in \Delta_N \right\} \tag{3.4}$$

Similar results as these of Proposition 23 are well-known for these problems, let us mention the book by Graf and Luschgy [32] as a reference for these problems. In this case, the optimal positions are barycenters of their *Voronoi cells* instead of the Laguerre ones:

$$\mathrm{Vor}_i(Y) = \left\{ x \in \mathbb{R}^d, \ \|x - y_i\|^2 \leq \|x - y_j\|^2 \ \middle| \ \text{for } j = 1, \dots, N \right\} \tag{3.5}$$

(which are Laguerre cells for a set of weights $\Phi = 0$) and the optimal masses α_i are given by the semi-discrete optimality conditions

$$\alpha_i = \rho(\mathrm{Vor}_i(Y)).$$

A reason (or at least a feature) to fix the masses in Eq. (3.1) is to obtain a distribution of points for which the local density of points is related to the sampled density ρ. Indeed, at regions where ρ becomes larger, the Dirac masses of μ correspond to very small (in the sense of their Lebesgue measure) Laguerre cells of ρ-measure $1/N$, allowing a lot more points to fit in these regions for the optimal measure. On the contrary, an optimal μ for Problem (3.4) features more similar shapes of cells, compensating for this with lower/higher values in front of the corresponding Dirac masses. In particular, it is well-known that for a minimizer of Problem (3.4),

$$\mu \in \mathrm{argmin} \left\{ W_2^2(\tilde{\mu}, \rho) \mid \mathrm{Card}(\mathrm{spt}(\tilde{\mu}) \leq N) \right\},$$

if one defines the uniform measure on the suport of μ:

$$\hat{\mu} = \frac{1}{\mathrm{Card}(\mu)} \sum_{y \in \mathrm{spt}(\mu)} \delta_y,$$

then, $\hat{\mu}$ narrowly converges towards $\rho^{\frac{2}{2+d}}$. The concave power showcases well the uniformization of the Dirac masses' positions in $\mathrm{spt}(\mu)$.

In the remainder of this chapter, for $Y \in (\mathbb{R}^d)^N$, we denote the uniform probability measure supported on the points in Y, $\mu_Y := \frac{1}{N} \sum_{i=1}^N \delta_{y_i}$.This way, our quantization energy rewrites:

$$\mathcal{F}_N(Y) = \frac{W_2^2(\rho, \mu_Y)}{2}.$$

Other than in dimension 1, neither Problem (3.1), nor its Voronoi counterpart (3.4) are convex with respect to the positions of the Dirac masses. In particular, $\mathcal{F}_N$ could admit several minimizers and, often, several critical points which are

not even local minimizers. For instance, if $\rho \equiv 1$ on the unit square $\Omega = [0,1]^2$, one can check (Fig. 3.1) that the point cloud

$$Y_N = \left(\left(\frac{1}{2N}, \frac{1}{2} \right), \left(\frac{3}{2N}, \frac{1}{2} \right), \dots, \left(\frac{2N-1}{2N}, \frac{1}{2} \right) \right)$$

represented on the left of the figure, or the $4N$-points cross-shaped one

$$\begin{aligned} Y_N = \Bigg(& \left(\frac{1}{2} + \frac{2}{3\sqrt{N}}, \frac{1}{2} \right), \left(\frac{1}{2} + \frac{2}{3\sqrt{N}}(2\sqrt{2}-1), \frac{1}{2} \right), \\ & \dots, \left(\frac{1}{2} + \frac{2}{3\sqrt{N}}(N\sqrt{N} - (N-1)\sqrt{N-1}), \frac{1}{2} \right), \\ & \left(\frac{1}{2} - \frac{2}{3\sqrt{N}}, \frac{1}{2} \right), \left(\frac{1}{2} - \frac{2}{3\sqrt{N}}(2\sqrt{2}-1), \frac{1}{2} \right), \\ & \dots, \left(\frac{1}{2} - \frac{2}{3\sqrt{N}}(N\sqrt{N} - (N-1)\sqrt{N-1}), \frac{1}{2} \right), \\ & \left(\frac{1}{2}, \frac{1}{2} - \frac{2}{3\sqrt{N}} \right), \left(\frac{1}{2}, \frac{1}{2} - \frac{2}{3\sqrt{N}}(2\sqrt{2}-1) \right), \\ & \dots, \left(\frac{1}{2}, \frac{1}{2} - \frac{2}{3\sqrt{N}}(N\sqrt{N} - (N-1)\sqrt{N-1}) \right), \\ & \left(\frac{1}{2}, \frac{1}{2} + \frac{2}{3\sqrt{N}} \right), \left(\frac{1}{2}, \frac{1}{2} + \frac{2}{3\sqrt{N}}(2\sqrt{2}-1) \right), \\ & \dots, \left(\frac{1}{2}, \frac{1}{2} + \frac{2}{3\sqrt{N}}(N\sqrt{N} - (N-1)\sqrt{N-1}) \right) \Bigg) \end{aligned}$$

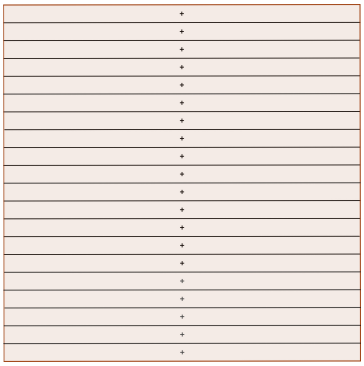
on the right are critical points of $\mathcal{F}_N$ but not minimizers. In fact, these critical points become arbitrarily bad as $N \to +\infty$ in the sense that

$$\lim_{N \to +\infty} \frac{\mathcal{F}_N(Y_N)}{\min \mathcal{F}_N} = +\infty.$$

On the other hand, actual minimizers correspond (as expected) to good approximations of ρ, in the sense that their quantization error vanishes as $N \to \infty$. This is a direct consequence of the upper bound given by the estimates (2.12) mentioned in Chapter 1:

$$\min \mathcal{F}_N = \min_{Y \in (\mathbb{R}^d)^N} \frac{1}{2} W_2^2 \left(\mu_Y, \rho \right) \lesssim \begin{cases} N^{-\frac{2}{d}} & \text{if } d > 2 \\ N^{-1} \log N & \text{if } d = 2 \\ N^{-1} & \text{if } d = 1. \end{cases} \tag{3.6}$$

Note here that a consequence of the observation in Section 3.3 below is that the constant in these bounds depends of ρ, and can very much degenerate when ρ is allowed to be a function of N (see Section 3.3).



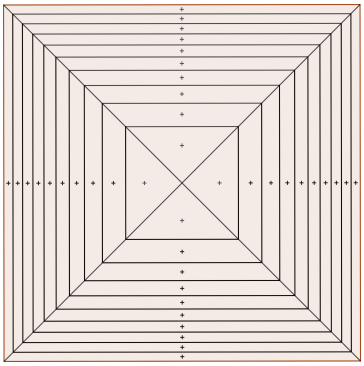


Figure 3.1: Critical points for $\mathcal{F}_{20}$ (on the left) and $\mathcal{F}_{40}$ (on the right) for a sampled density $\rho = \mathrm{d}x$. The Laguerre cells of each point are also drawn, and feature some very characteristic anisotropic dimensions.

3.2 Lloyd's algorithm for optimal uniform quantization

Considering the non-convex behavior mentioned in the previous section, it might seem-counter-intuitive that the main algorithm used to numerically approximate optimal quantization measures is a gradient descent algorithm, with large step-size. This algorithm was already recommended in Balzer et al. [7], which is commonly considered as the first use of Laguerre cells as an alternative to Voronoi ones to do quantization of measures. It consists, similarly to the algorithm used for non-uniform quantization from which it borrows its name, in sending each point of an initial point cloud to the ρ-barycenter of their respective Laguerre cells, and reiterating until some stopping criterion is attained:

Algorithm 1: Lloyd's Algorithm

Input: $N > 0$, $Y_N \in (\mathbb{R}^d)^N$
$\Phi \leftarrow (0)_N$;
$Y \leftarrow Y_N$;
$B \leftarrow Y_N$;
while *StoppingCriterion* **not** *reached* **do**
  Compute the optimal weights $\Phi \leftarrow \Phi^{\mathrm{opt}}(Y)$;
  Compute the barycenters of the Laguerre cells $B \leftarrow B_N(Y)$;
  Send each point to the corresponding barycenter $Y \leftarrow B$
Output: Y a quasi-critical point cloud

Let us come back to the gradient expression from Proposition 24 to see that

the algorithm above is also a discrete gradient descent, for the functional $N.\mathcal{F}_N$, which we summarize as the system:

$$\begin{cases} Y^0 \in (\mathbb{R}^d)^N \\ Y^{k+1} = Y^k + \tau_N(B_N(Y^k) - Y^k) \end{cases} \tag{3.7}$$

Taking $\tau_N = 1$, one indeed recovers the iterations of Lloyd's algorithm.

First, we mention some preliminary results for the possible limits of (3.7):

Proposition 25. *Let N be a fixed integer and $Y^0 \notin \mathbb{D}_N$. Then:*

- *The iterates $(Y^k)_{k\geq 0} \in (\mathbb{R}^d)^{\mathbb{N}}$ of the Lloyd algorithm are all well defined and in fact, the sequence $(Y^k)_{k\geq 0}$ belongs to a compact subset of $(\mathbb{R}^d)^N \setminus \mathbb{D}_N$.*
- *The energy $k \mapsto \mathcal{F}_N(Y^k)$ is decreasing, and $\lim_{k\to+\infty} \left\|\nabla \mathcal{F}_N(Y^k)\right\| = 0$.*

Proof. Given $Y = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N \setminus \mathbb{D}_N$, one has for any $i \in \{1, \dots, N\}$,

$$\begin{aligned} \int_{\mathrm{Lag}_i(Y)} \|x - y_i\|^2 \,\mathrm{d}\rho(x) &= \int_{\mathrm{Lag}_i(Y)} \|x - b_i(Y) + b_i(Y) - y_i\|^2 \,\mathrm{d}\rho(x) \\ &= \int_{\mathrm{Lag}_i(Y)} \|x - b_i(Y)\|^2 \,\mathrm{d}\rho(x) + \frac{1}{N} \|b_i(Y) - y_i\|^2 . \end{aligned}$$

Summing these equalities over i, we get

$$\begin{aligned} \frac{1}{N} \|B_N(Y) - Y\|^2 &= W_2^2(\rho, \mu_Y) - \sum_i \int_{\mathrm{Lag}_i(Y)} \|x - b_i(Y)\|^2 \,\mathrm{d}\rho(x) \\ &\leq W_2^2(\rho, \mu_Y) - W_2^2(\rho, \delta_{B_N(Y)}). \end{aligned}$$

Thus,

$$N \|\nabla \mathcal{F}_N(Y)\|^2 = \frac{1}{N} \|B(Y) - Y\|^2 \leq 2(\mathcal{F}_N(Y) - \mathcal{F}_N(B(Y))). \tag{3.8}$$

This implies that the values of $\mathcal{F}_N$ are decreasing during the Lloyd algorithm, as long as it can be performed.

Next up, let us show that the sequence evolves in a compact subset of $(\mathbb{R}^d)^N \setminus \mathbb{D}_N$, proving that these iterations can indeed be performed indefinitely:

Assume that $Y^k \in (\mathbb{R}^d)^N \setminus \mathbb{D}_N$. Since ρ is absolutely continuous, it is uniformly integrable which means that for every $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that for any set A with Lebesgue measure $|A| < \delta$ we have $\rho(A) < \varepsilon$. On the other hand, ρ is supported on the compact domain Ω and, as such, inside a ball $B(0, R)$. We claim that we have $|b_i(Y) - b_j(Y)| \geq r := \delta(\frac{1}{2N})/(\omega_{d-1}R^{d-1})$ and in fact, that

every barycenter $b_i(Y)$ is at distance at least $r/2$ from each face of the convex polytope $\mathrm{Lag}_i(Y)$.

Indeed, let us take $i \in \{1 \dots N\}$ and assume that one of the faces of $\mathrm{Lag}_i(Y^k)$ lies on the hyperplane $\{x_d = 0\}$ while the whole cell is located on top, inside $\{x_d \geq 0\}$. Let $s > 0$ be such that half of the mass of $\mathrm{Lag}_i(Y^k)$ is located below $x_d = s$:

$$\rho(\mathrm{Lag}_i(Y^k) \cap \{x_d \leq s\}) = \frac{1}{2N}$$

The Lebesgue measure of this "bottom half" is upper bounded by $\omega_{d-1}R^{d-1}s$ since $\Omega \subset B(0;R)$ and therefore, $s \geq r$ with the previous definition. Since half of the mass (according to ρ) of the cell $\mathrm{Lag}_i(Y)$ is above the level $x_d = s$ the x_d-coordinate of the barycenter is at least $r/2$. This shows that the ρ-barycenter lies at distance at least $r/2$ from each of its faces and therefore at distance r of the other barycenters. On the other hand, boundedness of the sequence $(Y^k)_k$ immediately follows from the fact that $\mathcal{F}_N(Y^k)$ is upper bounded by $\mathcal{F}_N(Y^0)$. Indeed, the second order moment of $\frac{1}{N}\sum_{i=1}^N \delta_{y_i^k}$ is bounded by its Wasserstein distance to ρ:

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N \left\|y_i^k\right\|^2 \leq& 2\left(\sum_{i=1}^N \int_{\mathrm{Lag}_i(Y)} \left\|y_i^k - x\right\|^2 + \|x\|^2 \,\mathrm{d}\rho(x)\right)\\ \leq& 2\left(W_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{y_i^k}, \rho\right) + \max_{x\in\Omega}\|x\|^2\right) \leq 2\mathcal{F}_N(Y^0) + C.\end{aligned}$$

As a consequence, the iterations Y^k lie in a compact subset of $(\mathbb{R}^d)^N \setminus \mathbb{D}_N$. Furthermore, since $(\mathcal{F}_N(Y_k))_{k\in\mathbb{N}}$ is bounded from below, we immediately get by applying (3.8) to Y^k and then summing in k, that $\sum_{k=1}^{+\infty} \left\|\nabla\mathcal{F}_N(Y^k)\right\|^2 < +\infty$ and in particular,

$$\lim_{k\to+\infty} \left\|\nabla\mathcal{F}_N(Y^k)\right\| = 0. \quad \square$$

For the non-uniform version of the quantization problem, convergence results for the algorithm (only to a critical point, but not up to a subsequence) can be found in [21] and [10]. These claims rely on the assumption that there is only a finite number of centroidal diagrams with the same energy (meaning, in our case, a finite number of critical point clouds for $\mathcal{F}_N$ with the same value of $\mathcal{F}_N$), which is usually a hypothesis of genericity. Under this assumption, we can also conclude positively to the convergence of the whole sequence Y^k to a critical point for $\mathcal{F}_N$, using the following proposition which implies in particular that the set of limit points is a finite connected set, i.e. a singleton:

Proposition 26. *Let N, $Y^0 \notin \mathbb{D}_N$ and $(Y^k)_{k\geq 0}$ be as in the previous Proposition 25.*

Then all limit-points of $(Y^k)_{k\geq 0}$ are critical points for $\mathcal{F}_N$ and $\mathcal{F}_N$ is constant on this limit set. Furthermore, this set of limit points is a compact connected set with empty interior.

Proof. For convenience, let us call A the set of limit points of $(Y^k)_{k\geq 0}$.

Since $\mathcal{F}_N$ is decreasing along the iterations of the Lloyd algorithm, all the limit points in A must have the same value of $\mathcal{F}_N$. Furthermore, the convergence $\left\|\nabla \mathcal{F}_N(Y^k)\right\| \to 0$ along these iterations, together with the fact that they remain in a compact set upon which $\mathcal{F}_N$ is $\mathcal{C}^1$ implies that any point of A must be a critical point.

Compactness of A is immediate from the fact that A is constituted of the limit points of a bounded sequence in $(\mathbb{R}^d)^N$ and therefore is closed and bounded, by a diagonal argument.

There is a similar result for the connectedness of such a set of limit points, but we will recall the proof:

Assume that we can take U open such that $A \subset U \cup (\mathring{U^c})$ and let us obtain a contradiction. Then, one can define an extractor ϕ by induction:

$$\begin{cases} \phi(0) = \min\{k \geq 0 \mid Y^k \in U,\ Y^{k+1} \in U^c\} \\ \phi(i+1) = \min\{k > \phi(i) \mid Y^k \in U,\ Y^{k+1} \in U^c\}. \end{cases}$$

$\phi(i)$ is well defined for $i \geq 0$, otherwise either $A \subset U$ or $A \subset (\mathring{U^c})$. Using compactness, and up to another extraction, we can assume that $Y^{\phi(i)}$ converges towards $a \in U \cap A$ whereas $Y^{\phi(i)+1}$ converges towards $b \in (\mathring{U^c}) \cap A$. But, we recall that for any i,

$$\left\|Y^{\phi(i)+1} - Y^{\phi(i)}\right\|^2 \leq 2N(\mathcal{F}_N(Y^{\phi(i)}) - \mathcal{F}_N(Y^{\phi(i)+1}))$$

and therefore, taking the limit as $i \to \infty$, we have a contradiction $a = b$, since the right-hand side goes to 0 (recall that F is constant and continuous at any point of A).

Finally, assume that there exists $a \in A$ and $r > 0$ such that $B(a, r) \subset A$. Take $a^k = Y^{\phi(k)}$ a subsequence of $(Y^k)_{k\geq 0}$ converging to a. Then there exists a rank K such that $a^k \in B(a, r)$ for any $k > K$ and after the rank $\phi(K)$, $(Y^k)_{k\geq \phi(K)}$ would be stationary, equal to a, since $\nabla \mathcal{F}_N = 0$ on $B(a, r)$. This is a contradiction to the fact that there are other limit points of A in this ball. □

We can observe on the "bad" critical points in Fig. 3.1 that the point cloud Y_N is highly concentrated, in the sense that the distance between two points in Y_N is at the lowest $\frac{1}{2N}$, whereas in an evenly distributed point cloud, one would expect the minimum distance between points to be of order $N^{-1/2}$ in this 2-dimensional setting.

In fact, this is a widely admitted starting condition, when using Lloyd algorithm (1), that the initial points be well spaced in the domain Ω. However, to our knowledge, no mathematical proof of this dependency on the initial conditions has been given in the literature so far.

From there, one could ask oneself whether or not it is easy to choose an initial Y^0 such that Lloyd algorithm ends up converging to an actual minimizer (or at least close to a good local one). It turns out that this is a highly nontrivial question (see Section 3.4 for our very partial answer). However, a very surprising experimental fact is that one step of the Lloyd algorithm is in fact enough to obtain a discrete measure very close to the density ρ (in the sense of $\mathcal{F}_N$), provided one did not start from an adversely chosen point cloud.

We now state the main result of this section, which quantifies the error on this 1-step approximation. We will use the following notation for $\varepsilon > 0$:

$$I_\varepsilon(Y) = \{i \in \{1, \dots, N\} \mid \forall j \neq i, \|y_i - y_j\| \geq \varepsilon\}$$

and

$$\mathbb{D}_{N,\varepsilon} = \{Y \in (\mathbb{R}^N)^d \mid \exists i \neq j, \|y_i - y_j\| \leq \varepsilon\}.$$

Note that $\mathbb{D}_{N,\varepsilon}$ is an ε-neighborhood around the generalized diagonal $\mathbb{D}_N$.

Theorem 27 (Quantization by barycenters)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a compact convex set, ρ a probability density on Ω and consider a point cloud $Y = (y_1, \dots, y_N)$ in $\Omega^N \setminus \mathbb{D}_N$. Then, for all $0 < \varepsilon \leq 1$,*

$$W_2^2\left(\rho, \frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y)}\right) \leq C_{d,\Omega}\left(\frac{\varepsilon^{1-d}}{N} + 1 - \frac{\mathrm{Card}(I_\varepsilon(Y))}{N}\right). \tag{3.9}$$

where $C_{d,\Omega} = \frac{2^{2d-1}}{\omega_{d-1}}(\mathrm{diam}(\Omega) + 1)^{d+1}$ and where ω_{d-1} is the volume of the unit ball in $\mathbb{R}^{d-1}$.

Notice here that, unlike in the rates of (3.6), the constant does not depend on ρ. However, this comes with a tradeoff, as the exponent in N becomes much worse.

The proof relies on a sort of concavity for the Laguerre cells, with respect to their weights. Let us recall our notation $A \oplus B$ for the Minkowski sum of sets: $A \oplus B = \{a + b \mid (a, b) \in A \times B\}$.

Proof. We denote $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$ and the corresponding measure on the barycenters $B(Y)$: $\mu_b = \frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y)}$.

Let $\phi^1 \in \mathbb{R}^N$ be a solution to the dual Kantorovich problem (1.8) between ρ and μ. We let $\phi^t = t\phi^1$ and we denote $\mathrm{Lag}_i^t = \mathrm{Lag}_i(Y, \phi^t) \cap \Omega'$ the ith Laguerre cell intersected with the slightly enlarged convex set $\Omega' = \Omega \oplus B(0,1)$. This way, $\mathrm{Lag}_i^1 \supseteq \mathrm{Lag}_i(Y) \cap \Omega$ whereas Lag_i^0 is in fact the intersection of the i-th Voronoi cell defined in (3.5) with Ω'.

