Interacting particles systems, Wasserstein gradient flow approach

Maxime Laborde

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Systèmes de particules en interaction, approche par flot de gradient dans l'espace de Wasserstein

École Doctorale de Dauphine — ED 543

Spécialité **Sciences**

COMPOSITION DU JURY :

M. Giuseppe BUTTAZZO
Université de Pise
Président du jury

M. Guillaume CARLIER
Université Paris-Dauphine
Directeur de thèse

M. Jean DOLBEAULT
Université Paris-Dauphine
Membre du jury

M. Daniel MATTHES
Technische Universität München
Rapporteur

M. Bertrand MAURY
Université Paris-Sud
Membre du jury

M. Michel PIERRE
ENS Cachan, Antenne de Bretagne
Rapporteur

Soutenue le 01.12.2016
par Maxime LABORDE

Dirigée par **Guillaume CARLIER**

DAUPHINE UNIVERSITÉ PARIS | PSL RESEARCH UNIVERSITY PARIS

Remerciements

Tout d'abord, je voudrais remercier très très chaleureusement Guillaume Carlier pour avoir accepté d'encadrer ma thèse. Durant ces trois années, j'ai appris énormément à son contact grâce à nos longues (et parfois tardives) discussions tant sur le plan mathématique que sur le plan humain. Il m'a permis de découvrir tout un domaine des mathématiques que j'ignorais et travailler avec lui est toujours un réel plaisir. Sa bonne humeur et son optimisme m'ont permis de supporter les moments de doutes. *Encore merci pour tout ce que tu m'as apporté.*

Je tiens également à remercier Daniel Matthes et Michel Pierre qui m'ont fait l'honneur d'être les rapporteurs de ma thèse ainsi que Giuseppe Buttazzo, Jean Dolbeault et Bertrand Maury pour avoir accepté d'être les membres de mon jury. Encore merci à Giuseppe Buttazzo qui m'a permis de découvrir Pise. Travailler avec lui est très stimulant et je le remercie pour sa patience.

Je suis aussi très reconnaissant envers tous les enseignants que j'ai eus pendant mes études, qui m'ont donné le goût de continuer dans le domaine des Mathématiques. En particulier, je voudrais remercier Filippo Santambrogio avec qui j'ai effectué mon stage de M2 et qui m'a proposé d'effectuer ma thèse avec Guillaume Carlier. Sa constante bonne humeur et ses conseils mathématiques m'ont été très profitables.

J'ai une pensée particulière pour Martial Agueh. Ces travaux m'ont beaucoup inspiré pendant ma thèse et j'ai eu l'honneur d'aller le voir à Victoria pendant un mois et demi. Nos discussions quotidiennes, toujours dans la bonne humeur, m'ont beaucoup aidé. Martial m'a permis, avec Guillaume, de me sentir comme chez moi très rapidement à Victoria.

Merci beaucoup à Jean-David Benamou pour nos discussions et sa patience durant mon apprentisage de Freefem. Il m'a été d'une aide très préciseuse.

Durant ma thèse, j'ai fait la connaissance de Thomas Gallouët et Léonard Monsaingeon. Récemment nous avons commencé à travailler ensemble. C'est un réel plaisir de travailler, de discuter avec eux et je les en remercie beaucoup.

Un grand merci aussi à toute l'équipe du CEREMADE. Merci en particulier à Thibaut, Isabelle, Jonathan, Luca, Raphaël B., Camille, Daniela, Pierre, Lénaïc, Jean, Fang, Marco, Quentin, Clara, Roméo, Julien, Aude, Raphaël D., Arnaud, Michaël, Shuoqing, William, Jonas, Bernhard, Amal, Benjamin, Vincent pour toutes les discussions, les pauses à rallonge et les sorties après le séminaire pour aller boire "un" verre. Je remercie également Isabelle Bellier, Marie Belle et César Faivre pour leur disponibilité et leur aide dans toutes les démarches administratives, en particulier pour l'organisation de l'école d'été.

Durant ces trois années, j'ai aussi effectué une mission d'enseignement à l'Université Paris-Dauphine et je souhaite remercier toute l'équipe du MIDO ainsi que les professeurs pour qui j'ai effectué des TD : Daniela Tonon, Olivier Glass et Boris Haspot. Ce fut vraiment très agréable de travailler avec eux.

Je remercie tous les membres de l'équipe INRIA MOKAPLAN, qui m'ont offert un super cadre de travail et en particulier à Simon et Luca sans qui le travail dans le bureau et la vie en dehors

n'auraient pas été pareils.

De nombreuses personnes m'ont beaucoup apporté et aidé pendant ces trois ans et avant et je voudrais ici les remercier. Je pense notamment à mes camarades de la fac d'Orsay, Guillaume, Malek, Sébastien, Nicolas, Emilie, Céline, Rémi, Ridha, Lucie, Lucile, Chloé..., à l'équipe de l'Huma, Emilie, Rémi, Benoit, Pilar, Luis..., à la Team Staff, Guillaume, Sandrine, Océane, Romain, Alex, Carole, Filou, Camille, Aude, Flore, Vincent..., aux camarades du Baker, Pierre, Dario, Margherita, Thibaut, Ivan, Camille, Riccardo, Davide, Carlo, etc.

Je voudrais remercier encore une fois Guillaume, Sandrine, Emilie, Rémi, Simon, Luca et Manitra pour m'avoir supporté dans les bons et surtout les mauvais moments !

J'ai aussi une grosse pensée pour mes amis Gâtinais avec qui c'est toujours un plaisir de faire la fête de retour à Parthenay ou ailleurs : Yoann, Simon, Pacôme, Babar, Roland, Rachel, Pauline B, Jonathan, Ben, Guigui, Antoine, Charlène, Pauline M, Camille, Anne, Clément P, Clément B, Pauline L, Justine, Caryl, Mickaël, P-A, etc.

Enfin pour finir je remercie ma famille qui m'a supporté pendant toute cette période. *Merci beaucoup pour le soutien que vous m'avez apporté depuis le début surtout dans les moments difficiles.*

Contents

Chapter 1

Introduction générale

Dans cette thèse, on s'intéresse à des sytèmes d'espèces en interaction. Ces dynamiques peuvent modéliser plusieurs phénomènes tels que les mouvements de foules, l'interaction entre espèces biologiques, la croissance de tumeur, l'aménagement urbain etc. Le point commun entre tous ces modèles est que les mouvements d'une population sont décrits par l'évolution de sa densité et non de chaque individu. Ces modèles sont appelés macroscopiques. De plus, dans de nombreux phénomènes comme les mouvements de foule ou la migration cellulaire, la taille de la population est fixée et reste inchangée et donc la théorie du transport optimal nous offre un bon outils pour étudier ces systèmes.

Le problème du transport optimal a été introduit par Monge au 18ème siècle, [95], et consiste à minimiser l'énergie nécessaire pour transporter une masse de volume donnée vers une autre masse de même volume. L'exemple classique introduit par Monge consiste à transporter une pile de sable vers un trou de même volume en minimisant la distance parcourue, c'est à dire, on cherche une fonction T telle que toute la masse de sable à un point x soit transférée à un point $T(x)$ du trou, tout en minimisant le déplacement moyen. Ce problème est resté sans réponse pendant de nombreuses années. En 1942, Kantorovich a proposé un problème relaxé plus facile à resoudre, [68]. Dans son modèle, la masse au point x peut s'étaler dans le trou et donc on ne recherche plus une fonction mais un plan de transport. Ce n'est qu'à la fin des année 80 que Brenier a résolu le problème de Monge pour le coût quadratique [21]. Depuis, la théorie du transport optimal a connu un regain d'activité avec de nombreuses applications et les livres de Villani, [116, 117], ou Santambrogio [112], sont des références très complètes sur ce sujet.

Une de ces applications est la résolution d'équations aux dérivées partielles (EDPs). En effet, le problème de transport optimal pour le coût quadratique défini une distance sur l'espace des mesures de probabilité et on appelle cette espace muni de cette distance l'espace de Wasserstein. Certaines équations paraboliques peuvent être interprétées comme des flots de gradient par rapport à cette distance. Jordan, Kinderlehrer et Otto ont été les premiers à avoir cette idée dans [65]. Dans cet article, ils montrent que l'équation de Fokker-Planck peut être vue comme le flot de gradient de l'Entropie et d'une énergie de potentiel dans l'espace de Wasserstein. Cette théorie, appelée flot gradient dans l'espace de Wasserstein, a été largement developpée ces derniers années pour montrer l'existence de solutions pour de nombreuses EDPs telles que l'équation des milieux poreux [98], les équations d'aggrégations [37], des équations paraboliques dégénérées [97, 1], l'équation de Keller-Segel [18] ou encore des équations du quatrième ordre [86].

Le but de cette thèse est d'étendre cette théorie pour étudier des équations ne provenant pas de flots gradient dans l'espace de Wasserstein, par exemple, les systèmes de populations avec des interactions non locales ou encore des équations de transport avec un champ de vecteur n'étant pas donné par le gradient d'un potentiel. Les interactions qu'on va étudier vont être de plusieurs natures. Tout d'abord, dans les chapitres 4 et 5, on va s'intéresser à des interactions non locales sur la vitesse qui peuvent modéliser l'attirance ou la répulsion d'une espèce envers une autre. Le chapitre 6 est consacré à l'étude de modèle de congestion pour une population comprenant différents types d'individus. Par exemple, lorsque deux foules se croisent la poulation globale des deux foules doit satisfaire des contraintes sur le nombre de personnes au mètre carré. Ce type de modèles a été introduit dans le cas d'une population par Maury, Roudneff-Chupin et Santambrogio [89], et

on va étendre ces résultats à une population divisée en deux groupes (ou plus) ayant chacun un comportement distinct. Le chapitre 7 étudie des systèmes où la taille des populations, fixée dans les précédents modèles, peut varier en fonction du comportement des autres espèces. Un exemple pour illustrer ce phénomène provient des systèmes proie-prédateur avec diffusion. En effet, si on se place dans le cas de deux espèces, une proie et un prédateur, chacune va avoir une diffusion qui lui est propre dans l'espace mais elle vont interagir car les prédateurs vont chercher à manger leur proie pour se nourrir. On appliquera aussi cette méthode pour des équations du type Hele-Shaw pouvant modèliser la croissance tumorale [102, 101]. Le dernier type d'interactions étudiés provient d'un modèle de planification urbaine. On considère deux espèces, les habitants et les services. Les habitants cherchent à se rapprocher des services dans le but de minimiser leur coût de transport et de plus sont soumis à une congestion car ils ne veulent pas habiter dans un endroit trop peuplé. De leur coté les services veulent aussi se trouver proches des habitations mais veulent être concentrés afin de minimiser leur coût de gestion. Ce genre de modèle a été introduit dans le cas statique dans [25, 31]. Dans le chapitre 8, considérant qu'une ville est en constante évolution, on s'est intéressé au modèle dynamique et on remarque ici, que l'interaction se situe dans un problème de transport optimal entre les deux espèces.

Tous les systèmes étudiés sont des extensions de la théorie des flots de gradients dans l'espaces de Wasserstein. Cette théorie consiste à construire par récurrence une suite de mesures de probabilité, $(\rho_h^k)_k \subset \mathcal{P}_2(\Omega)$, de la façon suivante

$$\rho_h^{k+1} \in \operatorname*{argmin}_{\rho\in\mathcal{P}_2(\Omega)} \frac{1}{2h} W_2^2(\rho, \rho_h^k) + \mathcal{E}(\rho), \tag{1.1}$$

où Ω est un sous-ensemble de $\mathbb{R}^n$, W_2 indique la distance de Wasserstein d'ordre 2 sur $\mathcal{P}_2(\Omega)$, $\mathcal{E} : \mathcal{P}_2(\Omega) \to \mathbb{R}$ est une fonctionnelle donnée et $h > 0$ est un pas de temps. Ce schéma est appelé schéma de JKO. D'après [65, 97, 4], on sait que l'interpolation en temps constante par morceaux va converger vers une solution de l'équation de continuité

$$\partial_t \rho - \operatorname{div}(\rho \nabla (F'(\rho) + V)) = 0,$$

dans le cas où

$$\mathcal{E}(\rho) = \int_\Omega F(\rho) + \int_\Omega V\rho.$$

Le premier chapitre 3 est basé sur un article en collaboration avec Jean-David Benamou et Guillaume Carlier. On y propose une nouvelle méthode numérique, nommée ALG2-JKO, pour résoudre des flots de gradient dans l'espace de Wasserstein en résolvant (1.1) itérativement. La difficulté dans ce problème réside dans le terme de Wasserstein qui implique de résoudre un problème de transport optimal, souvent très coûteux à chaque étape. Il existe plusieurs méthodes pour résoudre ce problèmes.

Notre méthode est basée sur la formulation dynamique de Benamou-Brenier du transport optimal,

$$W_2^2(\rho_0, \rho_1) = \inf \left\{ \int_0^1 \int_\Omega \frac{|m_t(x)|^2}{\rho_t}\, dxdt \,:\, \partial_t \rho_t + \operatorname{div}(m_t) = 0,\ \rho_{t=0,1} = \rho_0, \rho_1 \right\},$$

qui est convexe. Elle nous permet donc de proposer une formulation convexe de (1.1) à chaque étape qu'on résout à l'aide d'une méthode de Lagrangian augmenté en utilisant FreeFem++. Cette méthode a été testée sur l'équation des milieux poreux. De plus, elle est très maniable et on va pouvoir l'utiliser dans plusieurs chapitres de la thèse pour illustrer des résultats notamment sur l'interaction entre espèces (chapitre 4), les mouvements de foules (chapitre 6), les systèmes diffusifs proie-prédateur (chapitre 7) ou la croissance tumorale (chapitre 7), qui ne proviennent plus de flots de gradient dans l'espaces de Wasserstein mais qui sont des perturbations de cette méthode.