We will now prove an upper bound on the sum of the diameters of the cells $\mathrm{Lag}_i(Y)$ whose index lies in $I_\varepsilon(Y)$. First, we notice the following inclusion, which holds for any $t \in [0,1]$:

$$(1-t)\mathrm{Lag}_i^0 \oplus t\mathrm{Lag}_i^1 \subseteq \mathrm{Lag}_i^t, \tag{3.10}$$

Indeed, let $x^0 \in \mathrm{Lag}_i^0$ and $x^1 \in \mathrm{Lag}_i^1$, so that for all $j \in \{1,\dots,N\}$ and $k \in \{0,1\}$,

$$\left\|x^k - y_i\right\|^2 - \phi_i^k \leq \left\|x^k - y_j\right\|^2 - \phi_j^k.$$

Expanding the squares and substracting $\left\|x^k\right\|^2$ on both sides these inequalities become *linear* in ϕ_i^k, ϕ_j^k and x^k, implying as desired:

$$x^t = (1-t)x^0 + tx^1 \in \mathrm{Lag}_i^t.$$

For any index $i \in I_\varepsilon$, the point y_i is at distance at least ε from other points, implying that $B(y_i, \frac{\varepsilon}{2})$ is contained in the Voronoi cell $V_i(Y)$ defined using Ω'. Using that $\mathrm{Lag}_i^0 = V_i(Y) \cap \Omega'$, that $\Omega' = \Omega \oplus B(0,1)$ and that $y_i \in \Omega$, we deduce that Lag_i^0 contains the same ball. On the other hand, Lag_i^1 contains a segment S_i of length $\mathrm{diam}(\mathrm{Lag}_i^1)$ and inclusion (3.10) with $t = \frac{1}{2}$ gives

$$\frac{1}{2}(B(y_i, \varepsilon/2) \oplus S_i) \subseteq \mathrm{Lag}_i^{1/2}.$$

The Minkowski sum in the left-hand side contains in particular the product of a $(d-1)$-dimensional ball of radius $\varepsilon/2$ with an orthogonal segment with length $\mathrm{diam}(\mathrm{Lag}_i^1) \geq \mathrm{diam}(\mathrm{Lag}_i(Y))$. Thus,

$$\frac{1}{2^d}\left(\omega_{d-1}\frac{\varepsilon^{d-1}}{2^{d-1}}\mathrm{diam}(\mathrm{Lag}_i(Y))\right) \leq |\mathrm{Lag}_i^{1/2}|.$$

Using that the Power cells $\mathrm{Lag}_i^{\frac{1}{2}}$ form a tesselation of the domain Ω', we obtain

$$\sum_{i \in I_\varepsilon(Y)} \mathrm{diam}(\mathrm{Lag}_i(Y)) \leq \frac{2^{2d-1}}{\omega_{d-1}}|\Omega'|\varepsilon^{1-d} \leq \frac{2^{2d-1}}{\omega_{d-1}}(\mathrm{diam}(\Omega)+1)^d\varepsilon^{1-d} \tag{3.11}$$

We now estimate the transport cost between μ_b and the density ρ. The cost due to the points whose indices do not belong to $I_\varepsilon(Y)$ can be bounded in a crude way by

$$\sum_{i\notin I_\varepsilon(Y)} \int_{\mathrm{Lag}_i(Y)} \|x-b_i\|^2 \,\mathrm{d}\rho(x) \leq (1-\frac{\mathrm{Card} I_\varepsilon(Y)}{N})\mathrm{diam}(\Omega)^2.$$

Note that we used $\rho(\mathrm{Lag}_i(Y)) = \frac{1}{N}$. On the other hand, the transport cost associated with indices in $I_\varepsilon(Y)$ can be bounded using (3.11) and and the fact that $b_i(Y) \in \mathrm{Lag}_i(Y)$:

$$\begin{aligned}\sum_{i\in I_\varepsilon(Y)} \int_{\mathrm{Lag}_i(Y)} \|x-b_i\|^2 \,\mathrm{d}\rho(x) &\leq \frac{1}{N}\sum_{i\in I_\varepsilon(Y)} \mathrm{diam}(\mathrm{Lag}_i(Y))^2 \\ &\leq \frac{1}{N}\mathrm{diam}(\Omega)\sum_{i\in I_\varepsilon} \mathrm{diam}(\mathrm{Lag}_i(Y)) \\ &\leq \frac{2^{2d-1}}{\omega_{d-1}}(\mathrm{diam}(\Omega)+1)^{d+1}\frac{\varepsilon^{1-d}}{N}\end{aligned}$$

In conclusion, we obtain the desired estimate:

$$\begin{aligned}W_2^2(\rho,\mu_b) &\leq \frac{2^{2d-1}}{\omega_{d-1}}(\mathrm{diam}(\Omega)+1)^{d+1}\frac{\varepsilon^{1-d}}{N} + \mathrm{diam}(\Omega)^2\left(1-\frac{\mathrm{Card} I_\varepsilon}{N}\right) \\ &\leq \frac{2^{2d-1}}{\omega_{d-1}}(\mathrm{diam}(\Omega)+1)^{d+1}\left(\frac{\varepsilon^{1-d}}{N} + \left(1-\frac{\mathrm{Card} I_\varepsilon}{N}\right)\right).\end{aligned}$$ □

Note that this result is actually true for more general costs (not only for the quadratic one). The proof of such result make use of a very natural change of variable by taking the so-called c-exponential map. Let us consider a cost c on $\mathbb{R}^d \times \mathbb{R}^d$ such that (we use here the notations from Loeper [43]):

$(A0)$ c is $\mathcal{C}^4$ on $\Omega' \times \Omega$ with $\Omega' := \Omega + \bar{B}(0,\eta)$.

$(A1)$ c verifies the following twist condition on the whole domain Ω' (which is technically a twist condition on $c^* : (x,y) \mapsto c(y,x)$):

$$\forall y \in \Omega,\ x \in \Omega' \mapsto D_y c(x,y)$$

is injective.

$(A2)$ For all $(x,y) \in \Omega' \times \Omega$, $\det(D^2_{x,y} c(x,y)) \neq 0$.

This allows the definition for $Y \in \Omega^N$ of the c-exponential maps

$$\exp^c_{y_i} : p \in \Omega_i \to (-D_y c(., y_i))^{-1}(p) \in \Omega'$$

where the sets Ω_i (the c-exponential charts) are taken such that these maps are $C^{1,1}$ diffeormorphisms. Let us note here ([43], Definition 2.6) that, under conditions $A0 - A2$, these maps are bi-lipshitz with constants that do not depend on the point cloud considered: There exists $K_{c,\Omega} > 0$ such that for any $Y \in \Omega^N$, $i = 1 \dots N$ and $p^1 \neq p^2 \in \Omega_i$,

$$\frac{1}{K_{c,\Omega}} \leq \frac{\left\|\exp^c_{y_i}(p^1) - \exp^c_{y_i}(p^2)\right\|}{\|p^1 - p^2\|} \leq K_{c,\Omega} \tag{3.12}$$

We make the additional following assumption, which is a stronger version of Loeper's condition (Aw) of [43] Theorem 3.2,which is itself a reformulation of the celebrated Ma-Trudinger-Wang condition for the regularity of the transport, [44]). Note that this stronger condition is known from the work of Kim, Figalli and McCann [26] as The Non-Negative-Cross-Curvature (NNCC) condition and is equivalent, under the previous regularity conditions on c, to convexity of the set of c-concave functions (which is essentially the arguments that we use in our proof).

We assume that for any $Y \in \Omega^N$ and $i \in \{1 \dots N\}$, the exponential charts $\Omega_i = -D_x c(\Omega, y_i)$ are convex and that Loeper's functions:

$$p \in \Omega_i \mapsto c(\exp^c_{y_i}(p), y_i) - c(\exp^c_{y_i}(p), y_j) \tag{3.13}$$

are convex for any j.

This assumption is stronger than the usual Loeper's condition (Aw) where these functions were only assumed to be quasi-convex (meaning that their sub-level sets are convex). With these definitions, we can make the corresponding claim to Theorem 27:

Theorem 28. *Let c, Ω and $(Y \in \mathbb{R}^d)^N$ be defined as above in order to verify $A0 - A2$ and the (NNCC) condition. Then, there exists $C > 0$ such that for any ρ a probability density, $0 \leq \varepsilon \leq 1$:*

$$I_c\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{b_i(Y)}, \rho\right) \leq C_{c,d,\Omega}\left(\frac{\varepsilon^{1-d}}{N} + 1 - \frac{\mathrm{Card}(I_\varepsilon(Y))}{N}\right) \tag{3.14}$$

with a constant $C_{c,d,\Omega}$ that only depends on the cost c and the domain $\Omega \subset \mathbb{R}^d$.

Proof. The proof is very similar to that of Theorem 27, and the stronger Loeper's condition is used to obtain a similar concavity of the Laguerre cells with respect

to their weights, only this time through the exponential maps $\exp_i^c$. Indeed, with the notations of the proof of Theorem 27, the Laguerre cells are defined by

$$c(x, y_i) - c(x, y_j) \leq \phi_i - \phi_j$$

and the convexity of Loeper's functions gives us again the "convexity of the Laguerre cells", in the following sense:

if $x_0 = \exp_{y_i}^c(p_0) \in \mathrm{Lag}_i^0$, $x_1 = \exp_{y_i}^c(p_1) \in \mathrm{Lag}_i^1$, then

$$x_{1/2} = \exp_{y_i}^c \left(\frac{1}{2} p_0 + \frac{1}{2} p_1 \right) \in \mathrm{Lag}_i^{1/2}.$$

Because of the bi-Lipschitz condition (3.12), the assumption $i \in I_\varepsilon$ implies that $(\exp_{y_i}^c)^{-1}(\mathrm{Vor}_i(Y) \cap \Omega')$ contains a ball of radius $K_{c,\Omega} \frac{\varepsilon}{2}$. Similarly, for every i, $(\exp_{y_i}^c)^{-1}(\mathrm{Lag}_i^1 \cap \Omega')$ contains a segment of length $K_{c,\Omega} \mathrm{diam}(\mathrm{Lag}_i^1 \cap \Omega')$.

To finish the proof, we rephrase the same arguments on the size of the intermediate Laguerre cells, but in the exponential chart instead of Ω'. For convenience sake, we rewrite ${}^c\mathrm{Lag}_i^t := (\exp_{y_i}^c)^{-1}(\mathrm{Lag}_i^t \cap \Omega')$ and we obtain:

$$\frac{K_{c,\Omega}^d}{2^d} \left(\omega_{d-1} \left(\frac{\varepsilon}{2} \right)^{d-1} \mathrm{diam}({}^c\mathrm{Lag}_i^1) \right) \leq |{}^c\mathrm{Lag}_i^{1/2}|$$

Again, this gives us an upper bound on the sum of the diameters of the Laguerre cells, since the Lebesgue measure of the sets Ω_i are uniformly bounded by a constant that does not depend on Y (because of the bi-Lipschitz bound on the c-exponential maps, once again). The transport costs on the points whose indices are not in $I_\varepsilon(Y)$ are once again crudely upper bounded by $\left(1 - \frac{\mathrm{Card} I_\varepsilon(Y)}{N}\right) \sup_{\Omega^2} c$. The rest of the proof is identical, noting that $c(x, y_i) \leq L \|x - y_i\|$ with a constant L that depends only on c and Ω. □

Theorem 27 could be extended *mutatis mutandis* to the case where ρ is a general probability measure (i.e. not a density). However, this would imply some technical complications in the definition of the barycenters b_i by introducing a disintegration of ρ with respect to the transport plan π.

An immediate application is the case, mentioned earlier, where the points are spread in a pattern of dimension close enough to d, in some sense. More precisely, we assume that the distance between any pair of distinct points of $Y_N \in (\mathbb{R}^d)^N$ is bounded from below by $\varepsilon_N \geq C N^{-1/\beta}$, implying that $I_{\varepsilon_N}(Y_N) = N$. This corresponds to the value one could expect for a point cloud uniformly sampled from a set with Minkowski dimension β. When $\beta > d - 1$, the following corollary asserts that one step of Lloyd's algorithm is enough to approximate ρ, in the sense that the uniform measure $\delta_{B_N(Y_N)}$ over the barycenters converges towards ρ as $N \to +\infty$ (with an exponential convergence rate):

Corollary 28.1 (Quantization by barycenters, asymptotic case)**.** *Assume* $\varepsilon_N \geq C \cdot N^{-1/\beta}$ *with* $C, \beta > 0$*. Then, with* $\alpha = 1 - \frac{d-1}{\beta}$

$$\forall Y \in (\mathbb{R}^d)^N \setminus \mathbb{D}_{\varepsilon_N}, \quad W_2^2\left(\rho, \frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y)}\right) \leq \frac{C_{d,\Omega}}{C^{d-1}} N^{-\alpha}, \tag{3.15}$$

and in particular, if $\beta > d - 1$*,*

$$\lim_{N\to+\infty} \max_{Y\in(\mathbb{R}^d)^N\setminus\mathbb{D}_{\varepsilon_N}} W_2^2\left(\rho, \frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y)}\right) = 0. \tag{3.16}$$

Remark 7. Let us make a note here that the results of Theorem 27 are tight in the following two sense:

- (Optimality of the exponent when $\beta = d$) There is no reason to believe that the exponent in the upper bound (3.15) is optimal in general. However, it seems to be optimal in a "worst-case sense" when $\beta = d$:

  More precisely, for any dimension d, $n \in \mathbb{N}$ and $N = n^d$, there exists a separable probability density ρ_N over $X = [-1, 1]^d$ such that if Y_N is a uniform grid of size $n \times \cdots \times n = N$ in X, then

  $$W_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y^N)}, \rho_N\right) \geq CN^{-\frac{1}{d}},$$

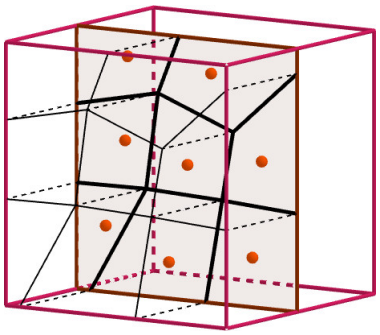
  where C is independent of N. Notice, on the other hand, that in this setting every point in Y_N is at distance at least $CN^{-1/d}$ from any other point in Y_N. Applying 28.1 with $\beta = d$, i.e. $\alpha = \frac{1}{d}$, we get

  $$W_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{b_i(Y^N)}, \rho_N\right) \leq C'N^{-\frac{1}{d}}.$$

  Comparing this upper bound with the previous lower bound, one sees that it is not possible to improve the exponent while keeping a constant independent of ρ.

- (Optimality of (3.16)) The assumption $\beta > d-1$ for (3.16) is tight:

  If ρ is the Lebesgue measure on $[0,1]^d$, it is possible to construct a point cloud Y_N with N points on the $(d-1)$-cube $\{\frac{1}{2}\} \times [0,1]^{d-1}$ such that distinct points in Y_N are at distance at least $\varepsilon_N \geq C \cdot N^{-1/(d-1)}$ (see figure on the right).

  Then, the barycenters $B(Y^N)$ are also contained in the lower-dimensional cube, so that

$$W_2^2\left(\rho, \frac{1}{N}\sum_{i=1}^{N} \delta_{b_i(Y^N)}\right) \geq \frac{1}{12}.$$

The next corollary is a probabilistic analogue of Corollary 28.1, assuming that the initial point cloud Y is drawn from a probability density σ on Ω. Note that σ can be distinct from ρ. The proof of this corollary relies on McDiarmid's inequality to quantify the proportion of ε-isolated points in a point cloud that is drawn randomly and independently according to σ.

Corollary 28.2 (Quantization by barycenters, probabilistic case)**.** *Consider a probability measure $\sigma \in \mathbb{L}^\infty(\Omega) \cap \mathcal{P}(\Omega)$ and let $X_1, ..., X_N$ be i.i.d. random variables with distribution σ. Then, there exists a constant $C > 0$ depending only on $\|\sigma\|_{\mathbb{L}^\infty}$ and d, such that for N large enough,*

$$\mathcal{P}\left(W_2^2\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{b_i(X)}, \rho\right) \lesssim N^{-\frac{1}{2d-1}}\right) \geq 1 - e^{-CN^{\frac{2d-3}{2d-1}}}$$

In a point cloud Y, we shall call *ε-isolated* the points y_i such that $i \in I_\varepsilon(Y)$, and *ε-connected* points y_i such that $i \notin I_\varepsilon(Y)$. Let us also introduce, in the hope of making the proofs clearer, the proportion of ε-isolated points in Y:

$$\kappa(Y) = \frac{1}{N}\mathrm{Card}(I_\varepsilon(Y)).$$

Lemma 5. *Let $X_1, \ldots, X_N$ be independent, $\mathbb{R}^d$-valued, random variables. Then, there is a constant $C_d > 0$ such that*

$$\mathcal{P}(\{|\kappa(X_1, \ldots, X_N) - \mathbb{E}(\kappa)| \geq \eta\}) \leq e^{-N\eta^2/C_d}.$$

Proof. This lemma is a consequence of McDiarmid's inequality. To apply this inequality, we need evaluate the amplitude of variation of the function κ along changes of one of the points x_i. Denote c_d the maximum cardinality of a subset S of the ball $B(0,\varepsilon)$ such that the distance between any distinct points in S is at least ε. By a scaling argument, one can check that c_d does not, in fact, depend on ε. To evaluate

$$|\kappa(x_1,\dots,x_i,\dots,x_N)-\kappa(x_1,\dots,\tilde{x}_i,\dots,x_N)|,$$

we first note that at most c_d points may become ε-isolated when removing x_i. To prove this, we remark that if a point x_j becomes ε-isolated when x_i is removed, this means that $\|x_i-x_j\|\leq\varepsilon$ and $\|x_j-x_k\|>\varepsilon$ for all $k\notin\{i,j\}$. The number of such j is bounded by c_d. Symmetrically, there may be at most c_d points becoming ε-connected under addition of $\hat{x}_i$. Finally, the point x_i itself may change status from ε-isolated to ε-connected. To summarize, we obtain that with $C_d=2c_d+1$,

$$|\kappa(x_1,\dots,x_i,\dots,x_N)-\kappa(x_1,\dots,\tilde{x}_i,\dots,x_N)|\leq\frac{1}{N}C_d.$$

The conclusion then directly follows from McDiarmid's inequality. □

Lemma 6. *Let $\sigma\in\mathbb{L}^\infty(\mathbb{R}^d)$ be a probability density and let $X_1,\dots,X_N$ be i.i.d. random variables with distribution σ. Then,*

$$\mathbb{E}(\kappa(X_1,\dots,X_N))\geq(1-\|\sigma\|_{\mathbb{L}^\infty}\omega_d\varepsilon^d)^{N-1}.$$

Proof. The probability that a point X_i belongs to the ball $B(X_j,\varepsilon)$ for some $j\neq i$ can be bounded from above by $\sigma(B(X_j,\varepsilon))\leq\|\sigma\|_{\mathbb{L}^\infty}\omega_d\varepsilon^d$, where ω_d is the volume of the d-dimensional unit ball. Thus, the probability that X_i is ε-isolated is larger than

$$(1-\|\sigma\|_{\mathbb{L}^\infty}\omega_d\varepsilon^d)^{N-1}.$$

We conclude by noting that

$$\mathbb{E}(\kappa(X_1,\dots,X_N))=\frac{1}{N}\sum_{1\leq i\leq N}\mathcal{P}(X_i\text{ is }\varepsilon\text{-isolated}).$$ □

Proof of Corollary 28.2. We apply the previous Lemma 6 with $\varepsilon_N=N^{-\frac{1}{\beta}}$ and $\beta=d-\frac{1}{2}$. The expectation of $\kappa(X_1,\dots,X_N)$ is lower bounded by:

$$\begin{aligned}\mathbb{E}(\kappa(X_1,\dots,X_N))\geq&\left(1-N^{-\frac{d}{\beta}}\|\sigma\|_{\mathbb{L}^\infty}\omega_d\right)^{N-1}\\ \geq&1-CN^{1-\frac{d}{\beta}}\end{aligned}$$

for large N, since $\beta < d$. By Lemma 5, for any $\eta > 0$,

$$\mathcal{P}(\kappa(X_1, \dots, X_N) \geq 1 - CN^{1-\frac{d}{\beta}} - \eta) \geq 1 - e^{-KN\eta^2},$$

for constants $C, K > 0$ depending only on $\|\sigma\|_{\mathbb{L}^\infty}$ and d. We choose $\eta = N^{-\frac{1}{2d-1}}$, so that η is of the same order as $N^{1-\frac{d}{\beta}}$ since $1 - \frac{d}{\beta} = -\frac{1}{2d-1}$.Thus, for a slightly different C,

$$\mathcal{P}(\kappa(X_1, \dots, X_N) \geq 1 - C\eta) \geq 1 - \mathrm{e}^{-KN\eta^2}.$$

Now, for $\omega_1, \dots, \omega_N$ such that

$$\kappa(X_1(\omega_1), \dots, X_N(\omega_N)) \geq 1 - C\eta,$$

Theorem 27 yields:

$$W_2^2\left(\delta_{B_N(X(\omega))}, \rho\right) \lesssim \frac{N^{\frac{d-1}{\beta}}}{N} + \eta \lesssim N^{-\frac{1}{2d-1}}$$

and such a disposition happens with probability at least

$$1 - \mathrm{e}^{-KN\eta^2} = 1 - \mathrm{e}^{-KN^{\frac{2d-3}{2d-1}}}.$$ □

Finally, theorem 27 can be interpred as a modified Polyak-Łojasiewicz-type (PŁ for short) inequality for the function $\mathcal{F}_N$. The usual PŁ inequality for a differentiable function $F : \mathbb{R}^D \to \mathbb{R}$ is of the form

$$\forall Y \in \mathbb{R}^D, \quad F(Y) - \min F \leq C \left\|\nabla F(Y)\right\|^2,$$

where C is a positive constant. This inequality has been originally used by Polyak [57] to prove convergence of gradient descent towards the global minimum of F. Note in particular that such an inequality implies that any critical point of F is a global minimum of F. As could be seen on Fig. 3.1, $\mathcal{F}_N$ has critical points that are not minimizers, so that we cannot expect the standard PŁ inequality to hold. What we get is a similar inequality relating $\mathcal{F}_N(Y)$ and $\|\nabla\mathcal{F}_N(Y)\|^2$ but with a term involving the minimimum distance between the points in place of $\min \mathcal{F}_N$.

Corollary 28.3 (Polyak-Łojasiewicz-type inequality)**.** *Let $Y \in (\mathbb{R}^d)^N \setminus \mathbb{D}_{N,\varepsilon}$. Then,*

$$\mathcal{F}_N(Y) - C_{d,\Omega}\frac{1}{N}\left(\frac{1}{\varepsilon}\right)^{d-1} \leq N \left\|\nabla\mathcal{F}_N(Y)\right\|^2 \tag{3.17}$$

Proof. We first note that by Proposition 24, we have

$$\left\|\nabla\mathcal{F}_N(Y)\right\|^2 = \frac{1}{N^2}\left\|B_N(Y) - Y\right\|^2.$$

We then use $W_2^2(\delta_{B_N(Y)}, \delta_Y) \leq \frac{1}{N} \|B_N(Y) - Y\|^2$ to get:

$$W_2^2(\rho, \delta_Y) \leq 2W_2^2(\rho, \delta_{B_N(Y)}) + 2N \|\nabla \mathcal{F}_N(Y)\|^2 .$$

Thus, using Theorem 27 to bound $W_2^2(\rho, \delta_{B_N(Y)})$ from above, we get the desired result. □

We note that when $\varepsilon \simeq \left(\frac{1}{N}\right)^{1/d}$, the term $\frac{1}{N}\left(\frac{1}{\varepsilon}\right)^{d-1}$ in (3.17) has order $\left(\frac{1}{N}\right)^{1/d}$. On the other hand, we recall (3.6), $\min \mathcal{F}_N \lesssim \left(\frac{1}{N}\right)^{2/d}$ when $d > 2$. Thus, inequality (3.17) is not truly a PŁ-inequality, as the order of the second term $\frac{1}{N}\left(\frac{1}{\varepsilon}\right)^{d-1}$ is not that of the minimum of F.

3.3 Numerical results

In this section, we report some experimental results in dimension $d = 2$. This is by no means a performance study, but rather a set of observations, related to the theoretical convergence results of the previous sections. For actual efficient implementations for uniform quantization of measure, and in particular speed comparisons, we refer the reader to articles on Blue Noise Sampling, [17] for computations using Lloyd's algorithm and [70] for a more general overview.

Gray-scale picture A first, somewhat toy-, application for optimal quantization is the sparse representation of a grey-scale image, via points clustering more closely in darker areas. On figure 3.2, we plotted the point clouds obtained after a single Lloyd step toward the density representing the image on the left (Puffin), starting from regular grids. Underneath the pictures are the graphs, in log-log scale of the quantization error with respect to the number of points. The observed rates of convergence, close to $N^{-1.00}$, are coherent with the theoretical estimate $O(\log(N)/N)$ of (3.6). In the background of the Puffin, one can also observe that some structure of the starting grid has been conserved after the Lloyd step. These artefacts are actually symptomatic of a remarkable phenomenon, which is that Laguerre cells with respect to a separable density on one side and a grid-structured point cloud on the other will also be aligned along a grid. This result is proven in the next chapter, Proposition 33.

Examples in separable cases: In this example, we showcase the optimality of the exponent in Corollary 28.1, even under assumptions of regularity on the underlying density.

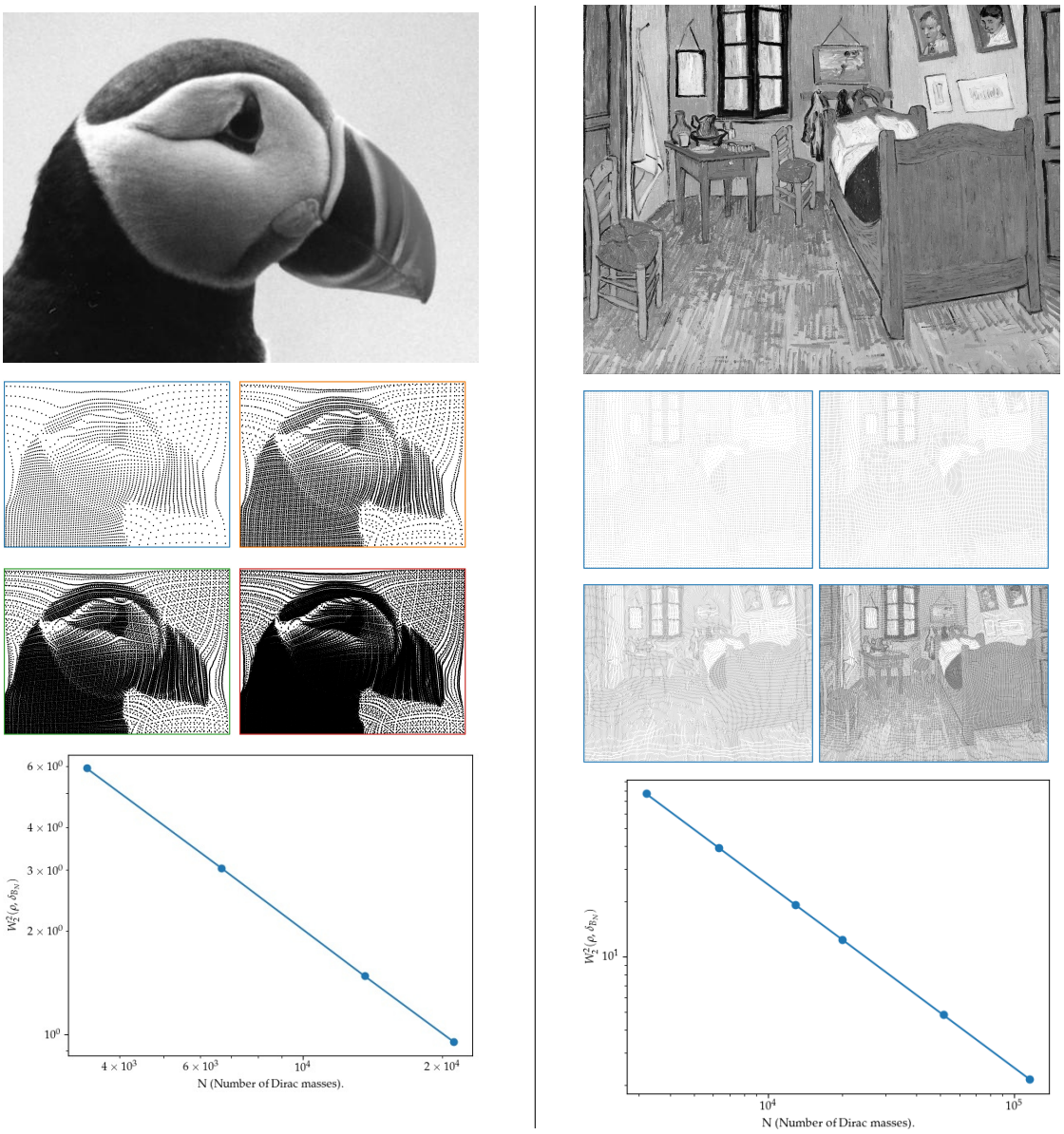


Figure 3.2: Approximative optimal quantization of a density ρ corresponding to a gray-scale image (Wikimedia Commons, CC BY-SA 3.0). (Middle) We display the point clouds obtained after one step of Lloyd's algorithm, starting from a regular grid of size $N \in \{3750, 7350, 15000, 43350\}$. (Bottom) Quantization error $W_2^2(\rho, \delta_{B_N})$ as a function of N the number of points, showing that $W_2^2(\rho, \delta_{B_N}) \simeq N^{-1.00}$.

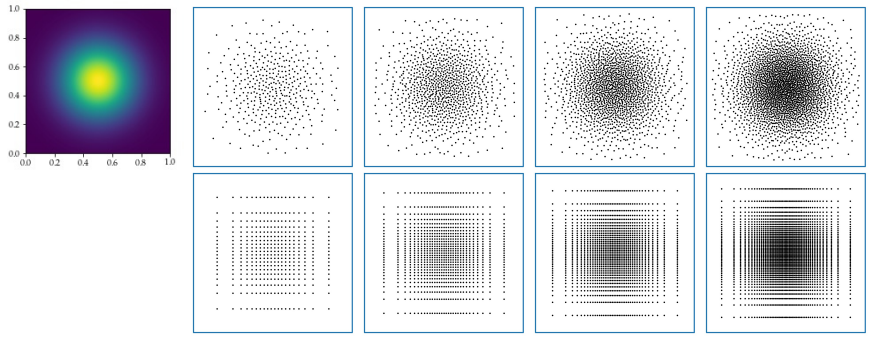


Figure 3.3: Point clouds obtained after one step of the Lloyd algorithm in the sampling of the Gaussian density represented on the top left. The point clouds on top were constructed starting from a random initial distribution of 400, 961,1600 and 2500 points. The ones on the bottom were reached from an initial point cloud aligned on a regular grid (of the same sizes).

We first observe what happens when the approximated density is a Gaussian one (truncated to the unit square $\Omega = [0,1]^2$), with low variance:

$$\rho(x,y) = \frac{1}{Z} e^{-8((x-\frac{1}{2})^2+(y-\frac{1}{2})^2)}$$

where Z is a normalization constant. In the experiments represented on the left column of Fig. 3.3, the initial point clouds Y_N^0 are randomly sampled uniformly in $[0,1]^2$, and the subsequent pictures represent the barycenters of the Laguerre cells. In this case, we observe a decrease rate for the quantization energy $N^{-0.95}$ with respect to the number of points, similar to the case of the gray scale images on Fig. 3.2.

However, when starting from a regular grid (pictures on the left of Fig. 3.3), we observed a similar grid-like formation for the barycenters, which is stable (meaning the Lloyd algorithm does not move the points after the first step). This formation also gave us a much worse quantization error, of the order of $N^{-0.78}$.

Of course, the cause of this behavior is the separability of the density in this case (meaning a Gaussian one). This comes down directly from the same configuration in dimension one, using the factorization of the Laguerre cells in the separable case (Proposition 33):

Proposition 29. *For any $\sigma > 0$, consider $\rho_\sigma := m_\sigma \mathrm{e}^{-\frac{|x|^2}{2\sigma^2}} \mathbb{1}_{[-1;1]} dx$ the truncated centered Gaussian density, where m_σ is taken so that ρ_σ has unit mass. Then, for every $\delta \in (0,1)$, there exists a constant $C > 0$ and a sequence of variances $(\sigma_N)_{N\in\mathbb{N}}$ such that*

$$\forall Y \in (\mathbb{R}^d)^N \setminus \mathbb{D}_N, \quad W_2^2\left(\delta_{B_N(Y)}, \rho_{\sigma_N}\right) \geq C N^{-(2-\delta)}$$

Proof. We denote $g : x \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{|x|^2}{2}}$ the density of the centered Gaussian distribution and F_g its cumulative distribution function, so that

$$m_\sigma^{-1} = \int_{-1}^{1} \mathrm{e}^{-\frac{|x|^2}{2\sigma^2}} dx = \sigma\sqrt{2\pi} \int_{-1/\sigma}^{1/\sigma} g(y)dy = \sqrt{2\pi}\sigma(F_g(1/\sigma) - F_g(-1/\sigma)) \quad (3.18)$$

Note that, whenever $\sigma \to 0$, we have $\sigma m_\sigma \to (\sqrt{2\pi})^{-1}$. We denote the cumulative distribution function of ρ_σ by $F_\sigma : [-1, 1] \to [0, 1]$. Given any point cloud $Y = (y_1, \dots, y_N)$ such that $y_1 \leq \cdots \leq y_N$, the Power cells $P_i(Y)$ is simply the segment

$$P_i(Y) = [F_\sigma^{-1}(i/N), F_\sigma^{-1}((i+1)/N)].$$

Since these segments do not depend on Y, we will denote them $(P_i)_{1\leq i\leq N}$. Finally, defining $b_i = N \int_{P_i} x \mathrm{d}\rho_\sigma(x)$ as the barycenter of the ith power cell and $\mu_B = \frac{1}{N}\sum_i \delta_{b_i}$, we have

$$\begin{aligned} W_2^2(\mu_B, \rho_\sigma) &= \sum_{i=1}^{N} \int_{P_i} (x - b_i)^2 \mathrm{d}\rho_\sigma(x) \\ &\geq \rho_\sigma(-1) \sum_{i=1}^{N} \int_{P_i} (x - b_i)^2 \mathrm{d}x \qquad (3.19) \\ &\geq C\rho_\sigma(-1) \sum_{i=1}^{N} (F_\sigma^{-1}((i+1)/N) - F_\sigma^{-1}(i/N))^3, \end{aligned}$$

where we used that ρ_σ attains its minimum at ± 1 to get the first inequality.

We now wish to provide an approximation for $F_\sigma^{-1}(t)$, $t \in [0, 1]$. We first note, using Taylor's formula, that we have

$$\begin{aligned} F_\sigma^{-1}(t) &= \sigma F_g^{-1}\left(F_g\left(\frac{-1}{\sigma}\right) + t\left[F_g\left(\frac{1}{\sigma}\right) - F_g\left(\frac{-1}{\sigma}\right)\right]\right) \\ &= \sigma F_g^{-1}\left(F_g\left(\frac{-1}{\sigma}\right) + \frac{t}{\sqrt{2\pi}\sigma m_\sigma}\right) \\ &= -1 + \sigma(F_g^{-1})'\left(F_g\left(\frac{-1}{\sigma}\right)\right)\frac{t}{\sqrt{2\pi}\sigma m_\sigma} + \frac{\sigma}{2}(F_g^{-1})''(s)\frac{t^2}{2\pi\sigma^2 m_\sigma^2} \end{aligned}$$

for some $s \in [F_g(-\frac{1}{\sigma}), F_g(-\frac{1}{\sigma}) + t(F_g(\frac{1}{\sigma}) - F_g(-\frac{1}{\sigma}))]$. But,

$$(F_g^{-1})'(t) = \frac{1}{g \circ F_g^{-1}(t)} = \sqrt{2\pi}\mathrm{e}^{\frac{|F_g^{-1}(t)|^2}{2}},$$

$$(F_g^{-1})''(t) = -\frac{g' \circ F_g^{-1}(t)}{\left(g \circ F_g^{-1}(t)\right)^3} = 2\pi F_g^{-1}(t)\mathrm{e}^{|F_g^{-1}(t)|^2},$$

and we see that

$$|F_\sigma^{-1}(t) - \left(-1 + \frac{t}{m_\sigma}\mathrm{e}^{\frac{1}{2\sigma^2}}\right)| \leq \mathrm{e}^{\frac{1}{\sigma^2}}\frac{t^2}{2\sigma^2 m_\sigma^2}$$

Therefore, if we denote $\varepsilon(\sigma, t)$ the second-order error in the above formula, i.e. $\varepsilon(\sigma,t) = \mathrm{e}^{\frac{1}{\sigma^2}}\frac{t^2}{2\sigma^2 m_\sigma^2}$, the size of the first Power cell $P_0(Y)$ is of order:

$$F_\sigma^{-1}(1/N) - F_\sigma^{-1}(0) = \frac{1}{Nm_\sigma}\mathrm{e}^{\frac{1}{2\sigma^2}} + O\left(\varepsilon\left(\sigma, \frac{1}{N}\right)\right).$$

We will choose σ_N depending on N in order for the first term in the left-hand side to dominate the second one:

$$\varepsilon\left(\sigma_N, \frac{1}{N}\right) = o\left(\frac{1}{Nm_\sigma}\mathrm{e}^{\frac{1}{2\sigma^2}}\right). \tag{3.20}$$

In this way, we have

$$\begin{aligned}(F_\sigma^{-1}(1/N) - F_\sigma^{-1}(0))^3\rho_\sigma(-1) &\geq \frac{c}{N^3 m_\sigma^3}\mathrm{e}^{\frac{3}{2\sigma^2}}m_\sigma\mathrm{e}^{-\frac{1}{2\sigma^2}} \\ &= \frac{c}{N^3 m_\sigma^2}\mathrm{e}^{\frac{1}{\sigma^2}}.\end{aligned} \tag{3.21}$$

We now choose $\sigma = \sigma_N$ such that $\mathrm{e}^{\frac{1}{2\sigma^2}} = N^\alpha$ for an exponent α to be chosen. We need $\alpha > 0$ so that $\sigma_N \to 0$. This last condition and (3.18) implies that m_{σ_N} is of order $\sqrt{\log N}$. This means that the condition (3.20) is satisfied if $\alpha < 1$ and N large enough.

The sum in (3.19) is lower bounded by its first term, (3.21), and we get

$$W_2^2(\delta_B, \rho_\sigma) \geq c\frac{1}{N^3 m_{\sigma_N}^2}\mathrm{e}^{\frac{1}{\sigma_N^2}} \geq C\left(\frac{N^{2\alpha-3}}{\ln(N)}\right)$$

for some constant $C > 0$, since σ depends logarithmically on N. Finally, if we want this last expression to be larger than $N^{-(2-\delta)}$ we can take for instance $2\alpha > 1 + \delta$ and N large enough. □

Corollary 29.1. *Fix $\delta \in (0,1)$. Given any $n \in \mathbb{N}$, consider an axis-aligned discrete grid of the form $Z_N = Y_1 \times \cdots \times Y_d$ in $\mathbb{R}^d$, with $N = \mathrm{Card}(Z_N) = n^d$, where each Y_j is a subset of $\mathbb{R}$ with cardinal n. Finally, define $\sigma_N := \sigma_{n,\delta}$ as in Proposition 29 Then we have*

$$W_2^2(\delta_{B_N(Z_N)}, \rho_{\sigma_N} \otimes \cdots \otimes \rho_{\sigma_N}) \geq CN^{-\frac{(2-\delta)}{d}},$$

where the constant C is independent of N.

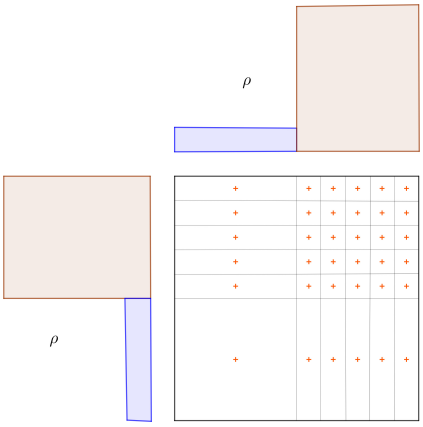


Figure 3.4: A critical point for $\mathcal{F}_N$, with ρ given by (3.22), which realises a good approximation in the sense of Theorem 27, but does not realize a minimum for $\mathcal{F}_N$, for large N.

Note that one can construct a much simpler example of a family of densities that cannot be approximated by discrete uniform measures at a better rate (which, however, does not share the smoothness of ρ_{σ_N}):

This can be done, as for the Gaussian example in dimension $d = 1$ and then generalized to higher dimensions using separable measures (see Fig. 3.4 below):

For $N \in \mathbb{N}^*$ set

$$\rho_N := \left(\frac{1}{N+1} \mathbb{1}_{[-1;0]} + \frac{N}{N+1} \mathbb{1}_{[0;1]} \right) \mathrm{d}x \tag{3.22}$$

Then, the only critical point for $\mathcal{F}_N$ (defined from the density ρ_N) is the cloud:

$$Y_N = \left(-\frac{1}{2}, \frac{1}{2N}, \frac{3}{2N}, \dots, \frac{2N-1}{2N} \right)$$

and

$$\mathcal{F}_N(Y_N) = \frac{1}{12N} + o_{N\to\infty}\left(\frac{1}{N}\right).$$

Now, for any d, consider

$$\rho_N^d := \rho_N^{\otimes d} = \rho_N \otimes \cdots \otimes \rho_N$$

on the domain $[-1;1]^d$. The point cloud $Y_N^{\otimes d}$ defined as the tensor product of the set of coordinates of Y_N d times with itself, is a critical point for $\mathcal{F}_N$ (this time with the density ρ_N^d) and it is the point cloud of barycenters $B(Y)$ for any

starting cloud Y aligned on the main axes, e.g. sampled on a regular grid. Using the decomposition of the Laguerre cells from Proposition 33, we obtain

$$\begin{aligned}\mathcal{F}_N(B(Y)) = \mathcal{F}_N(Y_N^{\otimes d}) =& \frac{d}{12N} + o_{N\to\infty}\left(\frac{1}{N}\right)\\ \geq& \frac{C}{(N^d)^{\frac{1}{d}}}\end{aligned}$$

which is the convergence rate of Corollary 28.1.