Les deux chapitres suivants 4 et 5 sont consacrés à l'édude d'existence et d'unicité pour des équations de continuité du type

$$\partial_t \rho + \operatorname{div}(\rho v[\rho]) = 0, \tag{1.2}$$

où le champ de vecteur $v[\rho]$ est le somme d'un terme régulier, qui peut se traiter en utilisant la méthode des caractéristiques, et d'un terme provenant d'un flot de gradient dans l'espace de Wasserstein, typiquement une diffusion.

Le chapitre 4 presente un premier résultat d'existence et d'unicité pour une classe de systèmes paraboliques avec diffusions non linéaires et interactions non locales,

$$\partial \rho_i - \Delta P_i(\rho_i) - \operatorname{div}(\rho_i \nabla V_i[\rho_1, \dots, \rho_l]) = 0,$$

avec $i \in [\![1, l]\!]$ sur Ω un ouvert de $\mathbb{R}^n$. Dans ce chapitre, le terme régulier est toujours donné par le gradient d'un potentiel mais ces systèmes ne sont pas des flots de gradient dans l'espace de Wasserstein produit car les V_i n'ont a priori aucun lien entre eux. Ils peuvent être considérés comme des perturbations régulières de cette théorie. Pour contourner cette difficulté, Di Francesco et Fagioli ont introduit, dans le cadre non diffusif, un schéma de JKO semi-implicite [54]. Il consiste à fixé le poteniel V dans l'énergie de potentiel avec les mesures définies à l'instant précédents, à chaque étape: On construit des suites $(\rho^k_{i,h}) \subset \mathcal{P}_2(\Omega)$ telles que $\rho^0_{i,h} = \rho_{i,0}$ et

$$\rho^{k+1}_{i,h} \in \operatorname*{argmin}_{\rho} \left\{ W_2^2(\rho, \rho^k_{i,h}) + 2h \left(\int_\Omega F_i(\rho) + \int_\Omega V_i[\rho^k_{1,h}, \dots, \rho^k_{l,h}]\rho \right) \right\}.$$

La principale difficultée de ce chapitre est d'obtenir une convergence forte afin de pouvoir passer à la limite dans le terme de diffusion non linéaire. Pour ce faire, on va utiliser deux méthodes différentes. La première, sur $\mathbb{R}^n$, est basée sur l'argument de *flow interchange* introduit par Matthes, McCann et Savaré [86] pour obtenir de la compacité en espace et sur une extension du lemme d'Aubin-Lions dûe à Rossi et Savaré [107]. La méthode de flow interchange présente cependant le désavantage d'utiliser la convexité géodésique de l'Entropie ce qui nous oblige à travailler sur un domaine convexe. Dans le cas d'un ouvert borné de $\mathbb{R}^n$, on propose une alternative pour obtenir la compacité en espace. En utilisant les équations d'Euler-Lagrange associés aux problèmes de minimisations, on obtient une estimation BV en espace sur une quantité non linéaire de $\rho_{i,h}$. On finit ce chapitre en donnant un résultat d'unicité qui utilise des arguments de convexité géodésique et des simulations numériques pour des sytèmes d'espèces en interactions en se servant de l'algorithme développé chapitre 3.

Comme on l'a vu dans le cas précédent, le champ de vecteur est donné par un gradient. Le chapitre 5 étend les résultats d'existence et d'unicité à des champs de vecteur généraux réguliers et est issue d'articles en collaboration avec Guillaume Carlier [35, 33]. C'est à dire, on cherche à résoudre

$$\partial \rho_i - \Delta P_i(\rho_i) - \operatorname{div}(\rho_i U_i[\rho_1, \dots, \rho_l]) = 0.$$

La première méthode pour obtenir ce résultat est complétement parabolique. On régularise l'équation (1.2) et on la réécrit de la manière suivante

$$\partial_t \rho - \operatorname{div}(a_\varepsilon(\rho)\nabla\rho) + b[\rho] \cdot \nabla\rho + c[\rho]\rho = 0,$$

avec $a_\varepsilon(\rho), b[\rho]$ et $c[\rho]$ appartenant à $L^\infty((0,T)\times\Omega)$ et $1/\varepsilon \geqslant a_\varepsilon(\rho) \geqslant \varepsilon$. On trouve une solution de l'équation régularisée par une méthode de point fixe puis en utilisant des méthodes classiques en équation parabolique, on obtient des estimations indépendantes du paramètre ε nous permettant de passer à la limite dans l'équation. Cette méthode se généralise facilement au cas des systèmes. De plus, on donne un résultat de contraction H^{-1} dans le cas où la diffusion n'est pas dégénérée.

Cette méthode, bien que simple, a le désavantage de ne pas être constructive et donc pour remédier à ce problème, on a, dans un second temps, développé une méthode de splitting dans l'espace de Wasserstein. L'idée de notre splitting consiste à utiliser la décomposition d'Helmholtz sur $U[\rho]$ i.e

$$U[\rho] = -W[\rho] + \nabla V[\rho],$$

où $W[\rho]$ est un champ de vecteur à divergence nulle. La partie à divergence nulle va être traitée par des phases de transport pur et on va utiliser le schéma de JKO semi-implicite développé chapitre 4 pour gérer le terme gradient. Dans le but de pouvoir appliquer la théorie de DiPerna-Lions, on va supposer que W vérifie une régularité Sobolev. L'avantage de ce splitting réside

dans le fait que la phase de transport pur conserve l'énergie interne ce qui va nous permettre de retrouver facilement les estimations habituelles dans la théorie des flots de gradient dans l'espace de Wasserstein (estimation sur l'énergie, estimation sur les moments, compacité en temps). De plus, on propose une méthode générale s'appliquant sur un ouvert de $\mathbb{R}^n$, pas nécessairement convexe ou borné, pour obtenir la compacité en espace nous permettant d'appliquer le théorème de Rossi et Savaré [107] afin de retrouver la convergence forte de la suite des solutions discrètes. On remarque que le théorème d'unicité des solutions démontré au chapitre précédent s'applique à ce cas. Signalons qu'un schéma de splitting avait déjà été introduit par Meszaros et Santambrogio dans [94] pour résoudre l'équation de Fokker-Planck munie d'une contrainte de densité $\rho \leqslant 1$. Leur méthode consiste à suivre l'équation de Fokker-Planck pendant une durée h puis de projeter par rapport à la distance de Wasserstein sur l'ensemble des densités inférieure à 1. Notre méthode permet aussi de contourner une des difficultés listées dans leur article (section 5, variant 3).

La fin du chapitre 5 est dédiée à l'extension de la méthode de splitting dans l'espace de Wasserstein pour des coûts plus généraux. On remplace la distance de Wasserstein d'ordre 2 dans le schéma de JKO par

$$W_c(\rho,\mu) := \inf\left\{ \iint_{\Omega\times\Omega} c(x-y)\, d\gamma(x,y) \,:\, \gamma \in \Pi(\rho,\mu)\right\},$$

où $c : \mathbb{R}^n \to [0,+\infty)$ est une fonction régulière, strictement convexe telle que $c(0) = 0$ et

$$\alpha|x|^q \leqslant c(x) \leqslant \beta(1+|x|^q),$$

pour certains $\alpha, \beta > 0$ et $q > 1$. En utilisant cette distance, on veut résoudre

$$\partial_t \rho - \operatorname{div}(\rho \nabla c^*(F'(\rho))) - \operatorname{div}(\rho U[\rho]) = 0,$$

avec c^* la transformée de Legendre de c. Sans le champs de vecteur U, ce système a déjà été étudié par Otto [97] (dans le cas où le coût est donné par $c(x) = |x|^p$) et par Agueh [1] dans le cas général. Notre méthode consiste à réutiliser le shéma de splitting mais en effectuant la phase de transport pur avec tout le vecteur $U[\rho]$. En supposant des bornes par en dessous et au dessus sur la condition initiale, on arrive à les itérer en temps, nous permettant de controler l'évolution de l'énergie interne lors de la phase de transport pur.

Le chapitre 6 est consacré à l'étude de systèmes où l'interaction se trouve dans le terme de diffusion. Ce champ de recherche a connu une grande activité ces dernières années, voir par exemple [79, 53, 66, 67, 72]. Ici, nous nous interesserons plus à une interaction modélisant la congestion de deux populations. Les modèles de congestion de foule dans le cadre macroscopique ont été proposés par Maury, Roudneff-Chupin et Santambrogio [89] dans le cadre d'une population pour modéliser, par exemple, l'évacuation d'une pièce. Ces modèles consistent à dire que la population suit un vecteur vitesse, mettons le gradient de la distance à une sortie, et ce vecteur est adapté dans les zones où la contrainte $\rho \leqslant 1$ est saturée. Les auteurs prouvent que ce problème a une structure de flot de gradient dans l'espace de Wasserstein. Plus récemment, Meszaros et Santambrogio dans [94] ont proposé un modèle de congestion dure où les individus sont soumis à une diffusion brownienne. Cela revient à montrer l'existence de solution de l'équation

$$\partial_t \rho - \Delta\rho - \operatorname{div}(\rho\nabla(V+p)) = 0, \qquad \rho \leqslant 1,$$

avec p un terme de pression satisfaisant

$$p \geqslant 0, \text{ and } p(1-\rho) = 0.$$

Une variante naturelle de ce système consiste à considérer deux populations, chacune d'entre elles ayant son propre potentiel mais étant sujet à une pression commune. Pour une diffusion linéaire, cela correspond à mettre un bruit brownien sur chaque espèce. La dynamique des mouvements de foules à deux espèces s'exprime par

$$\begin{cases} \partial_t \rho_1 - \Delta\rho_1 - \operatorname{div}(\rho_1(\nabla V_1 + \nabla p)) = 0, \\ \partial_t \rho_2 - \Delta\rho_2 - \operatorname{div}(\rho_2(\nabla V_2 + \nabla p)) = 0, \\ p \geqslant 0,\ \rho_1+\rho_2 \leqslant 1,\ p(1-\rho_1-\rho_2) = 0. \end{cases}$$

Ce système peut être vu comme le flot de gradient pour la distance de Wasserstein produit de

$$\mathcal{E}(\rho_1,\rho_2) := \sum_{i=1}^{2}\int_\Omega (\rho_i\log(\rho_i)+V_i\rho_i) + \int_\Omega \chi_{[0,1]}(\rho_1(x)+\rho_2(x))\,dx.$$

Plus généralement, on va étudier l'existence de solutions pour des systèmes de la forme

$$\begin{cases} \partial_t\rho_1 - \Delta\rho_1 - \operatorname{div}(\rho_1(\nabla V_1 + \nabla F_m'(\rho_1+\rho_2))) = 0, \\ \partial_t\rho_2 - \Delta\rho_2 - \operatorname{div}(\rho_2(\nabla V_2 + \nabla F_m'(\rho_1+\rho_2))) = 0, \end{cases}$$

avec pour $m\in[1,+\infty[$, $F_m:\mathbb{R}^+\to\mathbb{R}$ définie par

$$F_m(x) = \begin{cases} x\log x & \text{if } m=1, \\ \frac{x^m}{m-1} & \text{if } m>1. \end{cases}$$

Ce système est le flot de gradient pour la distance de Wasserstein produit de

$$\mathcal{E}(\rho_1,\rho_2) := \sum_{i=1}^{2}\int_\Omega (\rho_i\log(\rho_i)+V_i\rho_i) + \int_\Omega F_m(\rho_1(x)+\rho_2(x))\,dx.$$

La difficulté est de passer à la limite dans le terme de diffusion croisée. Pour ce faire, on va utilisé l'argument de flow interchange pour obtenir des estimations à la fois sur ρ_i et sur la somme $\rho_1+\rho_2$. À la fin du chapitre, on montrera des simulations numériques faites en utilisant l'algorithme défini chapitre 3.

Dans les systèmes précédement étudiés, il n'y avait pas de réaction et donc la masse était fixée ce qui permettait de travailler dans l'espace des mesures de probabilité muni de la distance de Wasserstein. Dans le chapitre 7, on va s'intéresser au cas des sytèmes de réaction-diffusion où les populations intéragissent entre elles via le terme de réaction. Ces systèmes apparaissent beaucoup en biologie, le modèle le plus simple étant les systèmes diffusifs proie-prédateur [96]. L'analyse de ce type de systèmes est déjà très développée, voir par exemple l'article de Michel Pierre [105] pour une vue d'ensemble sur le sujet. Dans ce chapitre, nous allons étudier deux façons différentes pour étendre la méthode des flots de gradient dans l'espace de Wasserstein aux équations de réaction-diffusion.