3.4 Applications to the initialization of some particle flows

We finish this chapter by studying the implication of Theorem 27 on the choice of initial datum for some minimizing movements, whose energy features a semi-discrete transport term. To clarify, we first consider the simple gradient flow of $\mathcal{F}_N$, giving us convergence bounds, but ones that cannot be used in practice and consist more in a proof of concept. We then apply the same techniques to the (more interesting) case of uncongested crowd motions as a Wasserstein gradient flow.

Gradient flow for the uniform quantization energy: The modified Polyak-Łojasiewicz inequality (3.17) suggests that the discrete gradient flow (3.7) will bring us close to a point cloud with low Wasserstein distance to ρ, provided we can guarantee that the the points in the clouds Y^k remain far from the generalized diagonal $\mathbb{D}_N$ during the iterations. We prove in Lemma 7 below that if $Y^{k+1} = Y^k - \tau_N \nabla\mathcal{F}_N(Y^k)$ and $\tau_N \in (0,1)$, then

$$\forall i \neq j, \quad \left\|y_i^{k+1} - y_j^{k+1}\right\| \geq (1-\tau_N)\left\|y_i^k - y_j^k\right\|. \tag{3.23}$$

We note that this inequality ensures that Y^k never touches the generalized diagonal $\mathbb{D}_N$, so that the gradient $\nabla\mathcal{F}_N(Y^k)$ is well-defined at each step (we already proved an analogous result for the limit case $\tau_N = 1$ in Proposition 25). Combining this inequality with Theorem 27, one can actually prove that if the points in the initial cloud Y_N^0 are not too close to each other, then a few steps of gradient discrete gradient descent leads to a discrete measure Y_N^k that is close to the target ρ. Precisely, we arrive at the following theorem:

Theorem 30. *Let $0 < \alpha < \frac{1}{d-1} - \frac{1}{d}$, $\varepsilon_N \gtrsim N^{-\frac{1}{d}-\alpha}$, and $Y_N^0 \in \Omega^N \setminus \mathbb{D}_{\varepsilon_N}$. Let $(Y_N^k)_k$ be the iterates of* (3.7) *starting from Y_N^0 with timestep $0 < \tau_N < 1$. We assume that*

$\lim_{N\to\infty} \tau_N = 0$ *and we set*

$$k_N = \left\lfloor \frac{1}{d\tau_N} \ln(\mathcal{F}_N(Y_N^0) N \varepsilon_N^{d-1}) \right\rfloor .$$

Then,

$$W_2^2\left(\rho, \delta_{Y_N^{k_N}}\right) = O_{N\to\infty}\left(W_2^2\left(\rho, \delta_{Y_N^0}\right)^{1-\frac{1}{d}} . N^{\frac{-1}{d^2}+\alpha\left(1-\frac{1}{d}\right)}\right). \tag{3.24}$$

Remark 8. Note that the exponential behavior implied by 3.23 and Lemma 7 is coherent with the estimates that are known in the absolutely continuous setting for the continuous gradient flow. When transitioning from discrete measures to probability densities, lower bounds on the distance between points become upper bounds on the density. The gradient flow $\dot{\mu}_t = -\frac{1}{2}\nabla_\mu W_2^2(\rho, \mu_t)$ has an explicit solution $\mu_t = \sigma_{1-e^{-t}}$, where σ is a constant-speed geodesic in the Wasserstein space with $\sigma_0 = \mu_0$ and $\sigma_1 = \rho$. In this case, a simple adaptation of the estimates in Theorem 2 in [62] shows the bound $\|\mu_t\|_{\mathbb{L}^\infty} \leq e^{td} \|\mu_0\|_{\mathbb{L}^\infty}$. Still in this absolutely continous setting, it is possible to remove the exponential growth if the target density is also bounded, as a consequence of *displacement convexity* [48, Theorem 2.2]. There seems to be no discrete counterpart to this argument, explaining in part the discrepancy between the exponent of N in (3.24) with the one obtained in Corollary 28.1.

Lemma 7. *Let $Y^0 \in (\mathbb{R}^d)^N \setminus \mathbb{D}_{N,\varepsilon_N}$ for some $\varepsilon_N > 0$. Then, the iterates $(Y^k)_{k\geq 0}$ of* (3.7) *satisfy for every $k \geq 0$, and for every $i \neq j$*

$$\left\|y_i^k - y_j^k\right\| \geq (1-\tau_N)^k \varepsilon_N \tag{3.25}$$

Proof. We consider the distance between two trajectories after k iterations:

$$e_k = \left\|y_i^k - y_j^k\right\| .$$

Assuming that $e_k > 0$, the convexity of the norm immediately gives us:

$$\begin{aligned} e_{k+1} - e_k \geq & \left(\frac{y_i^k - y_j^k}{\left\|y_i^k - y_j^k\right\|}\right) \cdot \left(y_i^{k+1} - y_j^{k+1} - \left(y_i^k - y_j^k\right)\right) \\ = & \tau_N \left(\frac{y_i^k - y_j^k}{\left\|y_i^k - y_j^k\right\|}\right) \cdot \left(b_i^k - b_j^k\right) - \tau_N \left\|y_i^k - y_j^k\right\| \end{aligned}$$

where we denoted $b_i^k := b_i(Y_N^k)$ the barycenter of the ith Laguerre cell $\mathrm{Lag}_i(Y_N^k)$ in the tesselation associated with the point cloud Y_N^k. Since each barycenter b_i^k

lies in its corresponding Laguerre cell, the scalar product $\left(y_i^k - y_j^k\right) \cdot \left(b_i^k - b_j^k\right)$ is non-negative: Indeed, for any $i \neq j$,

$$\left\|y_i^k - b_i^k\right\|^2 - \left\|y_j^k - b_i^k\right\|^2 \leq \phi_i^k - \phi_j^k$$

Summing this inequality with the same inequality with the roles of i and j reversed, we obtain:

$$\left(y_i^k - y_j^k\right) \cdot \left(b_i^k - b_j^k\right) \geq 0$$

thus giving us the geometric inequality $e_{k+1} \geq (1 - \tau_N)e_k$. Since Y_N^0 was chosen in $\Omega^N \setminus \mathbb{D}_{N,\varepsilon_N}$, this yields $e_k \geq (1 - \tau_N)^k e_0$ and inequality 3.25. □

Lemma 8. *For any $k \geq 0$*

$$\mathcal{F}_N(Y_N^k) \leq \mathcal{F}_N(Y_N^0)\eta_N^k + 2C_{d,\Omega}(1 - \eta_N)\frac{\varepsilon_N^{1-d}}{N}\frac{A_N^k - \eta_N^k}{A_N - \eta_N}, \tag{3.26}$$

where we denote $\eta_N = 1 - \frac{\tau_N}{2}(2 - \tau_N)$ and $A_N = (1 - \tau_N)^{1-d}$.

Proof. This is obtained in a very similar fashion as Lemma 7. For any $k \geq 0$, the semi-concavity of $\mathcal{F}_N$ yields the inequality:

$$\mathcal{F}_N(Y_N^{k+1}) - \frac{\left\|Y_N^{k+1}\right\|^2}{2N} - \left(\mathcal{F}_N(Y_N^k) - \frac{\left\|Y_N^k\right\|^2}{2N}\right) \leq \left(-\frac{B_N^k}{N}\right) \cdot \left(Y_N^{k+1} - Y_N^k\right)$$

with $B_N^k := B(Y_N^k)$ in accordance with the previous proof.

Rearranging the terms,

$$\begin{aligned}
\mathcal{F}_N(Y_N^{k+1}) - \mathcal{F}_N(Y_N^k) \leq & -\tau_N(1 - \frac{\tau_N}{2})\frac{\left\|B_N^k - Y_N^k\right\|^2}{N} \\
= & -\tau_N(1 - \frac{\tau_N}{2})W_2^2(\delta_{B_N^k}, \delta_{Y_N^k}) \\
\leq & \tau_N(1 - \frac{\tau_N}{2})\left(-\frac{1}{2}W_2^2(\delta_{Y_N^k}, \rho) + W_2^2(\rho, \delta_{B_N^k})\right)
\end{aligned}$$

by applying first the triangle inequality to $W_2(\delta_{B_N^k}, \delta_{Y_N^k})$. Using Theorem 27 and Lemma 7, this yields:

$$\begin{aligned}
\mathcal{F}_N(Y_N^{k+1}) \leq & (1 - \frac{\tau_N}{2}(2 - \tau_N))\mathcal{F}_N(Y_N^k) + C_{d,\Omega}\tau_N(2 - \tau_N)\frac{\varepsilon_N^{1-d}}{N}(1 - \tau_N)^{k(1-d)} \\
\leq & \eta_N \mathcal{F}_N(Y_N^k) + 2C_{d,\Omega}(1 - \eta_N)\frac{\varepsilon_N^{1-d}}{N}A_N^k.
\end{aligned}$$

and we simply iterate on k to end up with the bound claimed in Lemma 8. □

Proof of Theorem 30. To conclude, we simply make (order 1) expansions of the terms in (3.26). The definition of k_N in Theorem 30, although convoluted, was made so that both terms in the right-hand side of this inequality, $\mathcal{F}_N(Y_N^0)\eta_N^{k_N}$ and $(1-\eta_N)\frac{\varepsilon_N^{1-d}}{N}\frac{A_N^{k_N}-\eta_N^{k_N}}{A_N-\eta_N}$ have the same asymptotic decay to 0 (as $N \to +\infty$): With the notations of the previous proposition, we have for fixed N:

$$W_2^2\left(\rho, \delta_{Y_N^{k_N}}\right) \leq W_2^2\left(\rho, \delta_{Y_N^0}\right)\eta_N^{k_N} + 4C_{d,\Omega}\frac{(1-\eta_N)}{A_N-\eta_N}\frac{A_N^{k_N}-\eta_N^{k_N}}{N\varepsilon_N^{d-1}} \tag{3.27}$$

Define, for clarity's sake, the maximum duration of this gradient descent:

$$T_N = k_N\tau_N = \frac{1}{d}\ln\left(\mathcal{F}_N(Y_N^0)N\varepsilon_N^{d-1}\right) + O_{N\to\infty}(\tau_N).$$

Because of the assumption $\lim_{N\to\infty}\tau_N = 0$, we may write:

$$\frac{A_N^{k_N}-\eta^{k_N}}{N\varepsilon_N^{d-1}} = \frac{\mathrm{e}^{(d-1)T_N}}{N\varepsilon_N^{d-1}} + o_{N\to\infty}\left(\frac{T_N}{(N\varepsilon_N^{d-1})^{\frac{1}{d}}}\right)$$

as well as $\eta^{k_N} = \mathrm{e}^{-T_N} + o_{N\to\infty}\left(\frac{T_N}{(N\varepsilon_N^{d-1})^{\frac{1}{d}}}\right)$, and substituting T_N,

$$\begin{aligned} W_2^2\left(\rho, \delta_{Y_N^{k_N}}\right) \lesssim & \frac{W_2^2\left(\rho, \delta_{Y_N^0}\right)^{\frac{d-1}{d}}}{\left(N\varepsilon_N^{d-1}\right)^{\frac{1}{d}}} + o_{N\to\infty}\left(\frac{T_N}{(N\varepsilon_N^{d-1})^{\frac{1}{d}}}\right) \\ \lesssim & W_2^2\left(\rho, \delta_{Y_N^0}\right)^{1-\frac{1}{d}} N^{\frac{-1}{d^2}+\alpha\left(1-\frac{1}{d}\right)} \end{aligned}$$

□

Initialization of Wasserstein gradient flows: We now turn our attention to discretized congested crowd motions in the gradient flow setting of Mérigot et al. [40] and we make some observations for our model of crowd motions.

In the case of Mérigot et al., we recall that the motion of a crowd avoiding congestion is modelized by the gradient flow, in the Wasserstein space $\mathcal{P}(\Omega)$, of an energy $E(\mu) = F(\mu) + \int_\Omega V d\mu$ where F has the two, now familiar, forms:

$$F(\rho) = \int_\Omega \rho\log(\rho)$$

and

$$F(\rho) = \chi_{\rho\leq \mathrm{d}x}(\rho) = \int_\Omega \chi_{[0;1]}(\rho(x))\mathrm{d}x$$

(χ_A is the convex indicator function of the set A). Let us also note that the potential V is assumed to be $\mathcal{C}^1(\mathbb{R}^d)$.

As we mentioned, the discretization of the problem is almost identical to that of Section 2.2 and the authors look for discrete solutions in $\mathcal{C}^0([0;T],\mathcal{P}_N(\Omega))$ induced by curves $X^N=(x_1^N,\dots x_N^N)$ verifying the differential system:

$$\begin{cases}(x_i^N)'(t)=-\nabla_{x_i}\mathcal{F}_\varepsilon(X^N(t))-\nabla V(x_i(t))\\ X^N(0)=X_0^N\end{cases} \tag{3.28}$$

As in the case of discretized Mean Field Games, one expects to recover the optimal motion for the Wasserstein gradient flow of E, as $\varepsilon\to 0$ and $N\to\infty$.

We recall the general form of the main two results (one for each congestion penalty) of this paper, Theorem 2.1 and 3.1:

Theorem 31 (Leclerc, Mérigot, Santambrogio, Stra)**.** *Let $\mu^0\in\mathcal{P}(\Omega)$ be such that $F(\mu^0)<+\infty$. For every $N\in\mathbb{N}$, let $\varepsilon_N\in(0,\infty)$ and $\mu_N(0)\in\mathcal{P}_N(\mathbb{R}^d)$. Finally, let $X^N\in\mathcal{C}^1([0;T],(\mathbb{R}^d)^N)$ be a solution of* (3.28) *and $\mu_N:[0;T]\to\mathcal{P}_N(\mathbb{R}^d)$ be the corresponding curve of measures.*

Assume that

$$W_2^2(\mu^0,\mu_N(0))\le C\varepsilon_N,\qquad \lim_{N\to+\infty}\varepsilon_N=0 \tag{3.29}$$

and

$$\int_0^T W_2^2\left(\rho_N,\frac{1}{N}\sum_{i=1}^N\delta_{b_i(X^N)}\right)dt\le C\varepsilon_N \tag{3.30}$$

where ρ_N is the Moreau-Yosida projection of μ_N for F and a regularization parameter ε_N, $b_i(X^N)$ is the ρ_N-barycenter of the i-th optimal Laguerre cell defined by this regularization (see Chapter 1, Proposition 11) and $C>0$ is a constant independent of N.

Then, as $N\to\infty$ and up to subsequences, μ_N converges to $\mu\in\mathcal{C}^0([0;T],\mathcal{P}(\Omega))$ for the uniform convergence associated to the 2-Wasserstein distance, where μ is a weak solution of the Wasserstein gradient flow for E, starting at the measure μ^0.

Let us note here that the hypothesis (3.30) is one that is not satisfying and one would like (3.30) to automatically hold, depending on an adequate choice of ε_N. However, assuming that the left-hand side of this inequality will decay to 0 in a uniform way, as $N\to\infty$ (so that ε_N might be chosen à-priori) is non-trivial. The authors manage to get rid of this hypothesis in the 1-dimensional case (where

the barycenters $b_i(Y)$ are always at the optimum for $\mathcal{F}_N$, see Chapter 4, Proposition 32).

To find ε_N verifying the hypotheses of Theorem 31, we can use Theorem 27 in order to bound the distance between the discrete measure on the barycenters and ρ_N during the motion, similarly to the gradient flow case. This, of course also requires that a reverse-Gronwall inequality in the style of Lemma 7 exists for the position of the curves $x_1^N, \dots, x_N^N$. Luckily, this is the case, provided we have initial datum that was sufficiently spread, and that the potential term V is semi-concave (which was the case in their numerical computations, where V even had Lipschitz-continuous gradient).

Lemma 9. *Let X^N be defined for every N as in Theorem 31, with a $\mathcal{C}^1(\mathbb{R}^d)$, semi-concave potential V. Assume that there exists a constant $\eta > 0$, such that for any $i \neq j$, $\left\|x_i^N(0) - x_j^N(0)\right\| \geq \eta > 0$.*

Then, for any $t \in [0;T]$, $\left\|x_i^N(t) - x_j^N(t)\right\| \geq e^{-K_N t}\eta$

where $K_N = \frac{1}{\varepsilon_N} + L$ and L is a constant of semi-concavity of V

In particular, $W_2^2\left(\rho_N(t), \frac{1}{N}\sum_{i=1}^N \delta_{b_i(X^N(t))}\right) \leq C_{d,\Omega}\frac{e^{K_N(d-1)t}}{N\eta^{d-1}}$, with the constant $C_{d,\Omega}$ from Theorem 27.

Proof. The proof of the first part of the lemma is similar to that of Lemma 7. Each curve $(x_i^N)_i$ satisfies the first order differential equation:

$$(x_i^N)' = \frac{b_i(X^N) - x_i^N}{\varepsilon_N} - \nabla V(x_i^N)$$

Now, from the definition of L as a semi-concavity constant:

$$\text{For } x, y \in \mathbb{R}^d,\ (\nabla V(x) - \nabla V(y)) \cdot (x - y) \leq L\left\|x - y\right\|^2$$

and take $i \neq j$. Defining as in the previous lemma, but this time in a continuous setting, $e(t) = \frac{\left\|x_i^N(t) - x_j^N(t)\right\|^2}{2}$, one has, similarly to the gradient flow of $\mathcal{F}_N$,

$$\begin{aligned}
e'(t) &= \left((x_i^N)'(t) - (x_j^N)'(t)\right) \cdot \left(x_i^N(t) - x_j^N(t)\right) \\
&\geq \frac{1}{\varepsilon_N}\left[-\left\|x_i^N(t) - x_j^N(t)\right\|^2 + \left(b_i(X^N(t)) - b_j(X^N(t))\right) \cdot \left(x_i^N(t) - x_j^N(t)\right)\right] \\
&\quad - \left(\nabla V(x_i^N(t)) - \nabla V(x_j^N(t))\right) \cdot \left(x_i^N(t) - x_j^N(t)\right) \\
&\geq -2\left(\frac{1}{\varepsilon_N} + L\right) e(t)
\end{aligned}$$

since $\left(b_i(X^N(t)) - b_j(X^N(t))\right) \cdot \left(x_i^N(t) - x_j^N(t)\right)$, again, is non-negative. And, using this time a continuous reverse Gronwall's lemma,

$$\left\|x_i^N(t) - x_j^N(t)\right\|^2 \geq e^{-2K_N t}\eta$$

Taking the barycenters and applying Theorem 27, one obtains, for any $t \in [0;T]$,

$$W_2^2\left(\rho^N(t), \frac{1}{N}\sum_{i=1}^N \delta_{b_i(X^N(t))}\right) \leq \frac{C_{d,\Omega}e^{K_N(d-1)t}}{N\eta^{d-1}}. \quad \square$$

Choosing any ε_N such that

$$\int_0^T e^{\left(\frac{1}{\varepsilon_N}+L\right)(d-1)t}dt \leq CN\varepsilon_N\eta_N^{d-1}$$

will be suitable for Theorem 31, where C is a constant independant of N and η_N the smallest distance between two points of the support of a good N-point quantization measure of μ^0 (and we can take this measure for μ_N^0). One can set, for instance $\varepsilon_N = \frac{1}{\ln(N)+(d-1)\ln(\eta_N)} > 0$, since η_N is of order $N^{-\frac{1}{d}}$. These values are by no mean optimal to obtain convergence of the discretized gradient flow.

Remark 9. Sadly, these computations are difficult to adapt in the variational MFG setting. Indeed, in Section 2.2, we introduced a discrete measure on curves that minimized the energy:

$$J_\varepsilon(Q) := \int_\Gamma L(\gamma')dQ(\gamma) + \int_0^T F_\varepsilon(e_t\#Q)dt + G(Q)$$

over the space of uniform discrete measures on curves, $\mathcal{P}_N(\Gamma)$.

Assuming even that we are in a simple case where G is given by a $C_c^{1,1}$ potential V:

$$G(Q) = \int_0^T \int_\Gamma V(\gamma(t))dQ(\gamma)dt$$

Doing first order variations on the curves, $\gamma_i \to \gamma_i + h$ with $h \in \mathcal{C}_c^\infty([0;T],\mathbb{R}^d)$ tells us that if Q_N is a minimizer for $(\mathcal{M}_{N,\mu_N^0,\varepsilon_N})$, every $\gamma_i \in \mathrm{spt}(Q_N)$ verifies:

$$\gamma_i''(t) = \frac{\gamma_i(t) - b_i(\gamma_1(t),\dots,\gamma_N(t))}{N\varepsilon_N} + \nabla V(\gamma_i(t))$$

in the sense of distributions and therefore at almost every $t \in (0;T)$ since the left-hand side is $\mathbb{L}^2(\mathrm{d}t)$. Gronwall evaluations in the spirit of the two previous cases cannot be done as easily here. Indeed, we can notice that, if we define e as in the proof of Lemma 7 or Lemma 9, then, at least formally,

$$\begin{aligned} e''(t) =& 2((\gamma_i''(t) - \gamma_j''(t)) \cdot (\gamma_i(t) - \gamma_j(t)) + \left\|\gamma_i'(t) - \gamma_j'(t)\right\|^2) \\ =& \frac{\|\gamma_i(t) - \gamma_j(t)\|^2}{N\varepsilon_N} - \frac{(b_i(t) - b_j(t)) \cdot (\gamma_i(t) - \gamma_j(t))}{N\varepsilon_N} \\ &+ (\nabla V(\gamma_i(t)) - \nabla V(\gamma_j(t))) \cdot (\gamma_i(t) - \gamma_j(t)) + \left\|\gamma_i'(t) - \gamma_j'(t)\right\|^2 \end{aligned} \tag{3.31}$$

(we noted $b_i(t)$ the barycenter at time t: $b_i(\gamma_1(t), \dots, \gamma_N(t))$).

Notice that, in this case, the negative term $-(b_i(t) - b_j(t)) \cdot (\gamma_i(t) - \gamma_j(t))$ could be compensated by the positive one $\left\|\gamma_i'(t) - \gamma_j'(t)\right\|^2$ and we do not have enough information on either one. On the other hand, even if we had positive lower bounds for the distance between two curves of $\mathrm{spt}(Q_N)$, this would only give us a bound similar to (3.30) of Theorem 31, these do not translate easily into bounds on $J_{\varepsilon_N}(Q_N) - J(Q_{\min})$, that could be used for the proof of an upper bound result like that of Proposition 17.

Chapter 4

Lagrangian critical measures

4.1 Critical points for the quantization energy

We begin this chapter by recalling that for $N \in \mathbb{N}$, $Y \in (\mathbb{R}^d)^N$ is a critical point for F_N when every point y_i is the ρ-barycenter of its associated Laguerre cell:

$$y_i = N \int_{\mathrm{Lag}_i(Y)} x d\rho(x)$$

This condition is not sufficient most of the time for the point cloud to minimize F_N, with the notable exception of the one-dimensional case:

Proposition 32. *Let $\rho \in \mathcal{P}(\mathbb{R}) \cap \mathbb{L}^1(\mathbb{R})$. Then, $F_N : Y \mapsto W_2^2(\rho, \frac{1}{N}\sum_{i=1}^N \delta_{y_i})$ has a unique critical point, which is also its minimum.*

Proof. In dimension 1, the optimal transport between ρ and any $\mu \in \mathcal{P}_N(\mathbb{R})$ is given by the cumulative distribution function R of ρ, meaning that the Laguerre cells for this transport are the intervals:

$$\left[R^{-1}\left(\frac{i}{N}\right); R^{-1}\left(\frac{i+1}{N}\right)\right]$$

independently of the positions of the Dirac masses. This means that the only critical point for F_N is the one with each point at the ρ-barycenters of these intervals, and since F_N admits a minimizer, it is this critical point. □

Proposition 33. *Let $\rho \in \mathcal{P}(\mathbb{R}^d) \cap \mathbb{L}^1(\mathbb{R}^d)$ be the separable product of d real densities, $\rho = \rho_1 \otimes \cdots \otimes \rho_d$. For $N = n^d$, take d n-tuples $y^i = (y_1^i, \ldots, y_n^i)_{i=1\ldots d}$). Take $Y \in (\mathbb{R}^d)^N$ to be the cloud of points, aligned with the cartesian grid, with their i-th coordinate in $\{y_1^i, \ldots, y_n^i\}$ for every i.*

Then, the barycenters $B(Y)$ of the corresponding Laguerre cells are aligned on a parallel grid, and in fact, $B(Y)$ is a critical point for F_N.

Let us remark briefly that this is an example where the Lloyd algorithm of the previous chapter converges to a critical point in one step. It is also one where one can obtain the worst dependency in N of the quantization error (the value of F_N), meaning with the highest exponent $-1/d$ (see Section 3.3 of the previous chapter).

Proof. Let us write, in the spirit of our notations so far, $Y = (y_1, \dots, y_N)$, and $\mu^Y = \frac{1}{N}\sum_{k=1}^N \delta_{y_k}$. For $i = 1 \dots d$, take $\phi^i := (\phi^i_1, \dots, \phi^i_n)$ the optimal weights for the dual formulation of the optimal transport problem between ρ_i and the discrete measure $\frac{1}{n}\sum_{k=1}^n \delta_{y^i_k}$:

We wish to show that the optimal transport between ρ and μ^Y is given by the Laguerre cells: $\mathrm{Lag}_k(Y) := \mathrm{Lag}_k(Y, \Phi)$, where if $y_k = (y^1_{k_1}, \dots, y^d_{k_d})$, then $\phi_k = \phi^1_{k_1} + \dots + \phi^d_{k_d}$. To show this, we only need to prove the visual fact that these cells each contain the product of the corresponding 1-dimensional Laguerre cells:

$$\prod_{i=1}^d \mathrm{Lag}_k(y^i, \phi^i) \subset \mathrm{Lag}_k(Y, \Phi)$$

Indeed, the left-hand side of this inclusion has ρ-mass $1/N$, therefore, so does the right-hand side, as the total mass of ρ is 1 and this inclusion is an equality. But this inclusion is immediate from the definition of Φ and the separability of our cost (the squared euclidean norm):

If for $i = 1 \dots d$, $x^i \in \mathrm{Lag}_k(y^i, \phi^i)$ and $x = (x^1, \dots x^d)$, then for any $k'_1, \dots, k'_d$,

$$\begin{aligned} \|x - y_k\|^2 + \phi_k &= \sum_{i=1}^d \left\|x^i - y^i_{k_i}\right\|^2 + \phi^i_{k_i} \\ &\leq \sum_{i=1}^d \left\|x^i - y^i_{k'_i}\right\|^2 + \phi^i_{k'_i} \end{aligned} \tag{4.1}$$

and x is in $\mathrm{Lag}_k(Y, \Phi)$. Therefore each Laguerre cell $\mathrm{Lag}_k(Y, \Phi)$ has the correct mass $1/N$ and is optimal for the dual formulation of semi-discrete optimal transport.

Notice also that, from Proposition 32, these Laguerre cells do not depend on the actual coordinates of the Dirac masses, as long as these are aligned parallel to the main axes. Therefore the i-th coordinates of their ρ-barycenters are simply the barycenters, according to ρ_i of the 1-dimensional Laguerre cells $\mathrm{Lag}_k(y^i, \phi^i)$ and therefore, these barycenters are also aligned parallel to the main axes. □

Consider a sequence $(\mu^N)_{N\in\mathbb{N}}$ of discrete measures $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{y^N_i}$ supported on critical points $(Y^N)_{N\in\mathbb{N}}$ for F_N, with an increasing size N. Such a sequence is tight and therefore, applying Kolmogorov's theorem, one can assume

(up to considering a sub-sequence) that it narrowly converges towards a measure $\mu \in \mathcal{P}(\Omega)$. In the best case scenario, this measure is ρ, however, when the discrete measures are poorly chosen, such as those of Fig. 3.1 or Fig. 4.1 below, the limit measure can be singular. A consequence of Theorem 27 is that convergence to such a "bad" measure cannot happen, unless the minimum distance between two Dirac masses in the discrete measures μ^N goes to 0 faster than $N^{\frac{1}{d-1}}$. Indeed the points in this cloud are invariant through the Lloyd algorithm and therefore in this case, Theorem 27 gives a convergence rate of μ^N towards ρ for the 2-Wasserstein distance.

One can observe experimentally that the support of these narrow limits cannot be too irregular and in particular, seems to inherits the symmetry imposed by the optimality condition $Y = B(Y)$ in the discrete cases. This is the object of the next Section 4.2 and for now, let us cite some results on these limit measures, directly induced by the discrete settings.

In the remainder of this section, $(\mu^N)_{N\in\mathbb{N}}$ is a sequence of discrete measures supported on critical points of increasing size, as described above, and we assume that $(\mu^N)_{N\in\mathbb{N}}$ narrowly converges towards μ^∞ as $N \to \infty$. Furthermore, we assume that Ω is convex and that $\rho > 0$ almost everywhere on Ω. This last assumption allows us to assume that the measures μ^N are also supported inside $\mathrm{spt}(\rho)$, since they are supported on barycenters (and therefore so is μ^∞).

Proposition 34. *Assume that* $\mathrm{spt}(\mu^\infty)$ *has its connected components all connected by Lipschitz-continuous arcs (between any two points of the same component, one may draw a Lipshitz-continuous path). Then,*

$$\mu^\infty \in \mathrm{argmin}\{W_2^2(\rho,\mu) \mid \mu \in \mathcal{P}(\Omega),\ \mu^\infty(C) = \mu(C) \quad \textit{for any connected component } C \textit{ of } \mathrm{spt}(\mu^\infty)\}. \tag{4.2}$$

Proof. With these notation, let us fix N and consider $\Phi^N \in \mathbb{R}^N$ an optimal set of weights for the semi-discrete dual formulation of Optimal Transport (1.8). We first show that for any i, j, $|\phi_i^N - \phi_j^N| \leq \left\|y_i^N - y_j^N\right\|^2$: indeed, for such i, j, since y_j^N is inside its Laguerre cell, one has

$$\begin{aligned} -\phi_j^N &= \left\|y_j^N - y_j^N\right\|^2 - \phi_j^N \\ &\leq \mathrm{argmin}_{k=1\dots N} \left\|y_k^N - y_j^N\right\|^2 - \phi_k^N \\ &\leq \left\|y_i^N - y_j^N\right\|^2 - \phi_k^N \end{aligned}$$

The bound for the absolute value comes straightforwardly, exchanging i and j.

The sequence μ^N narrowly converges to μ^∞. On one hand, this tells us that every point of the support of μ^∞ is limit of a sequence of points in the supports of the measures μ^N:

$\forall y \in \mathrm{spt}(\mu^\infty),\ \exists (y^N)_{N\in\mathbb{N}} \in \Omega^{\mathbb{N}}$ s.t.

$$\forall N \in \mathbb{N},\ y^N \in \mathrm{spt}(\mu^N) \text{ and } \lim_{N\to\infty} y^N = y.$$

On the other hand, there exists a Kantorovich potential ϕ^∞ for the optimal transport from μ^∞ to ρ such that if $y^N_{i_N} \in \mathrm{spt}(\mu^N)$ converges to $y \in \mathrm{spt}(\mu^\infty)$ as $N \to \infty$, then $\phi^N_{i_N}$ converges to $\phi^\infty(y)$ (uniformly in y). One can see this by constructing Kantorovich potentials for the semi-discrete transport that are defined on the whole domain Ω with values ϕ^N_i at y^N_i (for every N, but, most importantly also for every $i = 1 \dots N$), see Remark 1 in Chapter 1. This sequence of potentials is still equicontinuous and equibounded, since they are c-concave, and we may assume, invoking Arzelà-Ascoli theorem, that a subsequence of it uniformly converges on Ω to a Kantorovich potential for the optimal transport between μ^∞ and ρ. For any $y, y' \in \mathrm{spt}(\mu^\infty)$, taking $y^N, y'^N \in \mathrm{spt}(\mu^N)$, such that $\lim_{N\to+\infty} y^N_{i_N} = y$, $\lim_{N\to+\infty} y^N_{j_N} = y'$, then for any N,

$$|\phi^N_{i_N} - \phi^N_{j_N}| \leq \left\|y^N_{i_N} - y^N_{j_N}\right\|^2$$

and taking the limit as $N \to \infty$, one has

$$|\phi^\infty(y) - \phi^\infty(y')| \leq \|y - y'\|^2 .$$

Now, assume that y and y' are connected, in the support of μ^∞, by a path with finite length $\gamma \in \mathcal{C}^0([0;1],\Omega)$ (i.e, a Lipschitz continuous curve, up to reparametrization). Taking for any $n \in \mathbb{N}$, a chain $y_0 = y = \gamma(0), y_1 = \gamma(1/n), \dots, y_n = y' = \gamma(1)$ on this path,

$$|\phi^\infty(y) - \phi^\infty(y')| = \sum_{i=0}^{n} \phi^\infty(y_i) - \phi^\infty(y_{i+1}) \leq L^2 \sum_{i=1}^{n} \frac{1}{n^2}$$

where L is the Lipschitz constant of the path, and letting $n \to \infty$, we arrive at

$$\phi^\infty(y) = \phi^\infty(y').$$

Now, take $\mu' \in \mathcal{P}(\Omega)$ such that for any connected component C of $\mathrm{spt}(\mu^\infty)$, $\mu'(C) = \mu^\infty(C)$. μ^∞ has finite mass and therefore, it only charges an at most countable number of connected (disjointed) sets $(C_1, C_2, \dots)$. Recall that any such component is also connected by Lipschitz paths and therefore, ϕ^∞ is constant on any of these components, $\phi^\infty \equiv \phi^\infty_i \in \mathbb{R}$ on C_i. Furthermore, ϕ^∞ is

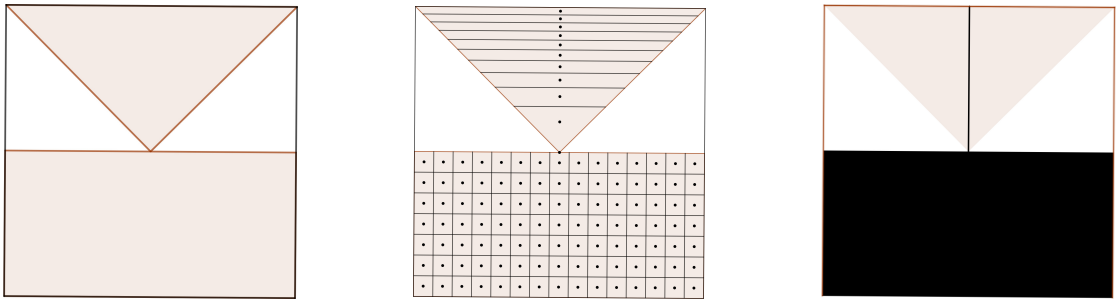


Figure 4.1: From left to right, the support of the probability density ρ (which is constant on this support), a critical point cloud for F_N which is not a minimizer and the limit measure obtained by adding points in a way that keeps the degeneracy in top part of the figure (see formula (3.1)). The limit measure μ^∞ is NOT uniform on the vertical segment (but it is on the lower rectangle due to Proposition 35).

lower bounded on $\mathrm{spt}(\mu^\infty)$, since it is continuous on this compact set, and, up to adding a constant to every Kantorovich potentials in the previous proof, we may assume that $\phi^\infty \geq 0$. Decomposing the integral of the cost along all connected components, we obtain,

$$\begin{aligned} W_2^2(\rho,\mu^\infty) &= \sum_i \int_{C_i} \phi^\infty \mathrm{d}\mu^\infty + \int_\Omega (\phi^\infty)^c(x)\mathrm{d}\rho(x) \\ &= \sum_i \phi_i^\infty \mu^\infty(C_i) + \int_\Omega (\phi^\infty)^c(x)\mathrm{d}\rho(x) \\ &= \sum_i \phi_i^\infty \mu'(C_i) + \int_\Omega (\phi^\infty)^c(x)\mathrm{d}\rho(x) \\ &= \int_\Omega \phi^\infty(y)\mathrm{d}\mu'(y) + \int_\Omega (\phi^\infty)^c(x)\mathrm{d}\rho(x) \leq W_2^2(\rho,\mu') \end{aligned} \tag{4.3}$$

and this exactly states that μ^∞ is a minimizer for the restricted Problem (4.2). □

On the other hand, the support of μ^∞ cannot be too large without forcing μ^∞ to actually be ρ on most of it:

Proposition 35. *Let us write $\mu^\infty = \mu^{\mathrm{ac}}\mathrm{d}x + \mu^{\mathrm{sing}}$ the decomposition of μ^∞ into absolutely continuous and singular part, such that $\mathrm{d}x \perp \mu^{\mathrm{sing}}$.*

Then, $\mu^{\mathrm{ac}}(x) = \rho(x)$ for $\mathrm{d}x$-almost every $x \in \mathrm{spt}(\mu^\infty)$. Furthermore, for any open set $A \subset \mathrm{spt}(\mu^\infty)$, $\mu^{\mathrm{sing}}(A) = 0$.

Proof. Let for any N, u_N be a Brenier potential for the transport from ρ to μ^N, i.e. a convex function on $\mathbb{R}^d$ such that the optimal transport map from ρ to μ^N is

∇u_N. In fact, we may assume $u_N(x) = \frac{\|x\|^2}{2} - \phi_N(x)$ ρ-almost everywhere, where ϕ_N is a c-concave Kantorovich potential for the same transport. This means that, as $N \to \infty$ and up to considering instead a subsequence, ϕ_N (respectively u_N) uniformly converges, on Ω, towards ϕ_∞ a Kantorovich potential (resp u_∞ a Brenier potential) for the transport from ρ to μ^∞.

Now, for any N, and $y \in \mathrm{spt}(\mu^N)$, $y \in \partial u_N^*(y)$, since these points are barycenters of the Laguerre cells $\partial u_N^*(y)$. Therefore, since every point in $\mathrm{spt}(\mu^\infty)$ is a limit (as $N \to \infty$) of points in $\mathrm{spt}(\mu^N)$, for any $y \in \mathrm{spt}(\mu^\infty)$, $y \in \partial u_\infty^*(y)$ and for $\mathrm{d}x$-almost every $y \in \mathrm{spt}(\mu^\infty)$, the subgradient is a singleton and $y = \nabla u_\infty^*(y)$. By duality, this immediately tells us that $\mathrm{d}x$-almost every $y \in \mathrm{spt}(\mu^\infty) \cap \mathrm{spt}(\rho)$ satisfies $y \in \nabla u_\infty(y)$ and making the same reasoning with u_∞ instead of u_∞^*, we get that the optimal transport between μ^∞ and ρ is given by the identity map in both direction, $\mathrm{d}x$-almost everywhere on $\mathrm{spt}(\mu_\infty) \subset \mathrm{spt}(\rho)$, hence the equality $\mu^{\mathrm{ac}}(x) = \rho(x)$.

On the other hand, take $x \in \mathrm{spt}(\mu^\infty)$ such that there exists a radius $r > 0$ with $B(x, r) \subset \mathrm{spt}(\mu^\infty)$. Then, as we mentioned earlier, for $\mathrm{d}x$-almost every y in this ball, $\nabla u_\infty(y)$ and $\nabla u_\infty^*(y)$ are well-defined and equal to y. Since u_∞ is a convex function, this implies that u_∞ (and therefore also u_∞^*) is equal to $\frac{\|.\|}{2}$ on this ball, and optimal transport between ρ and μ_∞ is given in both sense by the identity map on this ball (therefore, μ^{sing} does not charge any point in this set). □

The reader will notice that in this last result, we do not use the fact that μ^∞ was a limit of discrete measure, other than to guarantee that the optimal transport plan between μ^∞ and ρ charges the pair (x, x) for μ^∞-almost every $x \in \Omega$ (in a very informal way). Section 4.2 tries to formalize this assumption to prove that any measure that satisfies it has to be supported on a smooth structure. This is a very natural expectation, considering Theorem 27 and the fact that the Lloyd algorithm leaves stable certain lower dimensional subspaces (see Remark 7, second point) and it would allow us to understand better these "nasty" equilibrium situations and, in particular, how to avoid entering their zone of attraction when using this algorithm.

4.2 Lagrangian critical measures for W_2^2

A first order of business is defining a notion of critical point which is not dependent on being supported on a finite (or even countable) point cloud. Fortunately, the disintegration theorem applied to μ (a potential critical point) and ρ on one side, and the optimal transport plan γ between the two, allows for a very natural disintegration of ρ along μ. The fibers of this disintegration (see Theorem 42,

Appendix A) gives us an equivalent of the notion of Laguerre cells, in this continuous setting:

Proposition 36 (Disintegration along the transport plan)**.** *Let $\mu \in \mathcal{P}(\Omega)$ and T be the optimal transport map from ρ to μ. Then, there exists a family of measures $\{\rho_y\}_{y\in\Omega} \subset \mathcal{P}(\Omega)$ uniquely defined at least at μ-almost every y, such that*

- *For any Borel measurable set $B \subset \Omega$, $y \mapsto \rho_y(B)$ is a Borel measurable function.*
- *For μ-almost every y, ρ_y is supported on the fiber $T^{-1}(y)$.*
- *For every Borel-measurable function $f : \Omega \to \mathbb{R}^+ \cup \infty$:*

$$\int_\Omega f(x)\mathrm{d}\rho(x) = \int_\Omega \int_\Omega f(x)\mathrm{d}\rho_y(x)\mathrm{d}\mu(y)$$

or equivalently, for any Borel set A,

$$\rho(A) = \int_\Omega \rho_y(A)\mathrm{d}\mu(y).$$

Although this result can be demonstrated by applying the disintegration theorem 42 to μ and $\rho = T\#\mu$ for an optimal transport map T, the existence of such a disintegration for the transport plan γ instead of ρ can be proven without assuming the existence of T. We therefore prove Proposition 36 as a consequence in this case of the more general disintegration result. In particular, the disintegration of γ along μ is still valid even when it is not given by an optimal transport map T (however Proposition 36 would need to be written differently in that case).

Proof. With the notations of the proposition, let us recall that $\pi_2\#\gamma = \mu$ where π_2 is the projection of $\Omega \times \Omega$ on its second component.

Applying Theorem 42, we can disintegrate γ along μ, giving us the family of measures which we write $(\gamma_y)_{y\in\Omega} \in \mathcal{P}(\Omega^2)^\Omega$.

Then, according to the theorem, for μ-almost every $y \in \Omega$, γ_y is supported in $\pi_2^{-1}(y) \cap \mathrm{spt}(\gamma)$ and therefore, $\rho_y = \pi_1\#\gamma_y$ is supported on $T^{-1}(y)$.