La première méthode a été proposée par Kinderlehrer et Walkington dans [70] où ils présentent un algorithme de splitting pour résoudre les équations du type

$$\partial_t\rho - \operatorname{div}(\rho\nabla(F'(\rho)+V)) = f(\rho). \tag{1.3}$$

Le schéma consiste à introduire une densité intermédiaire où la masse est mise à jour puis d'utiliser le schéma de JKO en utilisant la densité intermédiaire comme instant précédent: si on se donne un pas de temps $h>0$ et ρ_h^k la solution au temps hk, on commence par construire

$$\tilde\rho_h^{k+1} := \rho_h^k + hf(\rho_h^k),$$

puis on définit ρ_h^{k+1} comme le minimum, sur les mesures positives de même masse que $\tilde\rho_h^{k+1}$, de

$$\rho\mapsto \frac{1}{2h}W_2^2(\rho,\tilde\rho_h^{k+1}) + \int (F(\rho)+V\rho).$$

Petrelli et Tudorascu démontrèrent dans [103] la convergence de ce schéma vers une solution de (1.3). La démonstration est basée sur un principe du maximum et sur le théorème de Fréchet-Kolmogorov pour obtenir une convergence forte. Dans la première partie de ce chapitre, section 7.1, on va étendre ce résultat aux systèmes avec un terme d'interaction dans la réaction en utilisant le schéma de splitting présenté plus haut. Tout d'abord, on étend le principe du maximum de Petrelli et Tudorascu au cas des systèmes. La convergence forte est retrouvée en utilisant la *bounded Lipschitz distance* pour obtenir de la compacité en temps alors que la compacité en espace est obtenue à l'aide des équations d'Euler-Lagrange associées aux problèmes de minimisation. La

fin de cette section est consacrée aux simulations numériques notamment sur des sytèmes proie-prédateur.

La seconde méthode est basée sur une méthode de splitting sur la distance de Wasserstein-Fisher-Rao. Cette métrique a été introduit récemment par trois équipes différentes [43, 44, 71, 80, 81]. Elle permet d'étendre la distance de Wasserstein aux mesures positives de masses différentes. Dans [63], Gallouët et Monsaingeon ont proposé une méthode de splitting pour résoudre les flots de gradient pour cette métrique. Ils ont remarqué qu'infinitésimalement, la métrique de Wasserstein-Fisher-Rao devrait être la somme orthogonale de la métrique de Wasserstein et celle de Fisher-Rao. Cela les a menés naturellement à diviser une étape de minimisation par rapport à la distance de Wasserstein-Fisher-Rao d'une énergie $\mathcal{E}$ en une sous-étape du schéma de JKO pour $\mathcal{E}$ et une sous-étape de minimisation par rapport à la distance de Fischer-Rao définie par

$$FR(\rho_0,\rho_1) := 4\int_\Omega \left|\sqrt{\frac{d\rho_0}{d\lambda}} - \sqrt{\frac{d\rho_1}{d\lambda}}\right|^2 d\lambda,$$

pour toute mesure de référence λ telle que ρ_0 et ρ_1 soient absolument continues par rapport à λ. Dans [63], ils démontrent que les solutions discrètes ainsi créées convergent vers la solution de

$$\partial_t\rho - \operatorname{div}(\rho\nabla(F'(\rho)+V)) = -\rho(F'(\rho)+V),$$

si $\mathcal{E}$ est de la forme

$$\mathcal{E}(\rho) := \int F(\rho) + \int V\rho.$$

Une variante naturelle de ce schéma est de minimiser des fonctionnelles différentes pour l'étape de Wasserstein et l'étape de Fischer-Rao. En collaboration avec Gallouët et Monsaingeon, nous proposons d'étudier ce type d'équations dans la seconde partie du chapitre 7. Puis on appliquera ce schéma pour retrouver l'existence de solutions faibles pour des modèles de croissance tumorale et les simuler numériquement en utlisant l'algorithme du chapitre 3 pour la partie JKO. Ces modèles ont été étudiés par Perthame, Tang et Vauchelet [102] et permettent de résoudre des équations du type Hele-Shaw.

Le dernier chapitre 8 étudie un système d'équations paraboliques où l'interaction est donnée par le potentiel de Kantorovich d'un problème multi-marges entre toutes les populations. Dans le cadre simple de deux populations, ce modèle a été inspiré d'un article récent de Kinderlehrer, Monsaingeon et Xu [69] où ils proposent une approche par flot de gradient dans l'espace de Wasserstein pour résoudre le système de Poisson-Nernst-Planck

$$\begin{cases} \partial_t\rho - \alpha\Delta\rho^m - \operatorname{div}(\rho\nabla(U+\varphi)) = 0, \\ \partial_t\mu - \beta\Delta\mu^m - \operatorname{div}(\mu\nabla(V-\varphi)) = 0, \\ -\Delta\varphi = \rho - \mu. \end{cases}$$

On s'est intéressé au cas "non linéaire" où ρ et μ sont couplés par l'équation de Monge-Ampère à la place de l'équation de Poisson,

$$\begin{cases} \partial_t\rho - \alpha\Delta\rho^m - \operatorname{div}(\rho\nabla U) - \operatorname{div}(\rho\nabla\varphi) = 0, \\ \partial_t\mu - \beta\Delta\mu^m - \operatorname{div}(\mu\nabla V) - \operatorname{div}(\mu\nabla\varphi^c) = 0, \\ \det(I - D^2\varphi)\mu(Id - \nabla\varphi) = \rho, \end{cases} \tag{1.4}$$

où φ^c est la c-transformée de φ, $\varphi^c(x) = \sup_y |x-y|^2 - \varphi(y)$ et $|x|^2 - \varphi$ est convexe. Le couple (φ,φ^c) est une paire de potentiel de Kantorovich pour le problème $W_2(\rho,\mu)$. Ce type de problème peut apparaître en aménagement urbain. Divers modèles ont été proposés dans le cas statique par exemple voir [25, 26, 36, 31, 32, 109, 108, 111]. Dans un cas simple, on considère une région Ω représentant une ville, une densité d'habitants $\rho \in \mathcal{P}(\Omega)$ et une densité de services $\mu \in \mathcal{P}(\Omega)$. Pour atteindre une configuration optimale, (ρ,μ) doit minimiser une fonctionelle $\mathcal{E}(\rho,\mu)$ modélisant plusieurs aspects. Tout d'abord, ρ et μ veulent minimiser un coût de transport pour venir des zones d'habitations aux zones où se trouvent les services. Ce coût peut être modélisé par $W_2^2(\rho,\mu)$. De

plus, les habitants veulent minimiser un terme de congestion dû au fait que la population ne veut pas se retrouver dans une zone trop peuplée. Ce terme de congestion est modélisé par

$$\mathcal{F}(\rho) := \int_\Omega F(\rho),$$

où $F : \mathbb{R}^+ \to \mathbb{R}^+$ est une fonction continue convexe superlinéaire. Les services, contrairement aux habitants, cherchent à se concentrer pour augmenter leur efficacité et diminuer les coûts de gestion. On modélise ce phénomène par la minimisation de

$$\mathcal{G}(\rho) = \iint_{\Omega\times\Omega} h(|x-y|)\, d\rho(x) d\rho(y),$$

avec h une fonction croissante. Donc ρ et μ veulent minimiser un coût total donnée par

$$\mathcal{E}(\rho,\mu) := W_2^2(\rho,\mu) + \mathcal{F}(\rho) + \mathcal{G}(\mu).$$

Partant du principe qu'une ville est en constante évolution, on s'est intéressé au problème dynamique c'est à dire au flot de gradient par rapport à la distance de Wasserstein produit de $\mathcal{E}$. Formellement, le flot de gradient converge vers une solution d'un système de la forme (1.4). Dans le chapitre 8, on va étudier l'existence des solutions de systèmes plus générales que (1.4), notament en augmentant le nombre de densités et en supposant que chaque population veut minimiser un problème de transport avec des coûts qui peuvent être différents. Cette dernière hypothèse est assez naturelle, par exemple les travailleurs doivent payer l'essence pour se rendre au travail tandis que les entreprises ne s'en soucient pas. On va montrer l'existence de solutions du système suivant

$$\partial_t \rho_i = \Delta P_i(\rho_i) + \operatorname{div}(\rho_i \nabla u_i),\ \rho_{i|t=0} = \rho_{i,0},$$

pour tout $i \in [\![1, l]\!]$, où u_i est un potentiel de Kantorovich associé au problème multi-marges

$$\mathcal{W}_{c_i}(\rho_1,\dots,\rho_l) := \inf\left\{ \int_{\Omega^l} c_i(x_1,\dots,x_l)\, d\gamma(x_1,\dots,x_l) \,:\, \gamma \in \Pi(\rho_1,\dots,\rho_l) \right\},$$

où $\Pi(\rho_1,\dots,\rho_l)$ représente l'ensemble des plans de transport entre $\rho_1,\dots,\rho_l$. Comme on utilise des coûts différents pour chaque densité, ce sytème n'est pas un flot de gradient et on va alors devoir utiliser un schéma de JKO semi-implicite pour le résoudre. On donnera ensuite un résultat d'unicité basée sur la convexité géodésique des problèmes de transport ainsi que des exemples de fonctionnelles vérifiant cette hypothése. Dans la dernière partie, nous donnerons quelques problèmes ouverts qui feront l'objet de recherches ultérieures.

Chapter 2

Introduction to optimal transportation and gradient flows

In this thesis, we wish to extend the usual theory of gradient flows in Wasserstain space to treat systems of parabolic equations with different forms of interaction. These systems appear in various domain as population dynamics, crowd modelling, tumor growth, urban planning etc. The main tool comes from the optimal transport theory. We recall here some results about optimal transportation theory and gradient flows theory. We refer to [116, 117, 112] to a detailed exposition of optimal transport and [4] for gradient flows.

2.1 Optimal transport

2.1.1 Generalities on optimal transport

The optimal transport problem was introduced in 1789 by Gaspard Monge in [95]. The problem is the following: given a pile of sand and a hole with same volume, we want to fill up the hole with the sand minimizing a cost function given by the euclidian distance in [95].

In a more mathematical framework, the pile of sand and the hole are represented by two probability measures ρ and μ defined respectively on complete and separable metric spaces X and Y. In the sequel, we will often work with probability measures defined on the same subset Ω of $\mathbb{R}^n$. The cost function is given by a countinuous (or lower semi continuous) map $c : X \times Y \to [0, +\infty)$. Then the problem consists in finding a map $T : X \to Y$ which pushes ρ to μ i.e $T_{\#}\rho = \mu$ where $T_{\#}\rho$ is called *image measure* or *push-forward* of ρ through T and is defined by

$$(T_{\#}\rho)(A) := \rho(T^{-1}(A)) \text{ for every measurable set } A,$$

or

$$\int_Y \phi \, d(T_{\#}\rho) = \int_X \phi \circ T \, d\rho \text{ for every measurable function } \phi,$$

such that T solves

$$(\mathcal{M}) \qquad \inf \left\{ \int_X c(x, T(x)) \, d\rho(x) \, : \, T_{\#}\rho = \mu \right\}. \tag{2.1}$$

A minimizer of this problem is called *optimal transport map*. We remark that (2.1) does not allow splitting of mass. Indeed, all the mass in x has to be send in $T(x)$ due to the constraint $T_{\#}\rho = \mu$.

Moreover, if X, Y are two subsets of $\mathbb{R}^n$ and ρ and μ are induced by two densities ($d\rho(x) = \rho(x)dx$ and $d\mu(y) = \mu(y)dy$ with $\rho \in L^1(X)$ and $\mu \in L^1(Y)$) assuming ρ, μ and T smooth and T injective, we can rewritte, using the change-of-variables formula, the constraint $T_{\#}\rho = \mu$ as

$$\mu(T(x)) \det(DT(x)) = \rho(x). \tag{2.2}$$

This equation is highly nonlinear and is the main difficulty to prove the existence of a minimizer in (2.1). Indeed, the usual approach is to take a minimizing sequence $(T_k)_k$ and prove a bound on it to have compactness. If we assume that Y is compact or $\mathrm{supp}(\mu)$ is compact, then T_k is bounded in L^∞ which implies the weak* convergence in L^∞ of the sequances to a limit T. Since c is assumed lower semi continuous, we have

$$\int_X c(x,T(x))\,d\rho(x) \leqslant \lim_k \int_X c(x,T_k(x))\,d\rho(x),$$

however, the weak convergence does not permit to pass to the limit in the nonlinear PDE (2.2),

To overcome this problem, Kantorovich introduced, in 1942 (see [68]), a relaxation of Monge's problem allowing the splitting of mass called *Monge-Kantorovich problem*,

$$(\mathcal{MK}) \qquad \inf\left\{\iint_{X\times Y} c(x,y)\,d\gamma(x,y) \,:\, \gamma\in\Pi(\rho,\mu)\right\}, \tag{2.3}$$

where $\Pi(\rho,\mu)$ is the set of probability measure on $X\times Y$ with marginals ρ and μ,

$$\Pi(\rho,\mu) := \left\{\gamma\in\mathcal{P}(X\times Y) \,:\, \pi_{x\#}\gamma=\rho \text{ and } \pi_{y\#}\gamma=\mu\right\},$$

and π_x and π_y are the projections of $X\times Y$ onto X and Y. Elements of $\Pi(\rho,\mu)$ are called *transport plans* between ρ and μ. Contrary to the Monge's problem, the distination of the mass in a point x is not specify and $\gamma(x,y)$, for $\gamma\in\Pi(\rho,\mu)$, coresponds to the amount of mass transferred from x to y and then the mass located in x may split into several parts. This definition generalizes the constraint for the Monge's problem. Given a transport map T, we define $\gamma_T := (Id,T)_\#\rho$ and it is easy to checked that γ_T is in $\Pi(\rho,\mu)$. This leads directly to

$$\inf(\mathcal{MK}) \leqslant \inf(\mathcal{M}).$$

Moreover, $\Pi(\rho,\mu)$ is never empty because $\rho\otimes\mu$ satisfies all the constraints and is tight so using Prokhorov theorem and the lower semi continuity of c, we obtain

Theorem 2.1. *Let X and Y be complete and separable metric spaces, $\rho\in\mathcal{P}(X),\mu\in\mathcal{P}(Y)$ and $c : X\times Y\to\mathbb{R}$ a lower semi continuous function. Then $(\mathcal{MK})$ admits at least one solution.*