Finally, for any Borel set $A \subset \Omega$,

$$\int_A \rho(x)\mathrm{d}x = \int_{A\times\Omega} \mathrm{d}\gamma = \int_\Omega \gamma_y(A \times \{y\})\mathrm{d}\mu(y) = \int_\Omega \rho_y(A)\mathrm{d}\mu(y)$$

(The computations are almost identical for the integral of a Borel map f on Ω). □

The measure ρ_y corresponds to the law of $X \sim \rho$, conditional to $Y = y$, when (X, Y) follows the law of the optimal transport plan between ρ and μ.

Note also that in the case when μ is discrete supported on $y_1, \dots, y_N$, ρ_y is simply $\rho \llcorner \mathrm{Lag}_i(Y)$, so that $\mathrm{spt}(\rho_y)$ can be seen as the equivalent of a Laguerre cell in the continuous limit case $N \to \infty$.

Equipped with this notion, we can now reformulate the condition of being the barycenter of one's Laguerre cell as a property which passes well to the narrow limit:

Definition 8. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $(\rho_y)_{y \in \mathbb{R}^d}$ be the disintegration of ρ along the optimal transport plan between ρ and μ mentioned in Proposition 36.*

We say that μ is Lagrangian critical for $W_2^2(., \rho)$ when

$$\text{for } \mu\text{-almost every } y \in \mathbb{R}^d,\ y = \int_\Omega x \mathrm{d}\rho_y(x).$$

In the remainder of this section, μ will always denote a Lagrangian critical (probability) measure for $W_2^2(., \rho)$.

As we are staying focused on the problem of optimal quantization (3.1), we will omit in the future the specification "for $W_2^2(., \rho)$" and only say that μ is Lagrangian critical when it verifies Definition 8. Furthermore, leaving for a moment the formalism of the disintegration theorem, we can give a more concise formulation of Definition 8:

Lemma 10. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ and γ be the optimal transport plan from ρ to μ.*

Then, μ is Lagrangian critical if and only if, for any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$:

$$\int_{\Omega \times \Omega} (y - x) \cdot \xi(y) \mathrm{d}\gamma(x, y) = 0$$

Proof. From Brenier's theorem, we know that the optimal transport plan from ρ to μ is given by a map $T : \Omega \to \Omega$,

$$\gamma = (Id, T)\#\rho.$$

Therefore, one can claim the following equivalent statements:

For μ-almost every $y \in \Omega$,

$$y = \int_\Omega x \mathrm{d}\rho_y(x)$$

$\iff$ For any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_\Omega (y - \int_\Omega x \mathrm{d}\rho_y(x)) \cdot \xi(y) \mathrm{d}\mu(y) = 0$$

$\iff$ For any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_{\Omega\times\Omega} (y - x) \cdot \xi(y) \mathrm{d}\rho_y(x) \mathrm{d}\mu(y) = 0$$

$\iff$ For any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_{\Omega\times\Omega} (T(x) - x) \cdot \xi(T(x)) \mathrm{d}\rho_y(x) \mathrm{d}\mu(y) = 0$$

$\iff$ For any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_{\Omega\times\Omega} (T(x) - x) \cdot \xi(T(x)) \mathrm{d}\rho(x) = 0$$

$\iff$ For any $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_{\Omega\times\Omega} (y - x) \cdot \xi(y) \mathrm{d}\gamma(x, y) = 0$$

□

Let us quickly mention an interpretation of our notion of critical measure in term of internal perturbations $\mu \mapsto (Id + \xi)\#\mu$ for a smooth map ξ:

Considering first outer perturbations of μ, $\mu \mapsto \mu + \chi$, for a small $\chi \in \mathbb{L}^1(\Omega)$, $\int_\Omega \mathrm{d}\chi = 0$, being a critical measure for these variations, namely

$$\frac{d}{d\varepsilon} W_2^2(\mu + \varepsilon\chi, \rho)\Big|_{\varepsilon=0} = 0$$

happens only at $\chi = 0$. Indeed, $\mu \mapsto W_2^2(\rho, \mu)$ is strictly convex since $\rho \ll \mathrm{d}x$ and therefore, there is only one critical measure for the outer variations, ρ (this is proposition 7.19, and in fact the whole section 7.2.2 of [63]).

On the other hand, for $\xi \in \mathcal{C}_b^0(\Omega, \mathbb{R}^d)$ and $\varepsilon > 0$ small enough (such that for all $x \in \Omega$, $x + \varepsilon\xi(x) \in \Omega + B(0; 1)$),

$$\begin{aligned} W_2^2((Id + \varepsilon\xi)\#\mu, \rho) - W_2^2(\mu, \rho) &\leq \int_\Omega \|T(x) + \varepsilon\xi(x) - x\|^2 \\ &\qquad - \|T(x) - x\|^2 \mathrm{d}\rho(x) \\ &\leq 2\varepsilon \int_\Omega (T(x) - x) \cdot \xi(x) \mathrm{d}\rho(x) \\ &\qquad + \varepsilon^2 \int_\Omega \|\xi(x)\|^2 \mathrm{d}\rho(x) \\ &\leq 2\varepsilon \int_\Omega (y - x) \cdot \xi(x) \mathrm{d}\gamma(x, y) + O_{\varepsilon\to 0}(\varepsilon^2) \end{aligned}$$

where T is the optimal transport map from ρ to μ and $\gamma = (Id, T)\#\rho$ is the optimal transport plan. Using $\mu_\varepsilon = (Id + \varepsilon\xi)\#\mu$ instead of μ, we obtain symmetrically,

$$W_2^2(\mu_\varepsilon, \rho) - W_2^2(\mu, \rho) \geq 2\varepsilon \int_{(\Omega + B(0;1))^2} (y - x) \cdot \xi(x) \mathrm{d}\gamma_\varepsilon(x, y) - O_{\varepsilon \to 0}(\varepsilon^2)$$

with γ_ε the unique optimal transport plan from ρ to μ_ε.

The (also unique) optimal transport plan $\gamma = (Id, T)\#\rho$ from ρ to μ is the narrow limit of γ_ε as $\varepsilon \to 0$ (see [63], theorem 1.50). Taking this limit in the equality above, we have that the derivative of $\mu \mapsto W_2^2(\mu, \rho)$ with respect to inner variations in the direction $\xi \in \mathcal{C}_b^0(\Omega)$ is

$$2 \int_{\Omega \times \Omega} (y - x) \cdot \xi(y) \mathrm{d}\gamma(x, y).$$

It is 0 for every ξ iff μ is Lagrangian critical.

Lemma 10 gives us the stability of the notion of Lagrangian critical measure, up to narrow limits. The argument is once again the same stability, but for transport plans:

Proposition 37. *For any $N \in \mathbb{N}$, a discrete measure $\mu = \frac{1}{N} \sum_{i=1}^N \delta_{y_i}$ such that $Y = (y_1, \ldots, y_N)$ is a critical point for F_N, is Lagrangian critical.*

Furthermore, any narrow limit of Lagrangian critical measures is itself Lagrangian critical.

Proof. As we already mentioned, for any $i = 1 \ldots N$, $\rho_{y_i} = \rho \llcorner \mathrm{Lag}_i(Y)$ with the formalism of Theorem 42. The first part of the proposition is then immediate since each point y_i is the ρ-barycenter of its Laguerre cell.

For the second part, take a sequence $(\mu_n)_{n \in \mathbb{N}}$ of Lagrangian critical measures (the n is not related to the size of the support of μ_n in this case) and assume it converges narrowly towards $\mu \in \mathcal{P}(\Omega)$. For $n \in \mathbb{N}$, we write γ_n the optimal transport plan from ρ to μ_n. Then $(\gamma_n)_{n \in \mathbb{N}}$ is tight since $(\mu_n)_{n \in \mathbb{N}}$ is, and (up to a sub-sequence) it narrowly converges towards the optimal transport plan γ from ρ to μ (this is theorem 1.50 of [63]). The conclusion comes from taking the narrow limit in Lemma 10: for $\xi \in \mathcal{C}^0(\Omega, \mathbb{R}^d)$,

$$\int_{\Omega \times \Omega} (y - x) \xi(y) \mathrm{d}\gamma(x, y) = \lim_{n \to \infty} \int_{\Omega \times \Omega} (y - x) \xi(y) \mathrm{d}\gamma_n(x, y) = 0 \quad \square$$

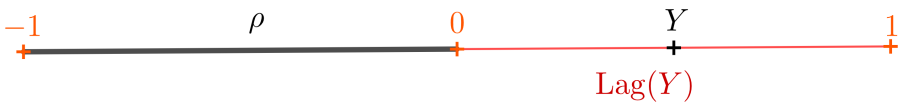


Figure 4.2: The support (in black) of a Lagrangian critical measure which cannot be obtained as a narrow limit of uniform discrete Lagrangian critical measures.

Let us note here that, with Definition 8, not every Lagrangian critical measure is limit of measures supported uniformly on a critical point for F_N. The simplest counter-example to this comes from the 1-dimensional case, $\Omega \subset \mathbb{R}$. In this case, the only Lagrangian critical measure which is such a limit is ρ, since there is only one critical point for F_N, for any N, and that critical point defines a discrete measure narrowly converging towards ρ, according to the convergence rates of (3.6). On the other hand, on the set $\Omega = [-1; 1]$, take

$$\rho(x) = \frac{1}{2}\mathbb{1}_{[-1;1]} \quad \text{and} \quad \mu = \rho\mathbb{1}_{[-1;0]}\mathrm{d}x + \frac{1}{2}\delta_{1/2},$$

(see Fig. 4.2), then μ is Lagrangian critical (but it is not ρ). Indeed, for any $-1 \leq y \leq 0$, $\rho_y = \delta_y$ and $\rho_{1/2} = \mathbb{1}_{[0;1]}$ and μ-almost every point $y \in \Omega$ is the ρ_y barycenter of Ω. On the other hand, Proposition 35 is also true for any Lagrangian critical measure, since, as we mentioned, we only use the fact that every point is a barycenter of points in the subgradient of a Brenier potential.

Restricted, as it is, between two parts of the support of ρ_y, we expect the support of a Lagrangian critical measure μ to be aligned on very regular curves. This is backed up by experiments, for instance, the limit measure of the discrete ones represented on Fig. 3.1 in the case of $\rho = \mathrm{d}x$ are supported on a finite union of segments. We can also observe on Fig. 4.1 that the (apparently) smooth submanifolds containing the support can be of several dimensions (dimensions 1 and 2 on the figure).

The following lemma gives a first mathematical statement of these observations. It states that, from μ-almost every point $y \in \Omega$, one can define a thin "trumpet" on each side of this point, such that locally, the support of μ is confined to these thin sets:

Lemma 11. *Let $\mu \in \mathcal{P}(\Omega)$ be Lagrangian-critical. Then, for μ-almost every $y \in \Omega$, there exists $C_y > 0$, $n \in \{0 \dots d\}$ and an n-dimensional linear subspace Π_y of $\mathbb{R}^d$ such that for any $w \in \Pi_y$, $\|w\| = 1$ and $z \in \mathrm{spt}\mu$,*

$$|w \cdot (y - z)| \leq C_y \|y - z\|^2$$

Proof. For μ-almost every $y \in \mathbb{R}^d$, we take $\Pi_y = \mathrm{span}\{x - y, x \in \mathrm{spt}(\rho_y)\}$ and $n = \dim(\Pi_y)$. Let us first notice that our statement is trivial (and almost empty) when $n = 0$ so that we may assume that there exists $z \neq y$ in $\mathrm{spt}(\rho_y)$.

Now, c-cyclical monotonicity states that for any $x \in \mathrm{spt}(\rho_y)$, and $z \in \mathrm{spt}(\mu)$, z cannot be in the ball of diameter $[x; y]$, otherwise said:

$$(x-y)\cdot(z-y) \leq \|y-z\|^2$$

We can find $N > n$ nonzero vectors $x_i - y$ such that for any $w \in \Pi_y$, $\|w\| = 1$, there exists i with $w.(x_i - y) > 0$. Indeed, consider the covering of the compact set $S(0,1) \cap \Pi_y$, with the open sets

$$U_x = \{w \in S(0,1) \cap \Pi_y,\ w \cdot (x-y) > 0\}$$

for $x \in \mathrm{spt}(\rho_y)$. It is a covering since y is the barycenter of Π_y with respect to ρ_y. One can extract from it a finite covering, which gives us $x_1, \dots x_N$ such that for any $w \in S(0,1) \cap \Pi_y$, $w.(x_i - y) > 0$. Notice then that the orthogonal in Π_y to $\mathrm{span}((x_i - y)_i)$ is $\{0\}$ and that family generates Π_y (and is not linearly independant hence $N > n$).

Now set C'_y to be the min positive correlation between any unit vector in Π_y and one of our reference vectors normalized $(x_i - y)/\|x_i - y\|$:

$$C'_y = \min_w \max_i \left\{ w \cdot \frac{(x_i - y)}{\|x_i - y\|},\ w \in \Pi_y,\ \ \|w\| = 1,\ i = 1 \dots N \right\}$$

Compactness ensures that C'_y exists, is attained and is strictly positive.

Now, for any $z \in \mathrm{spt}(\mu)$, and $w \in \Pi_y$, $\|w\| = 1$,

$$\begin{aligned} |w \cdot (z-y)| \leq & \left\| P_{\Pi_y}(z-y) \right\| \\ & \leq \max_i \frac{1}{C'_y \|x_i - y\|} (z-y)\cdot(x_i - y) \\ & \leq C_y \|z-y\|^2 \end{aligned} \tag{4.4}$$

where the constant only depends on y and $\mathrm{spt}(\rho_y)$. □

Let us illustrate here Lemma 11 a little: Considering the support of the limit measure μ_∞ represented on Fig. 4.1, one could think that this is the possible support for a Lagrangian critical measure associated with the uniform density on the whole square (instead of one supported only the non-convex set in light brown on the figure). However, this is in fact forbidden by the lemma. Indeed, points (of the support of ρ) in the sub-diagonal

$$F_< = \{(x,y) \mid x \in [0; 1/2],\ 1/2 < y < 1 - x\}$$

have to be sent by optimal transport to the border set

$$F_= = \{(x, 1/2) \mid x \in [0; 1/2]\}$$

(via an orthogonal projection). Taking such $x \in F_<$ and $y \in F_=$ its orthogonal projection, the vector $w = (x-y)/\|x-y\|$ is in Π_y. But, taking $z_\varepsilon = y - \varepsilon w$, z_ε should be in $\mathrm{spt}(\mu_\infty)$ for any $\varepsilon > 0$ small enough. This contradicts the lemma since $|w \cdot (z_\varepsilon - y)| = \varepsilon$ dominates $\|z_\varepsilon - y\|^2 = \varepsilon^2$ as $\varepsilon \to 0$. Therefore, Lagrangian critical measures associated with the Lebesgue measure on $\Omega = [0;1] \times [0;1]$ cannot have this support.

A direct corollary is that the support of μ is included in a countable union of Lipshitz manifolds. We note that this result is much weaker than what we are after, since Lipshitz-regularity of the support seems to be the least we can expect, looking at the supports of the Lagrangian critical measures in the previous pictures. On the other hand, it does give a lower-bound regarding the regularity we can expect on this set. Finally, this corollary also gives no structure (at least not directly) to the actual measure μ, allowing the support, even in these Lipschitz-continuous submanifolds to be very degenerate:

Corollary 37.1. *Let μ be a Lagrangian critical measure and denote*

$$\mathcal{E}_n := \{y \in \Omega \mid \Pi_y \textit{ exists and } \dim(\Pi_y) = d-n\}$$

Then, with the definition above, $\mathcal{E}_n$ is n-rectifiable for any $1 \leq n \leq d$, and in fact, there exists a countable family of Lipschitz-continuous maps $f_i : \mathbb{R}^n \to \mathbb{R}^d,\ i \in \mathbb{N}$ such that

$$\mathcal{E}_n \subset \bigcup_{i=0}^{\infty} f_i(\mathbb{R}^n)$$

Proof. Let $y \in \mathcal{E}_n$ such that $y \in \int_\Omega x \mathrm{d}\rho_y(x)$ and y verifies Lemma 11 (μ-almost any point of $\mathcal{E}_n$ does so). Let us fix for a time $\Pi = \Pi_y$ and $C = C_y$. Consider the tube orthogonal to Π and of width $\eta > 0$, inside Ω:

$$\Omega^y_{i_1,\dots,i_{d-n}} = \left(\bigoplus_{j=1}^{d-n} [i_j\eta; (i_j+1)\eta]\, e_j \oplus \bigoplus_{j=d-n+1}^{d} \mathbb{R} e_j \right) \cap \Omega$$

with $e_1, \dots e_{d-n}$ an orthogonal basis of Π, completed into an orthogonal basis of $\mathbb{R}^d$ by $e_{d-n+1}, \dots, e_d$, and $i_1, \dots, i_{d-n}$ chosen in $\mathbb{Z}$ such that $y \in \Omega^y_{i_1,\dots,i_{d-n}}$.

Then Lemma 11 implies that there exists a constant $C' = \sqrt{C\eta}$ such that, for any $w \in \Pi$, $\|w\| = 1$ and $z \in \mathrm{spt}(\mu) \cap \Omega^y_{i_1,\dots,i_{d-n}}$:

$$|w \cdot (y-z)| < C' \|y-z\|$$

In particular taking $\eta < \frac{1}{C}$ there cannot be any point of $\mathrm{spt}(\mu)$ inside Π and $\Omega^y_{i_1,\dots,i_{d-n}}$ (except for y), since such a point z would verify

$$\sup|w.(y-z)| = \|y-z\| < \|y-z\| \tag{4.5}$$

where the supremum is taken among unit vectors in Π. On the other hand, for any $z \in \mathrm{spt}(\mu) \cap \Omega^y_{i_1,\dots,i_{d-n}}$,

$$\|P_{\Pi^\perp}(z) - P_{\Pi^\perp}(y)\| \geq \sqrt{1-C'}\,\|z-y\| \tag{4.6}$$

(provided we took there again η such that $C' < 1$).

As is, Π and C' depend heavily on y, however, this last condition (4.6) is open in term of choosing the affine sub-space Π and the constant C. Therefore, up to shifting a little the inclination of Π_y and the value of C_y, and forgetting their actual definition from Lemma 11, we may assume that there exists an at most countable amount of such n-dimensional affine hyperplanes Π_y and constant C_y when one goes through μ-almost every point of $y \in \Omega$ (for instance, by approximating their parameters with rationals). But then, for such a subspace Π and constant C, the set $\mathcal{E}_{n,C',\Pi}$ of points y associated with this subspace and the constant $C' = \sqrt{C\eta}$, for $\eta < \frac{1}{C}$ rational is included in a finite union of graphs of lipshitz functions on $\mathbb{R}^n$. To see this, we only have to show that every "tube" $\mathcal{E}_{n,C',\Pi} \cap \Omega^y_{i_1,\dots,i_{d-n}}$ is included in the graph of such a Lipschitz function (since these sets are in countable amount). But all these points are associated with the same hyperplane Π and we have established in this case (in (4.5)) that the function:

$$f : x \mapsto y \in \mathcal{E}_{n,C',\Pi} \cap \Omega^\eta_{i_1,\dots,i_{d-n}} \text{ such that } P_{\Pi^\perp}(y) = x$$

is well defined on a subset of $\Pi_y^\perp$ (namely the orthogonal projection of $\mathcal{E}_{n,C',\Pi} \cap \Omega^y_{i_1,\dots,i_{d-n}}$ onto that sub-space) and Lipschitz-continuous with a constant at most $\sqrt{1-C'} > 0$ (this, on the other hand, was (4.6)). Up to extending this function to the whole of $\mathbb{R}^n$, we have the inclusion claimed in Corollary 37.1. □

Lemma 11 states that any point of Π_y at a distance less than $\frac{1}{2C_y}$ from y, has a unique projection (the point y itself) onto $\mathrm{spt}(\mu)$. This is very close to the notion of *reach positiveness* for the set $\mathrm{spt}(\mu)$:

A set $A \subset \Omega$ is said to have *positive reach*, $\varepsilon > 0$, when for any $x \in \mathbb{R}^d$, if the distance $d(x, A) < \varepsilon$, then x has a unique projection on A :

$$\mathrm{argmin}\{d(x,y) \mid y \in A\} \text{ is a singleton.}$$

The positive reach property would imply $C^{1,1}$ regularity of the different parts of this support, seen as manifolds (this is, for instance, (C) of [59]). However, we point out that, in the case of our Lagrangian critical measures, the reach positiveness property is only verified on the affine sub-space Π_y, which would not be enough.

Recall however the fact that the optimal transport from ρ to μ is given (ρ-almost everywhere) by the gradient of a convex function u:

$$\gamma = (Id, \nabla u)\#\rho$$

This implies that the optimal transport plan γ is supported on the "graph" of the subgradient of u^*:

$$\mathrm{spt}(\gamma) \subset \bigcup_{x\in\Omega} \partial u^*(x) \times x. \tag{4.7}$$

It might be obvious now, but the dimension of the support of μ near a point x will be the complementary of the dimension of $\partial u^*(x)$. Note that this last set is a convex one, and therefore, its dimension is simply the geometrical dimension of the affine space generated by this set. The truth of this claim, as well as the regularity that one can expect for the different parts of $\mathrm{spt}(\mu)$, are suggested by a remarkable theorem by Alberti (see Theorem 44 Appendix B). This result gives strong rectifiability to the set of points sharing the same dimension of subgradient for a given convex function.

For $0 \leq n \leq d$, denote $\mathcal{E}_{n,C}$ the set of all points in Ω verifying Lemma 11 with $C_y = C$ and $\dim(\Pi_y) = n$, such that

$$\mathcal{E}_n = \bigcup_C \mathcal{E}_{n,C}. \tag{4.8}$$

Notice here, that we can take the possible constants $(C_y)_{y\in\Omega}$ to be chosen in $\mathbb{Q}$ and therefore, (4.8) describes $\mathcal{E}_n$ as a countable (at most) union of sets which we expect to be $\mathcal{C}^2$ smooth, of dimension n (following Alberti's theorem).

From the proof of Lemma 11, we can claim that

$$\mathcal{E}_n \subset \{y \in \Omega \mid \dim(\partial u^*(y) \geq d - n\}. \tag{4.9}$$

Indeed, for $y \in \mathcal{E}_n$, μ-almost surely, the support of ρ_y generates the $d-n$ linear space Π_y, and, from inclusion (4.7), we can assume (again, μ-almost surely) $\mathrm{spt}(\rho_y) \subset \partial u^*(y)$ and therefore, this subgradient is of dimension at least $d-n$.

Alberti's theorem tells us that the right-hand side of inclusion (4.9) is contained in a $\mathcal{C}^2$, n-dimensional manifold, up to an $\mathcal{H}^n$-negligible set. If we were to show that $\mu \ll \mathcal{H}^n$ on the set in the left-hand side of the inclusion, we would get that the support of μ is included in a countable union of $\mathcal{C}^2$-submanifolds and that μ behaves "regularly" on these manifolds. Looking back at the examples of support of Lagrangian critical measures, this seems very likely that μ be absolutely continuous with respect to this n-dimensional Haussdorf measure (and one can check that this is indeed the case in these examples).

Proposition 38. *Assume that $y \mapsto \Pi_y$ is Lipshitz continuous on $\mathcal{E}_{n,C}$ for some $n \in \{1 \dots d\}$ and $C \in \mathbb{Q}$.*

Then, μ is absolutely continuous with respect to the n-dimensional Hausdorff measure $\mathcal{H}^n$ on $\mathcal{E}_{n,C}$. Furthermore, $\mathcal{E}_{n,C}$ is included in the countable union of $\mathcal{C}^2$ manifolds of dimension n, up to a μ-negligible set. If this is true for all $C \in \mathbb{Q}$, so is $\mathcal{E}_n$.

Proof. Let $C \in \mathbb{Q}$ and $A \subset \mathcal{E}_{n,C}$ be negligible for the Hausdorff measure with dimension $d-n$. We wish to prove that $\mu(A) = 0$. By construction of Π_y, it is sufficient to show that $B = \cup_{x \in A}(x + \Pi_x) \cap B(0,R)$ is Lebesgue-negligible for any $R > 0$ since all the μ-mass on A comes from ρ-mass on B for a large enough R.

Since A is negligible for $\mathcal{H}^n$, for any $\varepsilon > 0$, one can find a finite number of balls $B(x_i, r_i)$ such that $A \subset \bigcup_i B(x_i, r_i)$ and $\sum_i r_i^n \leq \varepsilon$. Denoting

$$A_i = A \cap B(x_i, r_i),$$

let us fix for a while $x \in A_i$. Then, for any $w \in \Pi_x$, the orthogonal projection v of w onto P_{x_i} verifies $|v| \leq R$ and

$$|w - v| \leq R d(\Pi_x, \Pi_{x_i}) \leq RLr_i$$

where L is the Lipshitz constant of $y \mapsto \Pi_y$ on $\mathcal{E}_{n,C}$. But then $\Pi_x \cap B(0,R)$ is contained in $(\Pi_{x_i} \cap B(0,R)) + B(0, RLr_i)$.

The sets $B_i = \cup_{x \in A_i}(x + \Pi_x) \cap B(0,R)$ cover B and we have just shown that for every i,

$$\begin{aligned} B_i \subseteq & B(x_i, r_i) + (\Pi_{x_i} \cap B(0,R)) + B(0, RLr_i) \\ = & (\Pi_{x_i} \cap B(0,R)) + B(0, (1+RL)r_i) \end{aligned} \tag{4.10}$$

Π_{x_i} is of dimension n, therefore, $\mathcal{L}(B_i) \lesssim r_i^{d-n}$, as it is contained in the sum of a n dimensional ball with a cube of length $(2 + 2RL)r_i$. But then,

$$\mathcal{L}(B) \leq \sum_i \mathcal{L}(B_i) \lesssim \sum_i r_i^{d-n} \leq \varepsilon,$$

and B is Lebesgue-negligible, therefore, A is μ-negligible and μ is absolutely continuous with respect to $\mathcal{H}^{d-n}$ on the set $\mathcal{E}_{n,C}$.

Now, to conclude, we simply notice that, $\mathcal{E}_{n,C}$ is included in the set of points in the subgradient of a convex function which are of dimension at least $k-n$. This function is in fact a Brenier potential ϕ associated with the transport from μ to ρ. Alberti's theorem tells us that the set $\{y \in \Omega,\ \dim(\partial\phi(y)) = n\}$ is included in a $\mathcal{C}^2$ submanifold of dimension $d-n$, up to removing a $\mathcal{H}^{d-n}$-negligible set. Therefore, the same is true for $\mathcal{E}_{n,C}$ up to removing a μ-negligible set. □

Proposition 39. *$y \mapsto \Pi_y$ is Lipshitz-continuous on $\mathcal{E}_{1,C}$ (resp $\mathcal{E}_{0,C}$, $\mathcal{E}_{d,C}$). In particular, $\mu \ll \mathcal{H}^1$ on this set (resp $\mu \ll \mathcal{H}^0$, $\mu \ll \mathcal{H}^d$) and $\mathcal{E}_1$ (resp $\mathcal{E}_0$, $\mathcal{E}_d$) is contained μ-almost everywhere in an at most countable union of C^2 paths (resp countable subset of $\mathbb{R}^d$, countable union of d-dimensional sub-manifolds).*

Proof. We start by proving the absolute continuity of μ on $\mathcal{E}_{0,C}$ as it ends up not needing the Lipschitz continuity. Indeed, we claim that the points in this set are strongly isolated (even from the whole support of μ):

Take $y \in \mathcal{E}_{0,C}$ for some $C > 0$, one has that for any $w \in S^{d-1}(0,1)$, $z \in \mathrm{spt}(\mu)$ and $z \neq y$:

$$|w \cdot (y - z)| \leq C \|y - z\|^2$$

For a vector w colinear to $y - z$ this yields

$$\|y - z\| \geq \frac{1}{C}$$

and there exists a ball around y which contains no other point of $\mathrm{spt}(\mu)$. Since $\mathcal{E}_{0,C}$ is a set of isolated points, it is at most countable and $\mu \ll \mathcal{H}^0$ on this set. The Lipschitz continuity of $y \mapsto \Pi_y$ is trivial since the lower bound on the radius of the ball does not depend on $y \in \mathcal{E}_{0,C}$.

The optimal transport plan between ρ and μ on $\mathcal{E}_{d,C}$ is given by the identity map in both directions, since $\partial u^*(y)$ is a singleton for any y in this set. This implies that $\mu = \rho$ on this set (and in particular, $\mu \ll \mathcal{H}^d$). Once again, the Lipschitz continuity condition on $y \mapsto \Pi_y$ is empty, since Π_y is $\{0\}$ for any $y \in \mathcal{E}_{d,C}$.

Finally, in the case $d = 1$, we use Proposition 38 and we have to show the Lipschitz continuity. It is a consequence of Lemma 11, since $\dim(\Pi_y) = d - 1$ μ-almost everywhere on $\mathcal{E}_{1,C}$.

Indeed take $y, z \in \mathcal{E}_{1,C}$. Then, for any $u \in B_{\mathbb{R}^d}(0,1)$,

$$u.(y - z) = P_{\Pi_y^\perp}(y - z).u + P_{\Pi_y}(y - z).u \tag{4.11}$$

and, since the last term is bounded by $C \|y - z\|^2$,

$$\left|\left|\left| P_{\mathbb{R}(y-z)} - P_{\Pi_y^\perp} \right|\right|\right| \leq C \|y - z\|$$

Doing the same at point z and summing these inequality yields:

$$\left|\left|\left| P_{\Pi_z^\perp} - P_{\Pi_y^\perp} \right|\right|\right| \leq C \|y - z\|$$

and the same inequality is true for $P_{\Pi_y} = Id - P_{\Pi_y^\perp}$ and P_{Π_z} □

The reader will notice that, in the case $d = 2$, Proposition 39 exhausts all possibilities and therefore, a Lagrangian critical measure μ is supported on a countable union of smooth manifolds of dimension 2 or less, and is regular with respect to the corresponding Hausdorff measure on each manifold.

This is still a far cry from the result we were hoping to obtain, and in particular, it does not prevent μ from having a very irregular support, even when Proposition 39 is valid (for instance, a Cantor set of non-zero mass on a part of dimension 1). On the other hand, one can easily construct examples of Lagrangian critical measures featuring either an infinite amount of connected components, or (in the case of Fig. 4.1) an infinite amount of components $E_{n,C}$ for the same dimension n but different "curvatures" C, which indicates that we cannot hope for much better than a countable union of regular manifolds for these supports.

An encouraging result, in this regard, is one of De Pauw [19] which, in spirit, reduces the problem to adding vectors to the spaces Π_y in order to obtain a Lipschitz selection of hyperplanes (of higher dimensions, but adding in this way a finite amount of directions). Indeed, the plans Π_y only need to be contained in a higher dimensional sub-space Π'_y, itself moving along $\mathrm{spt}(\mu)$ Lipschitz-continuously:

Under this weaker hypothesis, taking $A \subset \mathcal{E}_{k,C}$ negligible for the k dimensional Haussdorf measure, the set of points B for which $\dim(\Pi'_y) = d-k' > d-k$ verifies for any $y \in B$,

$$\mathcal{H}^{d-k'}(\Pi'_y \cap (y + \Pi_y)) = 0$$

as the sets measured are affine hyperplanes of dimensions strictly less than $d - k'$. Using the main result of [19], ρ gives no mass to $B + \Pi_B := \{b + x \mid b \in B, x \in \Pi_b\}$ and therefore, μ gives no mass to B. All that remains to do is use Proposition 38 on the remaining set. Of course, the larger space Π'_y would be made of Π_y and vectors contradicting the Lipschitz-continuity of the projection at y, but, so far, there does not seem to be a canonical way to add these vectors that does not depend heavily on the point y, preventing us from getting either global Lipschitz-continuity or a countable amount of such hyperplanes.

Appendix A

Measure theory

In this appendix, we develop several properties of measures, which we make ample use of in the rest of this thesis. To this avail, let us set a Polish (metric, complete and separable) space (X, d). For a more general overview of measure theory we refer the reader to Evans and Garzepy [24].

By a measure on (X, d), we mean an application from the Borelian tribe $\mathcal{T}$ to $\mathbb{R}$ (we do not consider infinite measures) verifying for any Borel set $A \subset X$:

- If $A = \bigcup_{i=1}^{\infty} A_i$ such that the Borel sets A_i are disjoints, for $i \neq j$, $A_i \cap A_j = \emptyset$:

$$\sum_{i=1}^{\infty} |\mu(A_i)| < +\infty \quad \text{and} \quad \mu(A) = \sum_{i=1}^{\infty} \mu(A_i).$$

- $\mu(A) = \sup\{\mu(K) \mid K \subset A,\ K \text{ compact}\}$. (Our measures are Radon measures by default).

The space of measures over X will be denoted $\mathcal{M}(X)$

We say that $\mu \in \mathcal{M}(X)$ is a probability measure when $\mu \geq 0$ (for any Borel set A, $\mu(A) \geq 0$) and $\mu(X) = 1$. We denote $\mathcal{P}(X)$ the (convex) space of probability measures over X. Let us also mention that we will write $\mathcal{M}_+(X)$ for the space of positive measures on X and $\mathcal{P}_N(X)$ for the space of probability measures on X which are discrete and supported on N points at most:

$$\mathcal{P}_N(X) = \left\{ \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \;\middle|\; (x_1, \ldots, x_N) \in X^N \right\}$$

Given a Borel map $T : X \to Y$, Y another Polish space, one can define the image of a measure $\mu \in \mathcal{M}(X)$ through the map T:

Definition 9. *Let $\rho \in \mathcal{M}(X)$ and T be a measurable mapping from X to Y. Then the push-forward measure of ρ along T is the measure $T\#\rho \in \mathcal{M}(Y)$ defined by:*

$$\text{For any Borel set } A \subset Y,\ T\#\rho(A) = \rho(T^{-1}(A))$$

or equivalently, for any $\phi \in \mathcal{C}_b^0(Y)$,

$$\int_Y \phi(y)\mathrm{d}(T\#\rho)(y) = \int_X \phi(T(x))\mathrm{d}\rho(x)$$

Note here that $T\#\mu$ has the same total mass as μ, only now in the space Y.

A.1 Narrow convergence of measures

The Radon property $\mu(A) = \sup\{\mu(K) \mid K \subset A,\ K \text{ compact}\}$ implies that any single measure is *tight*. More generally:

Definition 10. *A family $\{\mu_i\}_i$ of measures on the Polish space (X,d) is said to be* tight *when:*

For any $\varepsilon > 0$, there exists a compact set $K \subset X$ such that $\mu_i(X \setminus K) < \varepsilon$ for every $i \in I$

This notion of tightness implies sequential compactness for a specific topology on $\mathcal{M}(X)$, defined by duality with the space of real continuous bounded functions on X, $\mathcal{C}_b^0(X)$:

Definition 11. *Let $(\mu_n)_{n\in\mathbb{N}} \in \mathcal{M}(X)$, $\mu \in \mathcal{M}(X)$. Then we say that $(\mu_n)_{n\in\mathbb{N}}$ converges narrowly to μ, which we denote by $\mu_n \xrightarrow[n\to\infty]{} \mu$ when:*

$$\text{For any } \phi \in \mathcal{C}_b^0(X),\ \lim_{n\to\infty} \int_X \phi \mathrm{d}\mu_n = \int_X \phi \mathrm{d}\mu$$

Note here that $\mathcal{M}(X)$ is NOT the topological dual (for this convergence) of $\mathcal{C}_b^0(X)$ except when X is compact.

Compactness for tight sequences is given by the celebrated Prokhorov's theorem, which specifies, in some sense, the usual weak-$\star$ compactness of Banach-Alaoglu theorem for $\mathcal{P}(X)$, in the cases where X is not compact:

Theorem 40 (Prokhorov)**.** *If the sequence of probability measures $(\mu_n)_{n\in\mathbb{N}} \in \mathcal{P}(X)^{\mathbb{N}}$ is tight, then one may extract a sub-sequence $(\mu_{\phi(n)})_{n\in\mathbb{N}}$ which narrowly converges towards a measure $\mu \in \mathcal{P}(X)$:*

$$\mu_{\phi(n)} \xrightarrow[n\to\infty]{} \mu$$

Conversely, any sequence $(\mu_n)_{n\in\mathbb{N}}$ narrowly converging is tight.

A.2 Absolutely continuous measures and the disintegration theorem

Absolute continuity of a measure ρ with respect to a measure μ states that ρ sees at most the same sets as μ (and sometimes less):

Definition 12. *A measure $\rho \in \mathcal{M}(X)$ is said to be absolutely continuous with respect to $\mu \in \mathcal{M}(X)$ (written $\rho \ll \mu$) when for any Borel set A,*

$$\mu(A) = 0 \implies \rho(A) = 0.$$

On the other hand, ρ and μ are said to be mutually singular ($\rho \perp \mu$) when there exists a Borel set B such that $\rho(X \setminus B) = \mu(B) = 0$.

The following differentiation theorem by Lebesgue gives, for two measures, a useful decomposition of one into a part regular for the other and one singular:

Theorem 41. *Let μ, ν be two measures in $\mathcal{M}(X)$. Then there exists μ^{ac} and μ^{sing} in $\mathcal{M}(X)$ such that $\mu^{\mathrm{ac}} \ll \nu$, $\mu^{\mathrm{sing}} \perp \nu$ and,*

$$\mu = \mu^{\mathrm{ac}} + \mu^{\mathrm{sing}}$$

Although the image of a measure μ through a map, $T\#\mu$ is not always absolutely continuous with respect to μ (T can collapse the mass of μ too much), the following disintegration theorem allows for a representation of the image measure as the product of measures, one of which is the original measure μ:

Theorem 42. *Let $\rho \in \mathcal{M}(X)$ and $\pi : X \to Y$ be a Borel map (where Y is another Polish space). Let $\mu = \pi\#\rho$.*

Then there exists a family of measures $(\rho_y)_{y \in Y} \in \mathcal{M}(X)^Y$ such that:

- *For any Borel set $B \subset X$, $y \in Y \mapsto \rho_y(B)$ is a Borel measurable map.*
- *For μ-almost every y, ρ_y is supported on the fiber $\pi^{-1}(y)$.*
- *for every Borel-measurable function $f : X \to \mathbb{R}^+ \cup \infty$:*

$$\int_X f(x)\mathrm{d}\rho(x) = \int_Y \int_{\pi^{-1}(y)} f(x)\mathrm{d}\rho_y(x)\mathrm{d}\mu(y)$$

or equivalently,for any Borel set $A \subset X$,

$$\rho(A) = \int_Y \rho_y(A)\mathrm{d}\mu(y).$$

A.3 The Hausdorff measure for lower dimensional sets

We finish this appendix in the Euclidean space $X = \mathbb{R}^d$. The natural Borel measure on $\mathbb{R}^d$ to place on sets of dimension lower than d is the Hausdorff measure of the same dimension, and we use it at several points in this thesis. Note that it is not finite on sets of size too large, so it is not in $\mathcal{M}(\mathbb{R}^d)$ with our assumptions (however, this does not pose us any issues):

Definition 13. *Let $0 \leq k \leq d$ be an integer.*
The k-dimensional Haussdorf measure of a set $S \subset \mathbb{R}^d$ is:

$$\mathcal{H}^k(S) := \alpha(k) \lim_{\delta \downarrow 0} \inf \left\{ \sum_i (\mathrm{diam} U_i)^k \;\middle|\; S \subset \bigcup_i U_i,\ \mathrm{diam}(U_i) < \delta \right\}$$

where $\alpha(k)$ is the Lebesgue volume of the unit ball in $\mathbb{R}^k$.

This definition can be extended to k not being an integer, but we shall only consider Haussdorf measure on integer dimension in our computations.

Appendix B

Convex analysis

We recall in this appendix several facts on convex functions. For a more general overview of this branch of analysis, we refer the reader to Ekeland and Temam [23]. These functions are usually taken with value in $\bar{\mathbb{R}} := \mathbb{R} \cup +\infty$ as it is often convenient to give an infinite value to vectors which we want to forbid when doing minimization. We also assume that our functions are proper, meaning that they cannot be $+\infty$ everywhere (and the fact that they cannot be $-\infty$ is already assumed).

To fix things, if X is a real vector space, a function $f : X \to \bar{\mathbb{R}}$ is convex on X when for any $t \in [0; 1]$, $x, y \in X$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

It is concave when $-f$ is convex (the previous inequality is reversed).

Definition 14. *Let f be convex, bounded from below. We call domain of f the convex set:*

$$\mathrm{dom}(f) := \{x \in X \mid f(x) < +\infty\}$$

A simple illustration is given by the convex indicator function of a convex set C, in some way the simplest convex function of domain C:

$$\chi_C : x \mapsto \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$$

B.1 Convex duality:

Convex functions admit a very natural "weaker" notion of gradient, as a maximal slope for a hyperplane underneath their curve. This slope is naturally de-

fined as an element of E^*, the topological dual of E. We make here the remark, as a non-trivial example, that the space of (finite Radon) measures $\mathcal{M}(X)$ on a Polish compact space X is the topological dual of the space of real continuous (bounded) functions on X, $\mathcal{C}^0(X)$, through the dot-product

$$\mu, \phi \in \mathcal{M}(X) \times \mathcal{C}_b^0(X) \mapsto \int_X \phi(x) \mathrm{d}\mu(x),$$

meaning that any continuous linear functional over $\mathcal{C}^0(X)$ can be represented as the dot-product with a fixed measure over X. However, if X is not compact, $\mathcal{M}(X)$ is not this topological dual but the above duality would be sufficient to state most of the following theorems.

Definition 15. *Let $f : E \to \bar{\mathbb{R}}$ be convex.The subdifferential of f at $x \in E$ is the set of slopes:*

$$\partial f(x) := \{p \in E^* | \, \forall y \in E, \; f(y) \geq f(x) + p \cdot (y - x)\}$$

Similarly, for a concave function f, one can define its supergradient at $x \in E$ as the set $\partial^+ f(x) = -\partial(-f)(x)$. We use the notation ∂^+ in this second case only to make a difference with the notation for the subgradient. As we consider very few concave function in this thesis, this asymmetry of notations should not prove too disturbing.

It is an immediate but critical fact that, from this definition, $x \in E$ is a minimum of the convex function f if and only if $0 \in \partial f(x)$ and a symmetric fact is true for concave functions and their maxima.

As intuition dictates, the less vectors in its sub-gradient a convex function has, the smoother it will be. This is stated, in the simpler context of $X = \mathbb{R}^d$ equiped with the Euclidean topology, in the following proposition. The reader will note that most of it still holds true for much more general normed vector spaces, however, we have no use for such generalizations:

Proposition 43. *Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$. Then, if f has a non-empty sub-differential (with the previous definition) everywhere where it is finite, then f is convex.*

Furthermore, if $\partial f(x)$ is a singleton, then f is differentiable at x and

$$\partial f(x) = \{\nabla f(x)\}.$$

Finally, if there exists a continuous selection of sub-gradients

$$x \in \mathrm{dom}(f) \mapsto y_x \in \partial f(x),$$

then f is $\mathcal{C}^1$ on its domain.