Remark 2.2. *This result is false for the Monge's problem: take $\rho=\delta_0$, it can not exists a transport map between ρ and μ if μ is not a single Dirac mass.*

Another interesting aspect of the Monge-Kantorovich problem is the fact it is a linear problem under linear constraints then it is important to study the dual problem in order to exploit relations between dual and primal. Since, the constraint $\gamma\in\Pi(\rho,\mu)$ can be rewritten

$$\sup_{(\varphi,\psi)\in\mathcal{C}_b(X)\times\mathcal{C}_b(Y)} \int_X \varphi\,d\rho + \int_Y \psi\,d\mu - \iint_{X\times Y} (\varphi(x)+\psi(y))\,d\gamma(x,y) = 0,$$

we remark that if $\gamma\notin\Pi(\rho,\mu)$ the supremum becomes $+\infty$, then the Monge-Kantorovich problem becomes

$$\min_{\gamma\in\mathcal{M}^+(X\times Y)} \Bigg(\iint_{X\times Y} c(x,y)\,d\gamma(x,y) + \sup_{(\varphi,\psi)\in\mathcal{C}_b(X)\times\mathcal{C}_b(Y)} \int_X \varphi\,d\rho + \int_Y \psi\,d\mu - \iint_{X\times Y} (\varphi(x)+\psi(y))\,d\gamma(x,y)\Bigg).$$

Assuming we can interchange the infimum and the supremum, we get the dual problem

$$(\mathcal{D}) \qquad \sup\left\{\int_X \varphi\,d\rho + \int_Y \psi\,d\mu \,:\, \varphi\in\mathcal{C}_b(X),\psi\in\mathcal{C}_b(Y),\varphi(x)+\psi(y)\leqslant c(x,y) \text{ on } X\times Y\right\}. \tag{2.4}$$

If $(\mathcal{D})$ admits a solution (φ, ψ), then φ and ψ are called *Kantorovich potentials*. The existence of such a functions is not so obvious and comes from the fact that the supremum in $(\mathcal{D})$ can be taken on $\varphi \in c - conc(X)$ with $c - conc(X)$, the set of function $f : X \to \overline{\mathbb{R}}$ such there exists $g : Y \to \overline{\mathbb{R}}$ such that $f = g^c$ with $g^c : X \to \overline{\mathbb{R}}$ is the c-transform of g defined by

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

Theorem 2.3. *Problem $(\mathcal{D})$ admits a solution (φ, φ^c) where $\varphi \in c - conc(X)$.*

In fact the inversion "inf-sup" can be made rigorous using Fenchel-Rockafellar theorem (see [106]), then we have

Theorem 2.4. *Let X, Y be complete and separable metric spaces and let c be a lower semi-continuous nonnegative function on $X \times Y$ then*

$$\inf(\mathcal{MK}) = \sup(\mathcal{D}).$$

In addition the constraint on the dual problem is saturated a.e with respect to an optimal transport plan of $(\mathcal{MK})$.

Until now, we show the existence of solution for the relaxed problem and the dual of this one. Then we show the equality between $(\mathcal{MK})$ and $(\mathcal{D})$. Now we will come back to the Monge problem. In the sequel, we focus on the case where $X = Y = \Omega$ a subset of $\mathbb{R}^n$ and the cost function c is given by $c(x, y) = h(x - y)$ with h a strictly convex function. In this framework, Brenier (for $h(x) = |x|^2$ in [21]) and in a more general case Gangbo and McCann (h strictly convex in [64]) prove the existence and uniqueness of an optimal transport map.

Theorem 2.5. *Given $\rho, \mu \in \mathcal{P}(\Omega)$, such that the transport cost from ρ to μ is not always infinite. If ρ is absolutely continuous with respect to the Lebesgue measure and $\partial\Omega$ is negligible, then the optimal transport plan of $(\mathcal{MK})$ is unique and is on the form $\gamma_T = (Id, T)_{\#}\rho$. Moreover, there exists a Kantorovich potential φ. The potential φ and T are linked by*

$$T(x) = x - \nabla h^{-1}(\nabla\varphi(x)) \qquad a.e.$$

When the cost is quadratic, $c(x, y) = |x - y|^2$, we recover the result proved by Brenier in [21]. In this case, the map is given by $T(x) = x - \nabla\varphi = \nabla u$. It is easy to show that $u := |x|^2 - \varphi$ is a convex function using the fact that φ is in $c - conc(\Omega)$. Moreover, if ρ and μ are two densities with respect to the Lebesgue measure, then the Brenier's map ∇u satisfies the Monge-Ampère equation

$$\det(D^2 u)\mu(\nabla u) = \rho.$$

This equation is elliptic and highly degenrate. In [27, 28], Caffarelli proved the following result on the regularity of solution of the Monge-Ampère equation.

Theorem 2.6. *Let $\rho, \mu \in \mathcal{C}(\Omega)^{0,\alpha}$ $(0 < \alpha < 1)$ be Hölder-continuous functions on Ω, a convex bounded subset of $\mathbb{R}^n$, which are bounded from above and below by positive constants. Then the unique Brenier solution u belongs to $\mathcal{C}^{2,\alpha}(\Omega) \cap \mathcal{C}^{1,\alpha}(\overline{\Omega})$ and u satisfies the Monge-Ampère equation in the usual sense.*

2.1.2 Wasserstein space

In this section, we consider costs of the form $c(x - y) = |x - y|^p$, $p > 1$, in Ω a subset of $\mathbb{R}^n$. All the results and more details can be found in [116, 117, 112, 4]. We will used problems introduced previously to define a distance on the space of probability measures $\mathcal{P}(\Omega)$. If Ω is unbounded, we define the set

$$\mathcal{P}_p(\Omega) := \left\{ \rho \in \mathcal{P}(\Omega) \, : \, M_p(\rho) := \int_\Omega |x|^p \, d\rho(x) < +\infty \right\},$$

and we note $\mathcal{P}_p^{ac}(\Omega)$ the subset of $\mathcal{P}_p(\Omega)$ of probability measures on Ω absolutely continuous with respect to the Lebesgue measure. We remark that if Ω is bounded $\mathcal{P}_p(\Omega) = \mathcal{P}(\Omega)$, for all $p \in [1, +\infty[$. In the rest of the manuscript, since we will often work in $\mathcal{P}_2(\Omega)$, the second moment $M_2(\rho)$ will be denoted $M(\rho)$.

Definition 2.7. *For all $\rho, \mu \in \mathcal{P}_p(\Omega)$, the p-th Wasserstein distance between ρ and μ is defined by*

$$W_p(\rho,\mu) := \inf\left\{ \iint_{\Omega\times\Omega} |x-y|^p \, d\gamma(x,y) \ : \ \gamma \in \Pi(\rho,\mu) \right\}^{1/p}.$$

Since $\rho, \mu \in \mathcal{P}_p(\Omega)$, the p-th Wasserstein distance is finite then, provided that $\rho \in \mathcal{P}_p^{ac}(\Omega)$, theorem 2.5 gives that $W_2(\rho,\mu)$ admits a unique transport plan induced by a map T and, if $p = 2$, T is the gradient of a convex function.

Note that, since γ is a probability measure, the Hölder inequality implies that for all $1 \leqslant p \leqslant q < +\infty$,

$$\left(\iint_{\Omega\times\Omega} |x-y|^p \, d\gamma(x,y) \right)^{1/p} \leqslant \left(\iint_{\Omega\times\Omega} |x-y|^q \, d\gamma(x,y) \right)^{1/q},$$

which directly implies that $W_p(\rho,\mu) \leqslant W_q(\rho,\mu)$. In addition if Ω is bounded we obtain the opposite inequality,

$$W_q(\rho,\mu) \leqslant \mathrm{diam}(\Omega)^{\frac{q-p}{q}} W_p(\rho,\mu)^{q/p}.$$

Proposition 2.8. *The quantity W_p is indeed a distance on $\mathcal{P}_p(\Omega)$. We called the Wasserstein space of order p this space endowed with this distance.*

The difficulty is to prove the triangle inequality and we need of the next lemma using the disintegration of measures.

Lemma 2.9. *Given two measures $\gamma_1 \in \Pi(\rho,\nu)$ and $\gamma_2 \in \Pi(\nu,\mu)$, there exists a measure $\boldsymbol{\gamma} \in \mathcal{P}(\Omega\times\Omega\times\Omega)$ such that $\pi_{x,y\#}\boldsymbol{\gamma} = \gamma_1$, $\pi_{y,z\#}\boldsymbol{\gamma} = \gamma_2$.*

Proof of the triangle inequality. Let $\rho, \mu, \nu \in \mathcal{P}_p(\Omega)$, γ_1 an optimal transport plan for $W_p(\rho,\nu)$ and γ_2 an optimal transport plan for $W_p(\nu,\mu)$. Applying the previous lemma, there exists $\boldsymbol{\gamma} \in \mathcal{P}(\Omega\times\Omega\times\Omega)$ such that $\pi_{x,y\#}\boldsymbol{\gamma} = \gamma_1$, $\pi_{y,z\#}\boldsymbol{\gamma} = \gamma_2$. Moreover, $\pi_{x,z\#}\boldsymbol{\gamma}$ belongs to $\Pi(\rho,\mu)$, then immediately we get the result from the standard triangle inequality of the L^p distance. □

We have already seen that $W_p(\rho,\mu)$ is equal to is dual formulation,

$$W_p(\rho,\mu) = \sup_{\varphi \in c-conc(\Omega)} \left\{ \int_\Omega \varphi \, d\rho + \int_\Omega \varphi^c \, d\mu \right\}.$$

In the special case where $p = 1$, $\varphi^c = -\varphi$ and

$$W_1(\rho,\mu) := \sup\left\{ \int_\Omega \varphi \, d(\rho-\mu) \ : \ \varphi \in L^1(d|\rho-\mu|) \cap Lip_1(\Omega) \right\},$$

where $Lip_1(\Omega)$ is the set of 1-Lipschitz continuous functions. Then we have the following usefull inequality,

$$\int_\Omega \varphi d(\rho-\mu) \leqslant C W_1(\rho,\mu) \leqslant C W_2(\rho,\mu), \tag{2.5}$$

for all Lipschitz function φ.

We have defined a distance on the probability space $\mathcal{P}_p(\Omega)$, and now, we have to analyse the convergence in the the Wasserstein space with respect to this distance. We start to state a result on the stability of optimality

Proposition 2.10. *Let $(\rho_k), (\mu_k) \subset \mathcal{P}_p(\Omega)$ be two sequences narrowly converging to ρ, μ respectively, and γ_k an optimal transport plan in $W_p(\rho_k,\mu_k)$ such that $W_p(\rho_k,\mu_k)$ is bounded. Then (γ_k) is narrowly relatively compact in $\mathcal{P}(\Omega\times\Omega)$ and any narrow limit point γ is an optimal transport plan in $W_p(\rho,\mu)$ and*

$$W_p(\rho,\mu) \leqslant \liminf_{k\to+\infty} W_p(\rho_k,\mu_k).$$

Now, we state the following theorem which gives equivalence between the convergence with distance W_p and the narrow convergence.

Theorem 2.11. *$\mathcal{P}_p(\Omega)$ endowed with the p-Wasserstein distance is a separable metric space which is complete. A set $\mathcal{K} \subset \mathcal{P}_p(\Omega)$ is relatively compact iff it is p-uniformly integrable and tight. In particular, for a given sequence $(\rho_k) \subset \mathcal{P}_p(\Omega)$ we have*

$$\lim_{k\to+\infty} W_p(\rho_k, \rho) = 0 \Leftrightarrow \begin{cases} \rho_k \text{ narrowly converges to } \rho, \\ M_p(\rho_k) \to M_p(\rho). \end{cases}$$

In some chapter, we have to work in a product of probability space, $\mathcal{P}_p(\Omega)^l$. We define the product distance on $\mathcal{P}_p(\Omega)^l$ by

$$W_p(\boldsymbol{\rho}, \boldsymbol{\mu}) := \left(\sum_{i=1}^{l} W_p^p(\rho_i, \mu_i) \right)^{1/p}.$$

The previous results stated in this section holds for this distance.

2.1.3 Geodesics in Wassertein space and Benamou-Brenier formula

Another important aspect in Wassertein space is the understanding of curves and geodesics with respect to W_p. In the sequel, we focus on the case of the quadratic cost and we refer to [4, 112] for more detailed presentation. In this section we will recall the link between absolutely continuous curves in Wassertein space and solutions of the continuity equation

$$\partial_t \rho + \operatorname{div}(\rho v) = 0.$$

To do so, we need to recall some definitions from the analysis in metric space.

Definition 2.12.
- *A curve $(\rho_t)_{t\in[0,1]}$ in $\mathcal{P}_2(\Omega)$ is said absolutely continuous if there exists $g \in L^1([0,1])$ such that, for all $0 \leqslant s < t \leqslant 1$,*

$$W_2(\rho_s, \rho_t) \leqslant \int_s^t g(\tau)\, d\tau.$$

- *The metric derivative, $|\dot{\rho}|_{W_2}$ of a curve $(\rho_t)_{t\in[0,1]}$ in $\mathcal{P}_2(\Omega)$ is defined by*

$$|\dot{\rho}|_{W_2}(t) := \lim_{h\to 0} \frac{W_2(\rho_{t+h}, \rho_t)}{|h|},$$

provided the limit exists.