Obviously, the same is true for concave functions, replacing sub-differentials with super-differentials.

A very notable point is the following theorem by Alberti which strengthens the celebrated Rademacher theorem on Lipschitz continuous functions:

Theorem 44 (Alberti, [3]). *Let $u : \mathbb{R}^d \to \mathbb{R}$ be a convex function and $0 < k < d$ an integer. Then the set*

$$\{x \in \mathbb{R}^d \mid \dim(\partial u(x)) \geq k\}$$

is a $\mathcal{H}^{d-k}$-rectifiable set of class $\mathcal{C}^2$ and dimension $d-k$, i.e it can be covered (up to removing a $\mathcal{H}^{d-k}$-negligible set) by countably many $(d-k)$-dimensional submanifolds of class $\mathcal{C}^2$.

The dimension of the $\mathcal{C}^2$-manifolds are simply that of the underlying space $\mathbb{R}^{n-k}$ in their definition. On the other hand, the k-dimensional Hausdorff measure is, once again, the one given in Definition 13, Appendix A.

In particular, coupling this theorem with Proposition 43, we immediately get that any convex function f on $\mathbb{R}^d$ is differentiable except on a countable union of sets essentially of dimension $d-1$, and in dimension $d = 1$, a convex function is differentiable everywhere on its domain except on an at most countable set!

To a convex function on E, one can associate another on E^* through duality. Note that in this definition, no assumption is made on the reflexivity of E, the second definition being made in order to have a symmetrical transformation already:

Definition 16. *For a convex function $f : E \to \bar{\mathbb{R}}$ l.s.c and proper, the Legendre transform of f at $p \in E^*$ is defined as:*

$$f^*(p) = \sup_{x \in E} \langle p, x \rangle - f(x)$$

Symmetrically, the Legendre transform of f^* (or bitransform of f) is defined as:

$$f^{**}(x) = \sup_{p \in E^*} \langle p, x \rangle - f^*(p)$$

We will make the assumption on all our convex functions that they are all lower semi-continuous, which is the lowest regularity assumption one can make in order for a function to have a minimum on every compact set:

Definition 17. *For a Polish space X, a function $f : X \to \bar{\mathbb{R}}$ is lower semi-continuous (l.s.c.) when for any sequence $(x_n)_{n \in \mathbb{N}} \in X^{\mathbb{N}}$ and $x \in X$:*

$$\lim_{n \to \infty} x_n = x \implies f(x) \leq \liminf_{n \to \infty} f(x_n)$$

It is interesting to note, although we do not use it, that convex functions which are lower semi-continuous for the standard topology on E, are also l.s.c. for the weak-$\star$ topology (which is a stronger notion of continuity).

The lower semi-continuity hypothesis, together with the fact that we assumed that our convex functions were proper, guarantees that taking the bi-transform of a convex function f brings one back to the original function. Furthermore, in that case, the sub-gradients of f and f^* are reciprocates of one another in the following sense:

Theorem 45 (Fenchel-Moreau)**.** *Let $f : E \to \bar{\mathbb{R}}$ be a convex (proper) l.s.c. function, f^* its Legendre transform.*

Then, $f^{**} = f$

Furthermore, for any $x \in E$ and $p \in E^*$

$$p \in \partial f(x) \iff x \in \partial f^*(p)$$

We finish these reminders on convex duality with a central result in Legendre-Fenchel duality, the celebrated Fenchel-Rockafellar duality theorem. Through this result, one can rewrite a convex minimization problem into a concave one (over the dual space E^*) which, often, enjoys more regularity. Note that there exists a multitude of versions of this result and we give here the one that is best fitted to our needs:

Theorem 46 (Fenchel-Rockafellar, [60])**.** *Let f, g be convex l.s.c. functions over respectively the Banach vector spaces E and F and $A \in \mathcal{L}(E, F)$ continuous. Assume that g is continuous at a point of $A\mathrm{dom}(f)$.*

Then,

$$\inf_{x \in E} f(x) + g(Ax) = \max_{p \in E^*} -f^*(A^*p) - g^*(-p)$$

where A^* is the standard adjoint operator of A.

B.2 Different kinds of convexity

Often, especially when dealing with geometrical transport of measures, the above notion of convexity is not enough to describe all observed phenomena. We recall the following "alternative" definitions of convexity, which sometimes have to replace the "rougher" scalar one of the previous section.

The functional featuring semi-discrete optimal transport terms studied in this thesis are all a little less than concave, which is a natural consequence of their definition as Moreau envelopes:

Definition 18. *Let $(E, \|.\|)$ be a normed vector space. The function $f : E \to \bar{\mathbb{R}}$ is said to be α-semiconcave ($\alpha > 0$) when $f - \alpha\frac{\|.\|^2}{2}$ is a concave function.*

Similarly, f is α-semiconvex when $f + \alpha\frac{\|.\|^2}{2}$ is convex ($-f$ is semiconcave).

One can note that there is a stronger notions of concavity/convexity, in the same spirit (and we make some use of this regularity in the case of optimal transport costs):

Definition 19. *Let E be a convex subset of a normed vector space $(X, \|.\|)$. The function $f : E \to \bar{\mathbb{R}}$ is α-convex (or strongly convex) when $f - \alpha\frac{\|.\|^2}{2}$ is a convex function.*

Similarly, f is α-concave when $f + \alpha\frac{\|.\|^2}{2}$ is concave ($-f$ is strongly-convex).

Note that if f is a λ-convex function, it also verifies a stronger convexity inequality: For $x, y \in E^2$, and $t \in [0; 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \lambda\frac{t(1-t)}{2}\|x - y\|^2$$

Finally, we move completely to the (convex) space $\mathcal{P}(\Omega)$: A different kind of convexity, more suitable for optimal transport, is the one observed when following geodesics and generalized geodesics in this space. We refer the reader to the seminal work of Carlier and Agueh on Wasserstein barycenter for a definition of both notions and examples, [2]. We merely recall the definition of the second notion of convexity, as we make use of this property on the Wasserstein distance in the proof of Proposition 6:

Definition 20. *Let $\Omega \subset \mathbb{R}^d$ and $F : \mathcal{P}(\Omega) \to \bar{\mathbb{R}}$. We say that F is convex along generalized geodesics when, for any μ, ρ^0, ρ^1 in $\mathcal{P}(\Omega)$, $t \in [0; 1]$*

$$F(\rho_\mu^t) \leq tF(\rho^0) + (1-t)F(\rho^1)$$

where ρ_μ^t is the generalized geodesic between ρ^0 and ρ^1, with respect to μ defined by:

For any $\phi \in C^0(\Omega)$, $\displaystyle\int_\Omega \phi(x)\mathrm{d}\rho_\mu^t(x) = \int_{\Omega^3} \phi((1-t)x_1 + tx_3)d\gamma^0(x_2, x_1)d\gamma^1(x_2, x_3)$

and γ^i is an optimal transport plan between μ and ρ^i.

Let us recall briefly that when F has the integral form (1.14) considered in the end of Chapter 1, with the necessary assumptions on f to obtain lower semi-continuity and convexity, assuming furthermore:

- $f(0) = 0$.
- $\mathrm{dom}(f) \subset \mathbb{R}^+$
- The function $t \in \mathbb{R}_+^* \mapsto t^d f(t^{-d})$ is convex and non-increasing.

then F is convex along generalized geodesics (see [4], proposition 9.3.9).

We finish this appendix by mentioning that the 2-Wasserstein distance (on $\mathcal{P}_2(\mathbb{R}^d)$) is 2-strongly convex along generalized geodesics meaning, with the notations of Definition 20, that:

For any μ, ν, ρ^0, ρ^1 in $\mathcal{P}(\Omega)$, $t \in [0;1]$:

$$W_2^2(\nu, \rho_\mu^t) \leq tW_2^2(\nu, \rho^0) + (1-t)W_2^2(\nu, \rho^1) - t(1-t)W_2^2(\rho^0, \rho^1)$$

Bibliography

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010.

[2] M. Agueh and G. Carlier. Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.

[3] G. Alberti. On the structure of singular sets of convex functions. *Calculus of Variations and Partial Differential Equations*, 2:17–27, 1994.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2nd edition, 2008.

[5] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.

[6] F. Aurenhammer, F. Hoffmann, and B. Aronov. Minkowski-type theorems and least-squares clustering. *Algorithmica*, 20(1):61–76, 1998.

[7] M. Balzer, T. Schlömer, and O. Deussen. Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Transactions on Graphics (TOG)*, 28(3):1–8, 2009.

[8] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active Particles, Volume 1*, pages 141–171. Springer, 2017.

[9] S. Bobkov and M. Ledoux. *One-dimensional empirical measures, order statistics, and Kantorovich transport distances*, volume 261. American Mathematical Society, 2019.

[10] D. Bourne, P. J. J. Kok, S. M. Roper, and W. D. T. Spanjer. Laguerre tessellations and polycrystalline microstructures: A fast algorithm for generating grains of given volumes. *Philosophical Magazine*, 100(21):2677–2707, 2020.

[11] D. P. Bourne, B. Schmitzer, and B. Wirth. Semi-discrete unbalanced optimal transport and quantization. *arXiv: Optimization and Control*, 2018.

[12] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

[13] L. M. Briceno-Arias, D. Kalise, and F. J. Silva. Proximal methods for stationary mean field games with local couplings. *SIAM Journal on Control and Optimization*, 56(2):801–836, 2018.

[14] P. Cardaliaguet. Notes from P.-L. Lions' lectures at the Collège de France, Technical report. 2012.

[15] P. Cardaliaguet, A. R. Mészáros, and F. Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016.

[16] Y. Chen, T. T. Georgiou, and M. Pavon. On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *Journal of Optimization Theory and Applications*, 169(2):671–691, 2016.

[17] F. De Goes, K. Breeden, V. Ostromoukhov, and M. Desbrun. Blue noise through optimal transport. *ACM Transactions on Graphics (TOG)*, 31(6):1–11, 2012.

[18] F. De Gournay, J. Kahn, and L. Lebrat. Differentiation and regularity of semi-discrete optimal transport with respect to the parameters of the discrete measure. *Numerische Mathematik*, 141(2):429–453, February 2019.

[19] T. De Pauw. On Lebesgue null sets. *arXiv preprint arXiv:1904.12276*, 2019.

[20] G. De Philippis, A. R. Mészáros, F. Santambrogio, and B. Velichkov. BV Estimates in Optimal Transportation and Applications. *Archive for Rational Mechanics and Analysis*, 219(2):829–860, February 2016.

[21] Q. Du, M. Emelianenko, and L. Ju. Convergence of the Lloyd algorithm for computing centroidal voronoi tessellations. *SIAM journal on numerical analysis*, 44(1):102–119, 2006.

[22] R. Duits, S. P. L. Meesters, J.-M. Mirebeau, and J. M. Portegies. Optimal paths for variants of the 2d and 3d reeds–shepp car with applications in image analysis. *Journal of mathematical imaging and vision*, 60(6):816–848, 2018.

[23] I. Ekeland and R. Temam. *Convex analysis and variational problems*. SIAM, 1999.

[24] Lawrence C Evans and Ronald F Garzepy. *Measure theory and fine properties of functions*. Routledge, 2018.

[25] A. Figalli. The optimal partial transport problem. *Archive for rational mechanics and analysis*, 195(2):533–560, 2010.

[26] A. Figalli, Y.-H. Kim, and R. J McCann. When is multidimensional screening a convex program? *Journal of Economic Theory*, 146(2):454–478, 2011.

[27] A. Genevay, G. Peyré, and M. Cuturi. Learning generative models with Sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617. PMLR, 2018.

[28] I. Gentil, C. Léonard, and L. Ripani. About the analogy between optimal transport and minimal entropy. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 26, pages 569–600, 2017.

[29] A. Gersho and R. M. Gray. *Vector quantization and signal compression*. Springer Science & Business Media, 1992.

[30] N. Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(P_2(M), W_2)$. *Methods and Applications of Analysis*, 18(2):127–158, 2011.

[31] E. Giusti. *Direct methods in the calculus of variations*. World Scientific, 2003.

[32] S. Graf and H. Luschgy. *Foundations of quantization for probability distributions*. Springer, 2007.

[33] M. Huang, R. P. Malhamé, P. E. Caines, et al. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.

[34] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*, 29, 04 2000.

[35] Leonid V Kantorovich. On the translocation of masses. In *Dokl. Akad. Nauk. USSR (NS)*, volume 37, pages 199–201, 1942.

[36] J. Kitagawa, Q. Mérigot, and B. Thibert. Convergence of a Newton algorithm for semi-discrete optimal transport. *Journal of the European Mathematical Society*, 21(9):2603–2651, 2019.

[37] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[38] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. ii–horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006.

[39] H. Lavenant and F. Santambrogio. New estimates on the regularity of the pressure in density-constrained Mean Field Games. *Journal of the London Mathematical Society*, 2019.

[40] H. Leclerc, Q. Mérigot, F. Santambrogio, and F. Stra. Lagrangian discretization of crowd motion and linear diffusion. *SIAM Journal on Numerical Analysis*, 58(4):2093–2118, 2020.

[41] G. Leoni and M. Morini. Necessary and sufficient conditions for the chain rule in $W^{1,1}_{loc}(\mathbb{R}^N;\mathbb{R}^d)$ and $BV_{loc}(\mathbb{R}^N;\mathbb{R}^d)$. *Journal of the European Mathematical Society*, 9(2):219–252, 2007.

[42] S. Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. *Calculus of Variations and Partial Differential Equations*, 28(1):85–120, 2007.

[43] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta mathematica*, 202(2):241–283, 2009.

[44] X.-N. Ma, N. S. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transportation problem. *Archive for rational mechanics and analysis*, 177(2):151–183, 2005.

[45] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A macroscopic crowd motion model of gradient flow type. *Mathematical Models and Methods in Applied Sciences*, 20(10):1787–1821, 2010.

[46] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel. Handling congestion in crowd motion modeling. *Networks Heterogeneous Media*, 6(3):485–519, 2011.

[47] B. Maury and J. Venel. A discrete contact model for crowd motion. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(1):145–168, 2011.

[48] R. J McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[49] R. J. McCann. Polar factorization of maps on riemannian manifolds. *Geometric & Functional Analysis GAFA*, 11(3):589–608, 2001.

[50] Q. Mérigot and J.-M. Mirebeau. Minimal geodesics along volume-preserving maps, through semidiscrete optimal transport. *SIAM Journal on Numerical Analysis*, 54(6):3465–3492, 2016.

[51] Q. Merigot, F. Santambrogio, and C. Sarrazin. Non-asymptotic convergence bounds for wasserstein approximation using point clouds. *arXiv preprint arXiv:2106.07911*, 2021.

[52] J.-M. Mirebeau. Anisotropic fast-marching on cartesian grids using voronoi's first reduction of quadratic forms. *HAL (Preprint), April*, 2017.

[53] J.-M. Mirebeau. Fast-marching methods for curvature penalized shortest paths. *Journal of Mathematical Imaging and Vision*, 60(6):784–815, 2018.

[54] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*, 1781.

[55] Q. Mérigot and B. Thibert. Chapter 2 - optimal transport: discretization and algorithms. In Andrea Bonito and Ricardo H. Nochetto, editors, *Geometric Partial Differential Equations - Part II*, volume 22 of *Handbook of Numerical Analysis*, pages 133–212. Elsevier, 2021.

[56] G. Peyré and M. Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[57] B. T. Polyak. Gradient methods for minimizing functionals. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki*, 3(4):643–653, 1963.

[58] A. Prosinski and F. Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. *Stochastics*, 89(6-7):923–942, 2017.

[59] J. Rataj and L. Zajíček. On the structure of sets with positive reach. *Mathematische Nachrichten*, 290(11-12):1806–1829, Jan 2017.

[60] R. T. Rockafellar. Extension of fenchel'duality theorem for convex functions. *Duke mathematical journal*, 33(1):81–89, 1966.

[61] A. Roudneff. *Modelisation macroscopique de mouvements de foule*. PhD thesis, 2011. Thèse de doctorat dirigée par Maury, Bertrand Mathématiques Paris 11 2011.

[62] F. Santambrogio. Absolute continuity and summability of transport densities: simpler proofs and new estimates. *Calculus of variations and partial differential equations*, 36(3):343–354, 2009.

[63] F. Santambrogio. *Optimal transport for applied mathematicians*. Springer, 2015.

[64] F. Santambrogio. Lecture notes on variational mean field games. In *Mean Field Games*, pages 159–201. Springer, 2020.

[65] C. Sarrazin. Lagrangian discretization of variational mean field games. *arXiv preprint arXiv:2010.11519*, 2020.

[66] AN Tolstoi. Methods of finding the minimal total kilometrage in cargo transportation planning in space. *TransPress of the National Commissariat of Transportation*, 1:23–55, 1930.

[67] C. Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

[68] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[69] S.-Q. Xin, Bruno Lévy, Z. Chen, L. Chu, Y. Yu, C. Tu, and W. Wang. Centroidal power diagrams with capacity constraints: Computation, applications, and extension. *ACM Transactions on Graphics (TOG)*, 35(6):1–12, 2016.

[70] D.-M. Yan, J.-W. Guo, B. Wang, X.-P. Zhang, and P. Wonka. A survey of blue-noise sampling and its applications. *Journal of Computer Science and Technology*, 30(3):439–452, 2015.

université PARIS-SACLAY | ÉCOLE DOCTORALE de mathématiques Hadamard (EDMH)

Titre: Discrétisation Lagrangienne de problèmes variationnels dans des espaces de Wasserstein.
Mots clés: Transport optimal de mesures; Calcul des variations; Points critiques, Théorie des (analyse mathématiques); Principes variationnels; Problèmes extrémaux (mathématiques).

Résumé: Dans ce mémoire, nous étudions la discrétisation de problèmes variationnels, au moyen de méthodes du transport optimal semi-discret. Bien que les techniques simplifient grandement la résolution de ces problèmes elles introduisent également des termes non-convexes dans des problèmes qui étaient convexes avant discrétisation. Le modèle principalement étudié est celui d'une foule se déplaçant de façon à minimiser une énergie globale qu'elle accumule durant son mouvement. Cette évolution est modelée par un problème de jeu à champ moyen variationnel, pour la minimisation d'une énergie comprenant un terme qui pénalise la présence de zones de congestions. Nous approchons les solutions par les trajectoires d'un nombre fini d'individus, cependant ces foules discrètes ne sont pas admissible pour le problème de jeu à champ moyen, à cause justement du terme de congestion. Pour définir un problème similaire pour notre foule discrète, nous remplaçons le terme problématique par une régularisation de type Moreau-Yosida. Notre résultat principal affirme alors la convergence des minimiseurs du problème discret vers une solution du problème de jeu à champ moyen initial. Cependant, cette convergence est conditionnelle à un choix approprié de paramètres de régularisation et sa détermination est encore une question ouverte. Cela n'empêche pas la résolution du problème discret pour des choix arbitraires de ces paramètres, laquelle génère des trajectoires cohérentes avec le comportement imposé par le problème de jeu à champ moyen correspondant. La régularisation du terme de congestion, définie par une enveloppe de Moreau pour la distance de Wasserstein 2 introduit donc une non-convexité dans le problème discrétisé, et l'on est en droit de craindre d'aboutir à des situations critiques non-minimisantes en le résolvant numériquement. Afin de mieux comprendre ces dangers, nous avons étudié la structure de ces situations critiques pour le problème non convexe, plus simple mais néanmoins proche, de quantification optimale uniforme d'une mesure. Bien que nous ayons réussi à clarifier la forme de ces points critiques et en particulier le devenir de la discrétisation lorsque le nombre de particules croit à l'infini, la classification complète de ces configurations possibles demeure largement inachevée.

Title: Lagrangian discretization of variational problems in Wasserstein spaces.
Keywords: Optimal transport; Calculus of variations; Critical points; Variationnal principles; Extremal problems.

Abstract: In this thesis, we study the discretization of variational problems, via semi-discrete optimal transport methods. Although these techniques yield much simpler expressions for their solutions, they also introduce non-convex terms in what were convex problems before discretization. The main model is that of a crowd, which minimizes a global energy accumulated during its motion. This evolution is modeled by a variational mean field game problem, which energy features a term penalizing highly congested area. We approximate the solutions using the trajectories of a finite set of particles, however, such finite crowds are not admissible for this problem, due to the congestion term. In order to circumvent this issue, we consider a similar energy, however one where the congestion is penalized by a regularized version of the previous term. Our main result is then the convergence of these discrete minimizers towards a solution of the mean field game problem. However this statement requires the selection of an appropriate sequence of regularization parameters, which cannot be determined just yet. This is only partially an impediment, as the discretization used is very robust, and even for relatively large values of these parameters, the numerical experiments showcase a satisfying behavior regarding the modeled phenomenon. The regularization of the congestion term is defined as a Moreau envelope, using the 2-Wasserstein distance. These expressions introduce a non-convexity in the discretized problem which could make numerically computing their solutions difficult. In order to better understand these issues, we gather insightful information regarding the structure of critical points for the simpler but related non-convex problem of uniform optimal quantization. These critical configurations and in particular, their limits as measures when the discretization becomes more and more precise, remain a vastly unexplored part of these optimization problems.