Let us remark that if the curve $(\rho_t)_{t\in[0,1]}$ in $\mathcal{P}_2(\Omega)$ is a Lipschitz continuous curve then as a consequence of a result in the same spirit of Rademacher's theorem, the metric derivative exists for a.e $t \in [0,1]$ and we can replace g is the definition of absolute continuity by $|\dot{\rho}|_{W_2}$. Moreover, every absolutely continuous curve can be reparametrized in time and become Lipschitz continuous. Then the previous result holds for every absolutely continuous curve.

Now we can state the theorem of charaterization of absolutely continuous curves in Wasserstein space.

Theorem 2.13. *Let $(\rho_t)_{t\in[0,1]}$ be an absolutely continuous curve in $(\mathcal{P}_2(\Omega), W_2)$. Then for a.e $t \in [0,1]$, there exists a vector field $v_t \in L^2(\rho_t)$ such that the continuity equation $\partial_t\rho_t + \operatorname{div}(\rho_t v_t) = 0$ is satisfies in the sense of distribution and $\|v_t\|_{L^2(\rho_t)} \leqslant |\dot{\rho}|_{W_2}(t)$ for a.e $t \in [0,1]$.*

Conversely, if a narrowly continuous curve $\rho_t : [0,1] \to \mathcal{P}_2(\Omega)$ satisfies the continuity equation for some vector field v_t with $\|v_t\|_{L^2(\rho_t)} \in L^1([0,1])$, then ρ_t is absolutely continuous and $|\dot{\rho}|_{W_2}(t) \leqslant \|v_t\|_{L^2(\rho_t)}$, for a.e $t \in [0,1]$.

Now assuming that Ω is a convex subset of $\mathbb{R}^n$, McCann introduced in [93] the McCann interpolation between ρ_0 and ρ_1. It consists in taking γ an optimal transport plan between ρ_0 and ρ_1 for the 2-Wasserstein distance and define $\rho_t := ((1-t)x + ty)_{\#}\gamma$, for all $t \in [0,1]$. It is well known that this interpolation is a *constant speed geodesic* i.e

$$W_2(\rho_t, \rho_s) = (t-s)W_2(\rho_0, \rho_1), \text{ for all } 0 \leqslant s \leqslant t \leqslant 1.$$

This implies that the Wasserstein space is a *geodesic space*,

$$W_2(\rho_0,\rho_1)=\min\left\{\int_0^1 \mid \dot{\rho}\mid_{W_2}(t)\,dt\right\},$$

where the minimum is taken on the asolutely continuous curve $(\rho_t)_{t\in[0,1]}$ such that $\rho_{t=0}=\rho_0$ and $\rho_{t=1}=\rho_1$.

Then inspired by these results and by problems from the fluid mechanics, in the seminal paper, [11], Benamou and Brenier introduced a dynamic formulation of the Wassertein distance. This formulation is called the *Benamou-Brenier formula*

$$W_2^2(\rho_0,\rho_1)=\inf\left\{\int_0^1\int_\Omega |v_t|^2\,d\rho_t dt\,:\,\partial_t\rho_t+\mathrm{div}(\rho_t v_t)=0,\,\rho_{t=0}=\rho_0,\rho_{t=1}=\rho_1\right\}.$$

Now setting $E_t=\rho_t v_t$, the Benamou-Brenier formula can be rewritten as

$$W_2^2(\rho_0,\rho_1)=\inf\left\{\int_0^1\int_\Omega \psi(d\rho/d\mathcal{L},dE/d\mathcal{L})dxdt\,:\,\partial_t\rho_t+\mathrm{div}(E_t)=0,\,\rho_{t=0}=\rho_0,\rho_{t=1}=\rho_1\right\},$$

where $\psi:\mathbb{R}^{n+1}\to\mathbb{R}\cup+\infty$ is defined by

$$\psi(r,m):=\begin{cases}\frac{|m|^2}{r} & \text{if } (r,m)\in]0,+\infty[\times\mathbb{R}^n,\\ 0 & \text{if } (r,m)=(0,0),\\ +\infty & \text{otherwise,}\end{cases}$$

as in [11] and where $d\sigma/d\mathcal{L}$ is the Radon-Nikodym derivative of σ with respect to $\mathcal{L}_{|[0,T]\times\Omega}$. The advantage of this formulation is the convexity with respect to ρ and E. This formulation has been used in [11] to introduced a numerical scheme based on an augmented Lagrangian method. In chapter 3, we extend this method to solve numerically gradient flows in Wasserstein space.

2.2 Gradient flow theory in Wasserstein space

Here we recall the general theory on gradient flows and the connection with PDE's. First we give the definition in an euclidian setting and how we can extend this in a metric space. This exposition comes from the textbook of Santambrogio [112] and we refer to the textbook of Ambrosio, Gigli and Savaré [4] for a complete and detailed reference on the subject.

2.2.1 Gradient flows in metric space

In the euclidian case, given a function $\mathcal{F}:\mathbb{R}^n\to\mathbb{R}$ and a point $x_0\in\mathbb{R}^n$, we said that a curve x is a *gradient flow* of $\mathcal{F}$ starting at x_0 if at each time $x(t)$ goes in the direction where $\mathcal{F}$ decreases the most. More precisely, x is the solution of the Cauchy problem

$$\begin{cases} x'(t)=-\nabla\mathcal{F}(x(t)) & t>0,\\ x(0)=x_0.\end{cases}\tag{2.6}$$

If $\mathcal{F}$ is $\mathcal{C}^{1,1}(\mathbb{R}^n)$ then the Cauchy-Lipschitz theory implies that there exists a unique solution. However, this definition is very rigid because, we need strong assumption on the energy $\mathcal{F}$ and we need a gradient structure on the space.

An interesting remark is that the ODE (2.6) can be seen as the optimality condition of a minimizing problem. Indeed, given a time step $h>0$, using the Euler-implicit scheme of (2.6), we obtain

$$\frac{x_h^{k+1}-x_h^k}{h}=-\nabla\mathcal{F}(\rho_h^{k+1}),$$

for all $k\geqslant 0$. x_h^{k+1} is an approximation of the solution of (2.6) at time $h(k+1)$ and can be seen as the minimizer of

$$x\mapsto\frac{1}{2h}|x-x_h^k|^2+\mathcal{F}(x).$$

Then the sequence $(x_h^k)_k$ can be construct by induction, with $x_h^0 = x_0$, under weaker assumption on $\mathcal{F}$ (coercivity and lower semi-continuity). Moreover, we know that a piecewise constant interpolation of this sequence goes to converge to the solution of (2.6), when $h \searrow 0$.

Another advantage of this interpretation is that it can be easily generalized to metric space (X, d). Given a time step $h > 0$, a starting point $x_0 \in X$ and an energy functional $\mathcal{F} : X \to \mathbb{R}$, we define by induction a sequence $(x_h^k)_k$, by $x_h^0 = x_0$ and for all $k \geqslant 0$,

$$x_h^{k+1} \in \operatorname*{argmin}_{x \in X} \frac{1}{2h} d(x, x_h^k)^2 + \mathcal{F}(x).$$

Then we denote x_h the piecewise constant interpolation of this sequence,

$$x_h(t) := x_h^{k+1} \text{ if } t \in (kh, (k+1)h].$$

De Giorgi introduced in [51] the notion of *minimizing movements*:

Definition 2.14. *A curve $x : [0, T] \to X$ is said to be a Minimizing Movement if there exists a sequence of time steps $h_j \searrow 0$ such that the piecewise constant interpolations x_{h_j} uniformly converge to x on $[0, T]$.*

In the sequel, we will use this definition of minimizing movements for the one of gradient flow in a metric space. We observe that there exist others definitions for gradient flow which need stronger assumption on the energy $\mathcal{F}$. In several chapters, we will use the Evolution Variational Inequality (EVI) which needs convexity assumption on $\mathcal{F}$. The definition is given in chapter 5 and we refer to [4] for the equivalence of definitions in this case.

2.2.2 Application to Wasserstein space

Since we saw that the Wasserstein space $(\mathcal{P}_2(\Omega), W_2)$ is a metric space, we can apply this method in this space. It is well-known that many parabolic PDEs can be recovered as the gradient flows of well-chosen functionals. The seminal papers on this subject is due to Jordan, Kinderlehrer and Otto, see [65], where they proved that the Fokker-Planck equation is the gradient flow in the Wasserstein space of the energy

$$\mathcal{E} : \rho \mapsto \int_{\mathbb{R}^n} \rho \log(\rho) + \int_{\mathbb{R}^n} V \rho.$$

Their proof is based on the scheme of minimizing movements called *JKO scheme* in the Wasserstein setting: for a time step $h > 0$, and an initial condition $\rho_0 \in \mathcal{P}_2(\mathbb{R}^n)$, we define by induction a sequence $(\rho_h^k)_k$ as $\rho_h^0 := \rho_0$ and for all $k \geqslant 0$,

$$\rho_h^{k+1} := \operatorname*{argmin}_{\rho \in \mathcal{P}_2(\mathbb{R}^n)} \left\{ \frac{1}{2h} W_2^2(\rho_h^k, \rho) + \mathcal{E}(\rho) \right\}.$$

They proved that this sequence is well defined using the control of the Entropy by the second moment. Then the optimality condition of this problem at step $k+1$,

$$\nabla \log(\rho_h^{k+1}) + \nabla V = \frac{\nabla \varphi_h^{k+1}}{h} = \frac{I - T_h^{k+1}}{h} \qquad \rho_h^{k+1} - a.e,$$

where T_h^{k+1} and φ_h^{k+1} are respectively the optimal transport and a Kantorovich potential associated to $W_2(\rho_h^{k+1}, \rho_h^k)$. Then, they define a discreet velocity

$$v_h^{k+1} := \frac{T_h^{k+1} - I}{h} := -\frac{\nabla \rho_h^{k+1}}{\rho_h^{k+1}} - \nabla V.$$

They proved that the constant piecewise interpolations of $(\rho_h^k)_k$ and $(v_h^{k+1})_k$ solve

$$\partial_t \rho_h + \operatorname{div}(\rho_h v_h) = \mathcal{O}(h),$$

and converge to ρ and $v := -\frac{\nabla \rho}{\rho} - \nabla V$, which imply that

$$\partial_t \rho = -\operatorname{div}(\rho v) = \Delta \rho + \operatorname{div}(\rho \nabla V).$$

The theory of gradient flows in Wasserstein space has been developped rapidly in the last twenty years with many applications for example to porous media equation [98], aggregation equation [37], double degenerate diffusion equations [97], general degenerate parabolic equation [1] etc. The reference textbook of Ambrosio, Gigli and Savaré gives a very detailed account of this theory.

Chapter 3

An augmented Lagrangian approach to Wasserstein gradient flows

Taking advantage of the Benamou-Brenier dynamic formulation of optimal transport, we propose a convex formulation for each step of the JKO scheme for Wasserstein gradient flows which can be attacked by an augmented Lagrangian method which we call the ALG2-JKO scheme. We test the algorithm in particular on the porous medium equation. We also consider a semi implicit variant which enables us to treat nonlocal interactions.

This chapter is based on a joint work with J-D. Benamou and G. Carlier (see [17]) and we will use this method to illustrate theorical results proved in the other chapters.

3.1 Introduction

It is well-known since the seminal work of Jordan Kinderlehrer and Otto [65] that the Fokker-Planck equation

$$\partial_t \rho = \Delta \rho + \mathrm{div}(\rho \nabla V),\ \rho|_{t=0} = \rho_0 \tag{3.1}$$

where the initial condition ρ_0 is a probability density may be viewed as the Wasserstein gradient flow of the (relative) entropy functional

$$S_V(\rho) := \int_{\mathbb{R}^d} \rho(x) \log\Big(\frac{\rho(x)}{e^{-V(x)}}\Big) \mathrm{d}x. \tag{3.2}$$

More generally, given an internal energy E, a potential V and an interaction potential W, evolution equations of the form

$$\partial_t \rho = \mathrm{div}(\rho \nabla (E'(\rho) + V + W \star \rho)),\ \rho|_{t=0} = \rho_0 \tag{3.3}$$

is the Wasserstein gradient flow of the energy

$$\mathcal{E}(\rho) := \int_{\mathbb{R}^d} E(\rho(x)) \mathrm{d}x + \int_{\mathbb{R}^d} V(x)\rho(x) \mathrm{d}x + \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x-y)\rho(x)\rho(y) \mathrm{d}x \mathrm{d}y.$$

For instance, if $E(\rho) = \frac{1}{m-1}\rho^m$ and $V = W = 0$ one in particular recovers the porous medium equation $\partial_t \rho = \Delta \rho^m$, see the seminal work of Otto [98]. Convolution terms $\nabla W \star \rho$ in 3.3 arise naturally in aggregation equations [37] and models of granular media [41], [42].

The celebrated Jordan-Kinderlehrer-Otto (henceforth JKO) scheme consists, given a time-step $\tau > 0$ in constructing inductively, starting from ρ_0 a sequence of probability measures ρ^k by the implicit Euler scheme:

$$\rho^{k+1} \in \underset{\rho \in \mathcal{P}_2}{\mathrm{argmin}} \left\{ \frac{1}{2\tau} W_2^2(\rho, \rho^k) + \mathcal{E}(\rho) \right\} \tag{3.4}$$

where $\mathcal{P}_2$ denotes the set of probability measures on $\mathbb{R}^d$ having finite second moments and W_2^2 is the squared 2-Wasserstein distance defined for every $(\rho, \nu) \in \mathcal{P}_2 \times \mathcal{P}_2$ by

$$W_2(\rho, \nu) := \inf_{\gamma \in \Pi(\rho,\nu)} \Big\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mathrm{d}\gamma(x, y) \Big\}^{\frac{1}{2}}$$

where $\Pi(\mu, \nu)$ is the set of transport plans between ρ and ν i.e. the set of Borel probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ having μ and ν as marginals. When $\mathcal{E} = S_V$ is given by 3.2, Jordan, Kinderlehrer and Otto [65] proved that one recovers the solution of the Fokker-Planck equation 3.1 by letting τ tend to 0 in the JKO scheme. Similar convergence results hold for the more general equation 3.3 under suitable assumptions on E, V and W. The theory of Wasserstein gradient flows is by now well-developed and it is detailed in the textbooks of Ambrosio, Gigli and Savaré [4], Villani [116], [117] and Santambrogio [112].

We remark that the JKO scheme 3.4 is constructive and it is very natural and tempting to try to apply it for numerical purposes. The positivity, mass conservation and energy dissipation are inbuilt in the JKO scheme and non trivial to preserve with non-linear finite-difference or finite volume schemes (see [29] and references therein). Also some JKO gradient flows, like congested crowd motions [89], cannot be formulated as nonlinear PDEs and the JKO semi-discretisation is the only numerical option.

A serious difficulty with this approach is in the Wasserstein term which involves solving a costly optimal transport problem at each step. In dimension one, this is not really an issue since the optimal transport is essentially a rearrangement problem, and in fact, this 1-D numerical approach was proposed in the early work of Kinderlehrer and Walkington [70] and was used repeatedly. See in particular the recent work of Osberger and Matthes [88] for application to fourth-order evolution PDEs of thin films type. In higher dimensions, the optimisation problem in 3.4 is much more complicated because the optimal transport is given by Brenier's map, the gradient of a convex potential which solves some Monge-Ampère equation.

At least three categories of approaches have been followed to solve 3.4 numerically. A first "Lagrangian" strategy based on Brenier's Theorem is to formulate the problem in terms of the transport map or its potential instead of the density ρ to avoid dealing with the positivity and mass constraints. It also allows a more consistent discretisation of the mass when the density concentrates or dilates. This is done for instance in Carrillo and Moll [40] who proposed a Lagrangian scheme, based on a gradient flow for evolving diffeomorphisms (not necessarily the optimal transport maps) related to a system of evolution equations which is very nonlinear since it involves cofactors. Düring, Matthes and Milišić [59] and Osberger and Matthes [87] used a Galerkin discretisation of the potential. As illustrated in [38], where another Lagrangian method is introduced, a difficulty with the Lagrangian approach is the construction of a discrete density to be used in the internal energy. A semi-discrete solution to this problem has been proposed in [9] based on optimal maps, a discretisation of the Monge-Ampère operator and techniques of computational geometry. This method is provably convergent and enables one to use of a Newton method. Note that using monotone finite difference Monge-Ampère solvers as introduced in [16], [15] could be an option but it does not seem to have been tried.

A second strategy, which is Eulerian, is to use the Monge-Kantorovich linear relaxation of the Wasserstein distance in 3.4. The size of the discretisation is very limited by the linear programming approach. However, Peyré [104] recently generalised entropic regularisation techniques that are computationally efficient in optimal transport [14] to treat JKO gradient flows.

In the present paper we investigate a third approach, also Eulerian, based on replacing the Wasserstein distance with the Benamou-Brenier formulation [11]. This idea has already been used in [22], [23], [8], [10] either for JKO steps or in optimisation problems where the Wasserstein distance intervenes. Our original contribution, initiated in [13] on a general class of optimal transport problems and variational mean field games [76], is to extend this convex reformulation and its augmented Lagrangian numerical resolution based on the algorithm ALG2 of Glowinski and Fortin [61] to solve a succession of problems of the form 3.4, we will call this method ALG2-JKO. We also show that the method can be adapted to treat systems (which are not necessarily gradient flows) using the relaxation introduced in [54] by Di Francesco and Fagioli and extended to the diffusive case in [35].

The Benamou-Brenier formulation induces an extra time dimension in each of the JKO steps. The resulting extra cost because of the discretisation of the inner (Benamou-Brenier) time dimension is usually considered a draw back. In the ALG2-JKO scheme however, since the successive JKO density time snapshots are close only a very few inner timesteps are needed in practice. The ALG2 augmented Lagrangian method is very robust, can deal with non-smooth energies but remains a proximal splitting first order method and converges slowly [99].

The chapter is organized as follows. In section 3.2, we describe the ALG2-JKO scheme. In section 3.3, we illustrate the algorithm on two examples: the porous medium equation and a model of crowd motion with diffusion introduced by Santambrogio and Mészáros in [94].

3.2 The ALG2-JKO scheme

Let us consider one step of the JKO scheme 3.4 in the case where the energy $\mathcal{E}$ is of the form

$$\mathcal{E}(\rho) := \int_{\mathbb{R}^d} E(\rho(x))\mathrm{d}x + \int_{\mathbb{R}^d} V(x)\rho(x)\mathrm{d}x$$

with E a convex internal energy (typical cases being the entropy or a convex power), which corresponds to the time discretization of the PDE:

$$\partial_t \rho = \mathrm{div}(\rho\nabla(E'(\rho) + V)),\ \rho|_{t=0} = \rho_0. \tag{3.5}$$

Our goal is to rewrite 3.4 as a tractable convex problem. To do so, we use the Benamou-Brenier dynamic formula [11] to rewrite W_2^2 as

$$W_2^2(\rho,\nu) := \inf\left\{\int_0^1\int_{\mathbb{R}^d} \frac{|m_t(x)|^2}{\mu_t(x)}\mathrm{d}x\mathrm{d}t\ :\ \partial_t\mu + \mathrm{div}(m) = 0,\ \mu|_{t=0,1} = \rho,\nu\right\} \tag{3.6}$$

which is a convex variational program (it is implicit that the energy above is set to $+\infty$ whenever μ becomes negative or when $\mu = 0$ and $m \neq 0$ so that momentum m can be written as $m = \mu v$ that is m vanishes where μ does and then $|m|^2/\mu = \mu|v|^2$ is the kinetic energy).

Thanks to 3.6 one can rewrite one step of the JKO scheme 3.4 as the convex minimization:

$$\inf_{(\mu_t,m_t,\mu_1=\mu_t(1,.))} \frac{1}{2\tau}\int_0^1\int_{\mathbb{R}^d} \frac{|m_t(x)|^2}{\mu_t(x)}\mathrm{d}x\mathrm{d}t + \mathcal{E}(\mu_1) \tag{3.7}$$

subject to the constraints that $\mu \geq 0$, $m = 0$ when $\mu = 0$ and the linear constraint

$$\partial_t\mu + \mathrm{div}(m) = 0,\ \mu|_{t=0} = \rho^k. \tag{3.8}$$

One then recovers $\rho^{k+1} = \mu_1$ (and actually even an interpolation $(\mu_t)_{t\in[0,1]}$ between ρ^k and ρ^{k+1}).

Of course we can consider variants, for instance the periodic (in space) case or the case of a smooth bounded domain Ω of $\mathbb{R}^d$. In the latter case, we have to supplement the PDE 3.5 with the Neumann boundary condition:

$$\nabla(E'(\rho) + V))\cdot\nu = 0,\ \text{on } \partial\Omega \tag{3.9}$$

this amounts to modify 3.7-3.8 as

$$\inf_{(\mu_t,m_t)} \frac{1}{2}\int_0^1\int_{\Omega} \frac{|m_t(x)|^2}{\mu_t(x)}\mathrm{d}x\mathrm{d}t + \tau\mathcal{E}(\mu_1) \tag{3.10}$$

subject to the constraints that $\mu \geq 0$, $m = 0$ when $\mu = 0$ and the linear constraint

$$\partial_t\mu + \mathrm{div}(m) = 0,\ \mu|_{t=0} = \rho^k,\ m\cdot\nu = 0 \text{ on } \partial\Omega. \tag{3.11}$$

3.2.1 Augmented Lagrangian formulation

Convex time-dependent problems like 3.10 subject to a divergence constraint 3.11 appear in various contexts, they are actually particular cases of deterministic Mean-Field Games (a class of games with a continuum of players introduced by Lions and Lasry [74], [75]). Such problems can be solved by Augmented Lagrangian methods, see in particular [13] for applications to Mean-Field Games, Papadakis, Peyré and Oudet [99] for connections with proximal schemes and Buttazzo, Jimenez and Oudet [24] for applications to congested transport. We now recall the principle of the Augmented Lagrangian approach and explain how to use it in the JKO framework.

As was observed by Benamou and Brenier [11] the convex lower semicontinuous 1-homogeneous function defined for $(\mu, m) \in \mathbb{R} \times \mathbb{R}^d$ by:

$$\Phi(\mu, m) := \begin{cases} \frac{|m|^2}{2\mu}, \text{ if } \mu > 0, \\ 0, \text{ if } \mu = 0 \text{ and } m = 0 \\ +\infty, \text{ otherwise.} \end{cases}$$

is the support function of the convex set

$$K := \{(a, b) \in \mathbb{R}^{d+1}, \ a + \frac{1}{2}|b|^2 \leq 0\} \tag{3.12}$$

i.e.

$$\Phi(\mu, m) = \sup_{(a,b)\in K} \{a\mu + b \cdot m\}.$$

Rewriting 3.10-3.11 in Lagrangian form as

$$\inf_{\sigma=(\mu,m,\mu_1)} \Big\{ \int_0^1 \int_\Omega \Phi(\mu, m) + \tau \mathcal{E}(\mu_1) + \\ \sup_\phi \Big\{ \int_\Omega \phi(1, .)\mu_1 - \int_\Omega \phi(0, .)\rho^k - \int_0^1 \int_\Omega (\partial_t \phi \mu + \nabla \phi \cdot m) \Big\} \Big\}$$

and then switching the inf and the sup and using the fact that the Legendre transform of Φ is 0 on K and $+\infty$ outside, we formally obtain (see for instance [24] for a rigorous derivation) that the convex problem 3.10-3.11 is dual to:

$$\inf_{\phi=\phi(t,x)} \{ \int_\Omega \phi(0, .)\rho^k + \tau \mathcal{E}^* \Big(- \frac{\phi(1, .)}{\tau} \Big) \ : \ \partial_t \phi + \frac{1}{2}|\nabla \phi|^2 \leq 0\} \tag{3.13}$$

where $\mathcal{E}^*$ is the Legendre tranform of $\mathcal{E}$ (extended by $+\infty$ on $(-\infty, 0]$):

$$\mathcal{E}^*(c) := \sup_{\mu \geq 0} \{ \int_\Omega ((c(x) - V(x))\mu(x) - E(\mu(x)) \mathrm{d}x\} = \int_\Omega \mathcal{E}^*(x, c(x)) \mathrm{d}x$$

where, slightly abusing notations, we have set

$$\mathcal{E}^*(x, c) := \sup_{\mu \geq 0} \Big\{ (c - V(x))\mu - E(\mu) \Big\}.$$

We then rewrite the dual as

$$\inf_{\phi=\phi(t,x)} J(\phi) := F(\phi) + G(\Lambda \phi) \tag{3.14}$$

where

$$\Lambda \phi := (D\phi, -\phi(1, .)) = ((\partial_t \phi, \nabla \phi), -\phi(1, .)), \ F(\phi) = \int_\Omega \phi(0, .)\rho^k$$

and for $q = (a, b, c)$

$$G(q) = \int_0^1 \int_\Omega \chi_K(a, b) \mathrm{d}x \mathrm{d}t + \tau \mathcal{E}^* \Big(\frac{c}{\tau} \Big)$$

where χ_K denotes the indicator function

$$\chi_K(a,b) = \begin{cases} 0, \text{ if } (a,b) \in K \\ +\infty, \text{otherwise.} \end{cases}$$

Now the variables $\sigma := (\mu, m, \mu_1)$ play the role of Lagrange multipliers associated to the constraint $q = \Lambda\phi$ i.e. $a = \partial_t\phi$, $b = \nabla\phi$ and $c = -\phi(1,.)$, note in particular that μ_1 is a multiplier associated to the constraint $c = -\phi(1,.)$ it coincides with $\mu(1,.)$ for the saddle-point but not necessarily along the iterations of the augmented Lagrangian algorithm below.

The primal-dual extremality relations are formally equivalent to finding a saddle-point of the Lagrangian

$$L(\phi, q, \sigma) := F(\phi) + G(q) + \sigma \cdot (\Lambda\phi - q), \tag{3.15}$$

in the sense that (ϕ, σ) satisfies the optimality conditions of 3.14 and 3.10-3.11 respectively if and only if

$$(\phi, q, \sigma) = (\phi, \Lambda\phi, \sigma)$$

is a saddle-point of L. Now for $r > 0$, we consider the augmented Lagrangian function

$$L_r(\phi, q, \sigma) := F(\phi) + G(q) + \sigma \cdot (\Lambda\phi - q) + \frac{r}{2}|\Lambda\phi - q|^2 \tag{3.16}$$

where $q = (a, b, c)$, $\sigma = (\mu, m, \mu_1)$,

$$\sigma \cdot (\Lambda\phi - q) = \int_0^1 \int_\Omega \Big(\mu(t,x)(\partial_t\phi(t,x) - a(t,x)) + m(t,x) \cdot (\nabla\phi(t,x) - b(t,x))\Big)\mathrm{d}x\mathrm{d}t + \int_\Omega \mu_1(x)(-\phi(1,x) - c(x))\mathrm{d}x$$

and

$$|\Lambda\phi - q|^2 = \int_0^1 \int_\Omega (|\partial_t\phi(t,x) - a(t,x)|^2 + |\nabla\phi(t,x) - b(t,x)|^2)\mathrm{d}x\mathrm{d}t + \int_\Omega (\phi(1,x) + c(x))^2\mathrm{d}x$$

and recall (see for instance [61], [62]) that being a saddle-point of L is equivalent to being a saddle-point of L_r.

The augmented Lagrangian algorithm ALG2 consists, starting from (ϕ^0, q^0, σ^0) to generate inductively a sequence (ϕ^n, q^n, σ^n) as follows:

- **Step 1:** minimization with respect to ϕ:

$$\phi^{n+1} := \operatorname*{argmin}_\phi \Big\{F(\phi) + \sigma^n \cdot \Lambda\phi + \frac{r}{2}|\Lambda\phi - q^n|^2\Big\}, \tag{3.17}$$

- **Step 2:** minimization with respect to q:

$$q^{n+1} := \operatorname*{argmin}_q \Big\{G(q) - \sigma^n \cdot q + \frac{r}{2}|\Lambda\phi^{n+1} - q|^2\Big\}, \tag{3.18}$$

- **Step 3:** update the multiplier by the gradient ascent formula

$$\sigma^{n+1} = \sigma^n + r(\Lambda\phi^{n+1} - q^{n+1}). \tag{3.19}$$

The convergence of ALG2 to a saddle-point is well documented see in particular the general results of Bertsekas and Eckstein [60] in finite dimensions. We therefore have to understand that in the problems above, we have already projected the potentials in 3.14 on a finite-dimensional space of finite-elements and therefore deal with a finite-dimensional problem for which existence of a saddle-point is rather standard and convergence follows from [60]. Once we have reached a minimizer $\sigma = (\mu, m, \mu_1)$ for 3.10-3.11 by ALG2, we recover the density of a single JKO step by $\rho^{k+1} = \mu_1$.

3.2.2 Details for the three steps

Step 1 corresponds to a linear elliptic problem in t and x,

$$-r\Delta_{t,x}\phi^{n+1} = \operatorname*{div}_{t,x}((\mu^n, m^n) - r(a^n, b^n)), \text{ in } (0,1)\times\Omega, \tag{3.20}$$

together with the boundary conditions

$$r\partial_t\phi^{n+1}(0,.) = \rho^k - \mu^n(0,.) + ra^n(0,.), \tag{3.21}$$

$$r(\partial_t\phi^{n+1}(1,.) + \phi^{n+1}(1,.)) = \mu_1^n - \mu^n(1,.) + r(a^n(1,.) - c^n(.)), \tag{3.22}$$

and

$$r\frac{\partial\phi^{n+1}}{\partial\nu} + (m^n - rb^n)\cdot\nu = 0 \text{ on } \partial\Omega. \tag{3.23}$$

(or periodic boundary conditions if Ω is replaced by the flat torus).

Step 2 splits into two convex pointwise (i.e. for every t and x) minimization subproblems, the first one (minimization with respect to (a,b)) is a projection problem onto the parabola K:

$$\inf_{(a,b)\in K} |D\phi^{n+1}(t,x) + \frac{1}{r}(\mu^n(t,x), m^n(t,x)) - (a,b)|^2 \tag{3.24}$$

i.e.

$$(a^{n+1}(t,x), b^{n+1}(t,x)) = P_K\Big(D\phi^{n+1}(t,x) + \frac{1}{r}(\mu^n(t,x), m^n(t,x))\Big)$$

where the projection P_K onto K is explicit (see [11] or [99]):

$$P_K(\alpha,\beta) = \begin{cases} (\alpha,\beta), \text{ if } (\alpha,\beta)\in K, \\ (\alpha-\lambda, \frac{\beta}{1+\lambda}), \text{ with } \lambda>0 \text{ root of 3.25 otherwise} \end{cases}$$

where 3.25 is the cubic equation (with a single positive root if $(\alpha,\beta)\notin K$):

$$\alpha(1+\lambda)^2 - \lambda(1+\lambda)^2 + \frac{1}{2}|\beta|^2 = 0. \tag{3.25}$$

The second subproblem gives the update for c which is obtained by solving for each $x\in\Omega$

$$c^{n+1}(x) = \operatorname*{argmin}_{c\in\mathbb{R}}\Big\{\frac{r}{2}|\phi^{n+1}(1,x) - \frac{1}{r}\mu_1^n(x) + c|^2 + \tau\mathcal{E}^*\Big(x, \frac{c}{\tau}\Big)\Big\}. \tag{3.26}$$

Remark 3.1. *Given a convex lower semicontinuous function $f\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, we recall that the proximal operator of f, prox_f is defined by*

$$\operatorname{prox}_f(y_0) := \operatorname*{argmin}_{y\in\mathbb{R}^n}\Big\{\frac{1}{2}|y-y_0|^2 + f(y)\Big\},\ \forall y_0\in\overset{n}{\mathbb{R}}$$

so that 3.26 can be rewritten as

$$c^{n+1}(x) = \operatorname{prox}_{\frac{\tau}{r}\mathcal{E}^*(x,\frac{\cdot}{\tau})}\Big(-\phi^{n+1}(1,x) + \frac{1}{r}\mu_1^n(x)\Big).$$

Thanks to the well-known (and actually easy to check) Moreau's identity

$$\operatorname{prox}_f(y) = y - \operatorname{prox}_{f^*}(y), \tag{3.27}$$

we see that it is not necessary to compute E^* to solve (3.26) if the computation of prox_E turns out to be easier.

Step 3 is an explicit update which may be detailed as

$$\mu^{n+1} = \mu^n + r(\partial_t \phi^{n+1} - a^{n+1}), \tag{3.28}$$

$$m^{n+1} = m^n + r(\nabla \phi^{n+1} - b^{n+1}), \tag{3.29}$$

$$\mu_1^{n+1} = \mu_1^n - r(\phi^{n+1}(1,.) + c^{n+1}(.)). \tag{3.30}$$

Note that only the minimization with respect to c 3.26 in step 2 depends on the form of the energy $\mathcal{E}$, we shall give details for this step in each application given in the sequel.

In the discretisation of this algorithm, we use $P2$ finite elements (in time and space) for ϕ and $P1$ finite elements for σ so that in 3.24, in fact, one has to understand $D\phi^{n+1}$ as its projection onto $P1$ finite elements. It was implemented in FreeFeem++[1]. In practice, a discretization with 32×32 triangles in space and 4 inner timesteps needs a few hundreds iterations of ALG2 for each JKO time step. This is a few minutes on a standard laptop. Larger discretizations can be done using FreeFem mpi version which uses for instance MUMPS parallel linear solver[2].

3.3 Applications

We now present two appplications of the ALG2-JKO scheme: the first one deals with the porous medium equation and the second one with a diffusive model of crowd motion recently introduced in [94]. In the sequel, all the simulations are made using a discretization 16×16 in space. We use 8 internal time steps in each JKO step and make sure the ALG2 iterations converge "reasonably" (a 10^{-6} tolerance is prescribed in the optimality system). There are 200 JKO steps ($\tau = 0.01$).

3.3.1 Application to the porous medium equation

The porous medium equation

$$\partial_t \rho = \mathrm{div}(\rho \nabla (\frac{m}{m-1}\rho^{m-1} + V)) = \Delta \rho^m + \mathrm{div}(\rho \nabla V) \tag{3.31}$$

corresponds to

$$E(\rho) := \frac{1}{m-1}\rho^m,\ \mathcal{E}(\rho) = \int_\Omega E(\rho) + \int_\Omega \rho V.$$

In this case

$$\mathcal{E}^*(c) = \Big(\frac{m-1}{m}\Big)^{\frac{m}{m-1}} \int_\Omega \Big((c(x) - V(x))_+\Big)^{\frac{m}{m-1}} \mathrm{d}x$$

and then 3.26 consists in a pointwise minimization problem: given $x \in \Omega$, setting $\overline{c} = -\phi^{n+1}(1,x) + \frac{\mu_1^n(x)}{r}$ and $V = V(x)$, we have to solve

$$\inf_{c \in \mathbb{R}} \Big\{ \frac{1}{2}|c - \overline{c}|^2 + \theta \Big((c - \tau V)_+\Big)^{\frac{m}{m-1}} \Big\} \text{ with } \theta := \frac{1}{r\tau^{\frac{1}{m-1}}} \Big(\frac{m-1}{m}\Big)^{\frac{m}{m-1}} \tag{3.32}$$

whose solution is

$$c = \begin{cases} \overline{c}, \text{ if } \overline{c} \le \tau V, \\ \text{the root in } (\tau V, +\infty) \text{ of 3.33 otherwise} \end{cases}$$

where 3.33 is the equation

$$\overline{c} = c + \frac{\theta m}{m-1}(c - \tau V)^{\frac{1}{m-1}}. \tag{3.33}$$

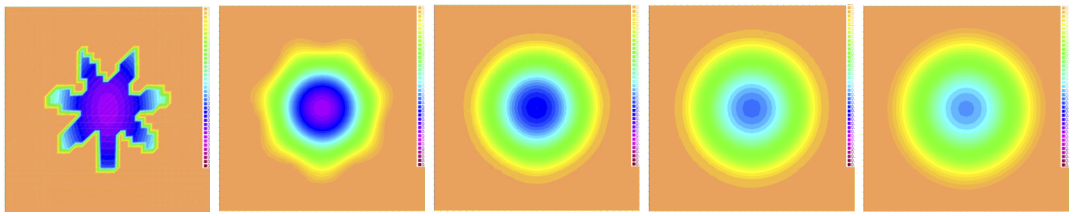
The case of a linear diffusion (Fokker-Planck) corresponds to an entropic internal energy $E(\rho) = \rho \log(\rho)$, in which case by similar computations, one finds $c = c(x)$ by solving

$$\overline{c} = c + \frac{1}{r} e^{\frac{c}{\tau} - V - 1}. \tag{3.34}$$

[1] http://www.freefem.org
[2] http://mumps.enseeiht.fr/

The fact that there is a dichotomy for the porous medium case (in contrast with the linear diffusion case leading to 3.34) corresponds to the finite speed of propagation and support containment in this case.

Figure 3.1 shows the evolution of the density for $m = 3$ and $V = \frac{|x|^2}{2}$. As expected (see [41]), we converge towards the stationary Barenblatt profile $BB(x) = (\frac{m-1}{2\,m}\max(1-|x|^2,0))^{\frac{1}{m-1}}$. Figure 3.2 gives quantitative information on the convergence towards the stationary profile, decrease of energy and conservation of mass (only external JKO time steps $k\tau$ are represented here).



$t = 0$ $t = 0.15$ $t = 0.6$ $t = 0.9$ $t = 1.5$

Figure 3.1: *Density at different time steps for the porous medium equation with a confining potential*

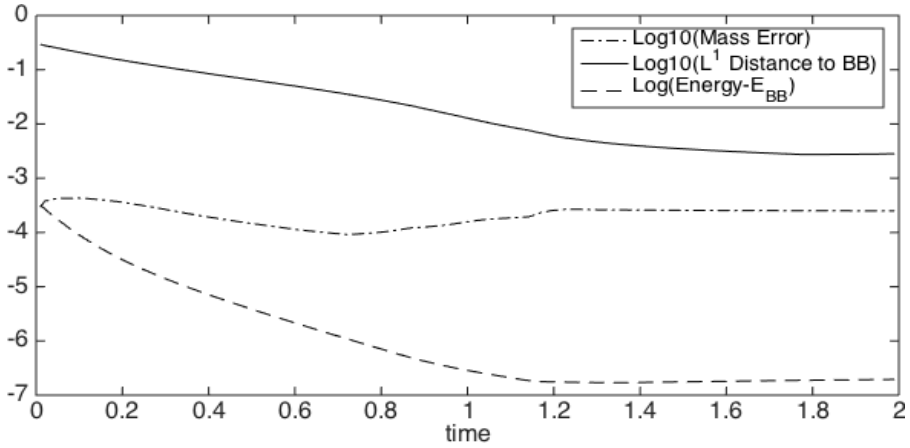


Figure 3.2: *This simulation shows the time evolution towards the stationary Barenblatt profile for $m = 3$. The first curve shows the mass of the density remains within the discretization error to the constant value (in log scale). The second curve is the L^1 distance of the density to the Barenblatt exact profile $BB(x) = (\frac{m-1}{2\,m}\max(1-|x|^2,0))^{\frac{1}{m-1}}$, reached up to discretization error again (in log scale).The last curve shows the decrease of the difference between the density energy (potential + entropy) $\mathcal{E}(\rho)$ and Barenblatt energy $\mathcal{E}(BB)$ in log scale. As expected (see [41], [4]) the decrease behaves like $-2\,t$ down to numerical discretization error.*

3.3.2 Application to crowd motions

In [94], Santambrogio and Mészáros considered a model of congested crowd motion with diffusion which leads to

$$\partial_t \rho - \Delta\rho = \mathrm{div}(\rho(\nabla V + \nabla p)),\ p \geq 0,\ \rho \leq 1,\ p(1-\rho) = 0, \tag{3.35}$$

with no flux boundary condition. In the nondiffusive case (no Laplacian in the left hand side), this model is due to Maury, Roudneff-Chupin and Santambrogio [89] who made it clear that it has a gradient flow structure. The diffusive case 3.35, of course also has a gradient flow structure for the following energy

$$\mathcal{E}(\rho) = \int_\Omega E(\rho(x))\mathrm{d}x + \int_\Omega V(x)\rho(x)\mathrm{d}x$$

with

$$E(\rho) := \rho\log(\rho) + \chi_{[0,1]}(\rho) = \begin{cases} \rho\log(\rho), \text{ if } \rho \in [0,1] \\ +\infty, \text{otherwise.} \end{cases}$$

A direct computation gives

$$\mathcal{E}^*(x,c) = \max\left(\frac{c}{\tau} - V(x), e^{\frac{c}{\tau}-V(x)-1}\right)$$

and then, again setting, $\overline{c} = -\phi^{n+1}(1,x) + \frac{\mu_1^n(x)}{r}$ and $V = V(x)$, 3.26 becomes

$$\min_c \left\{\frac{1}{2}|c-\overline{c}|^2 + \frac{\tau}{r}\max\left(\frac{c}{\tau} - V, e^{\frac{c}{\tau}-V-1}\right)\right\}$$

whose solution is explicit:

$$c = \begin{cases} \overline{c} - \frac{1}{r}, \text{ if } \overline{c} \geq \tau(1+V) + \frac{1}{r}, \\ \text{the root of 3.34, otherwise.} \end{cases}$$

The alternative in the previous formula somehow corresponds to the pressure p being on/off.

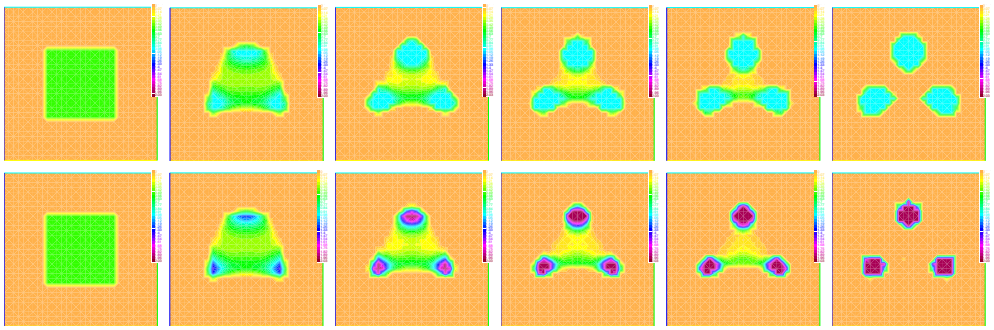


Figure 3.3: *Crowd motion evolution of one species with potential* $V((x_1,x_2)) = 30(((x_1^2)-1/4)^2 + (x_2+1/4)^2)((x_2-1/2)^2+x_1^2)$. *The first row represents the evolution under the constraint* $\rho \leqslant 1$ *and the second the evolution under the constraint* $\rho \leqslant 2$.

In figure 3.3, we represent the evolution of one species, the potential has three minima (hot spots where the crowd wants to go) but with two different density constraints. When the density threshold is higher (second row $\rho \leqslant 2$) then, at the end, the density is more concentrated around the three minima of the potential.

Chapter 4

Drift interactions: Potential case

This chapter presents existence and uniqueness results for a class of parabolic systems with non linear diffusion and nonlocal interaction. These systems can be viewed as regular perturbations of Wasserstein gradient flows. Here we extend results known in the periodic case ([35]) to the whole space or on a smooth bounded domain. Existence is obtained using a semi-implicit Jordan-Kinderlehrer-Otto scheme and uniqueness follows from a displacement convexity argument.

This work is based on *On some non linear evolution systems which are perturbations of Wasserstein gradient flows*, [73] and numerical simulations come from a joint work with J-D. Benamou and G. Carlier [17].

4.1 Introduction

In this chapter, we study existence and uniqueness of solutions for systems of the form

$$\begin{cases} \partial_t \rho_i - \operatorname{div}(\rho_i \nabla(V_i[\boldsymbol{\rho}])) - \alpha_i \operatorname{div}(\rho_i \nabla F_i'(\rho_i)) = 0 & \text{on } \mathbb{R}^+ \times \Omega, \\ \rho_i(0,\cdot) = \rho_{i,0} & \text{on } \Omega, \end{cases} \tag{4.1}$$

where $i \in [\![1, l]\!]$ ($l \in \mathbb{N}^*$), $\Omega = \mathbb{R}^n$ or is a bounded set of $\mathbb{R}^n$ and $\boldsymbol{\rho} := (\rho_1, \dots, \rho_l)$ is a collection of densities. Our motivation for this system comes from its appearance in modeling interacting species.

In the case of $\nabla(V_i[\boldsymbol{\rho}]) = 0$ or $V_i[\boldsymbol{\rho}]$ does not depend on $\boldsymbol{\rho}$, this system can be seen as a gradient flow in the product Wasserstein space i.e $\nabla F_i'(\rho_i)$ can be seen as the first variation of a functional $\mathcal{F}_i$ defined on measures. This theory started with the work of Jordan, Kinderlehrer and Otto in [65] where they discovered that the Fokker-Planck equation can be seen as the gradient flow of $\int_{\mathbb{R}^n} \rho \log \rho + \int_{\mathbb{R}^n} V\rho$. The method that they used to prove this result is often called JKO scheme. Now, it is well-known that the gradient flow method permits to prove the existence of solution under very weak assumptions on the initial condition for several evolution equations, such as the heat equation [65], the porous media equation [97], degenerate parabolic equations [1], Keller-Segel equation [18]. The general theory of gradient flow has been very much developed and is detailed in the book of Ambrosio, Gigli and Savaré, [4], which is the main reference in this domain.

However, this method is very restrictive if we want to treat the case of systems with several interaction potentials. Indeed, Di Francesco and Fagioli show in the first part of [54] that we have to take the same (or proportional) interaction potentials, of the form $V[\rho] = W * \rho$ for all densities. They prove an existence/uniqueness result of (4.1) using gradient flow theory in a product Wasserstein space without diffusion ($\alpha_i = 0$) and with $l = 2$, $V_1[\rho_1, \rho_2] := W_{1,1} * \rho_1 + W_{1,2} * \rho_2$ and $V_2[\rho_1, \rho_2] := W_{2,2} * \rho_2 + W_{2,1} * \rho_1$ where $W_{1,2}$ and $W_{2,1}$ are proportional. Nevertheless in the second part of [54], they introduce a new semi-implicit JKO scheme to treat the case where $W_{1,2}$ and $W_{2,1}$ are not proportional. In other words, they use the usual JKO scheme freezing the measure in $V_i[\boldsymbol{\rho}]$.

The purpose of this paper is to add a nonlinear diffusion in the system studied in [54]. Unfortunately, this term requires strong convergence to pass to the limit. This can be obtained using an extension of Aubin-Lions lemma proved by Rossi and Savaré in [107] and recalled in theorem 4.15. This theorem requires separately time-compactness and space-compactness to obtain a

strong convergence in $L^m((0,T)\times\Omega)$. The time-compactness follows from classical estimate on the Wasserstein distance in the JKO scheme. The difficulty is to prove the space-compactness. In the following, we show this result on the whole space $\mathbb{R}^n$ or on a smooth bounded domain. On the one hand in $\mathbb{R}^n$, we use the powerful flow interchange argument of Matthes, McCann and Savaré [86] and also used in the work of Di Francesco and Matthes [55]. On the other hand in a bounded domain, the flow interchange argument is very restrictive because it forces us to work in a convex domain and to impose some boundary condition on $V_i[\boldsymbol{\rho}]$. To avoid these assumptions, we establish a BV estimate to obtain compactness in space and then to find the strong convergence needed.

This chapter is composed of seven sections. In section 4.2, we state our main result, theorem 4.3. Sections 4.3, 4.4 and 4.5 are devoted to prove theorem 4.3. In section 4.3, we introduce a semi-implict JKO scheme, as in [54], and resulting standard estimates. Then, in section 4.4, we recall the flow interchange theory developed in [86] and we find a stronger estimate on the solution's gradient, which can be done by differentiating the energy along the heat flow. In section 4.5, we establish convergence results and we prove theorem 4.3. Section 4.6 deals with the case of a bounded domain. In the final section 4.7, we show uniqueness of (4.1) using a displacement convexity argument.

4.2 Main result

Let $l\in\mathbb{N}^*$ and for all $i\in[\![1,l]\!]$, we define $V_i:\mathcal{P}(\mathbb{R}^n)^l\to\mathcal{C}^2(\mathbb{R}^n)$ continuous such that:

- For all $\boldsymbol{\rho}=(\rho_1,\dots,\rho_l)\in\mathcal{P}(\mathbb{R}^n)^l$,
$$V_i[\boldsymbol{\rho}]\geqslant 0, \tag{4.2}$$
- There exists $C>0$ such that for all $\boldsymbol{\rho}\in\mathcal{P}(\mathbb{R}^n)^l$,
$$\|\nabla(V_i[\boldsymbol{\rho}])\|_{L^\infty(\mathbb{R}^n)}+\|D^2(V_i[\boldsymbol{\rho}])\|_{L^\infty(\mathbb{R}^n)}\leqslant C, \tag{4.3}$$
i.e $V_i[\boldsymbol{\rho}]$ and $\nabla(V_i[\boldsymbol{\rho}])$ are Lipschitz functions and the Lipschitz constants do not depend on the measure.
- There exists $C>0$ such that for all $\boldsymbol{\nu},\boldsymbol{\sigma}\in\mathcal{P}(\mathbb{R}^n)^l$,
$$\|\nabla(V_i[\boldsymbol{\nu}])-\nabla(V_i[\boldsymbol{\sigma}])\|_{L^\infty(\mathbb{R}^n)}\leqslant CW_2(\boldsymbol{\nu},\boldsymbol{\sigma}). \tag{4.4}$$

Remark 4.1. *The assumption* (4.2) *can be replaced by $V_i[\boldsymbol{\rho}]$ is bounded by below uniformly in $\boldsymbol{\rho}$.*

Let $m\geqslant 1$, we define the class of functions $\mathcal{H}_m$ by

$$\mathcal{H}_m:=\{x\mapsto x\log(x)\}\text{ if } m=1,$$

and, if $m>1$, $\mathcal{H}_m$ is the class of strictly convex superlinear functions $F:\mathbb{R}^+\to\mathbb{R}$ which satisfy

$$F(0)=F'(0)=0,\qquad F''(x)\geqslant Cx^{m-2}\text{ and }P(x):=xF'(x)-F(x)\leqslant C(x+x^m). \tag{4.5}$$

The two first assumptions imply that if $m>1$ and $F\in\mathcal{H}_m$ then F controls x^m.

Before giving a definition of solution of (4.1), we recall that the nonlinear diffusion term can be rewrite as

$$\operatorname{div}(\rho\nabla F'(\rho))=\Delta P(\rho),$$

where $P(x):=xF'(x)-F(x)$ is the pressure associated to F.

Definition 4.2. *We say that* $(\rho_1, \dots, \rho_l) : [0, +\infty[\to \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$ *is a weak solution of* (4.1) *if for all* $i \in [\![1, l]\!]$, $\rho_i \in \mathcal{C}([0, T], \mathcal{P}_2^{ac}(\mathbb{R}^n))$, $P_i(\rho_i) \in L^1(]0, T[\times \mathbb{R}^n)$ *for all* $T < \infty$ *and for all* $\varphi_1, \dots, \varphi_l \in \mathcal{C}_c^\infty([0, +\infty[\times \mathbb{R}^n)$,

$$\int_0^{+\infty} \int_{\mathbb{R}^n} \left[\left(\partial_t \varphi_i - \nabla \varphi_i \cdot \nabla (V_i[\boldsymbol{\rho}]) \right) \rho_i + \alpha_i \Delta \varphi_i P_i(\rho_i) \right] = - \int_{\mathbb{R}^n} \varphi_i(0, x) \rho_{i,0}(x).$$

With this definition of solution we have the following result

Theorem 4.3. *For all* $i \in [\![1, l]\!]$, *let* $F_i \in \mathcal{H}_{m_i}$, *with* $m_i \geqslant 1$, *and* V_i *satisfy* (4.2), (4.3) *and* (4.4). *Let* $\alpha_1, \dots, \alpha_l$ *positive constants. If* $\rho_{i,0} \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$ *satisfy*

$$\mathcal{F}_i(\rho_{i,0}) + \mathcal{V}_i(\rho_{i,0} | \boldsymbol{\rho_0}) < +\infty, \tag{4.6}$$

with

$$\mathcal{F}_i(\rho) := \begin{cases} \int_{\mathbb{R}^n} F_i(\rho(x))\, dx & \text{if } \rho \ll \mathcal{L}^n, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and } \mathcal{V}_i(\rho | \boldsymbol{\mu}) := \int_{\mathbb{R}} V_i[\boldsymbol{\mu}] \rho \, dx.$$

then there exist $(\rho_1, \dots, \rho_l) : [0, +\infty[\to \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$, *continuous with respect to* W_2, *weak solution of* (4.1).

Remark 4.4. *In the following, to simplify the proof, we take* $\alpha_i = 1$.

4.3 Semi-implicit JKO scheme

In this section, we introduce the semi-implicit JKO scheme, as [54], and we find the first estimates as in the usual JKO scheme.

Let $h > 0$ be a time step, we construct l sequences with the following iterative discrete scheme: for all $i \in [\![1, l]\!]$, $\rho_{i,h}^0 = \rho_{i,0}$ and for all $k \geqslant 1$, $\rho_{i,h}^k$ minimizes

$$\mathcal{E}_{i,h}(\rho | \boldsymbol{\rho}_h^{k-1}) := W_2^2(\rho, \rho_{i,h}^{k-1}) + 2h \left(\mathcal{F}_i(\rho) + \mathcal{V}_i(\rho | \boldsymbol{\rho}_h^{k-1}) \right),$$

on $\rho \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$, with $\boldsymbol{\rho}_h^{k-1} = (\rho_{1,h}^{k-1}, \dots, \rho_{l,h}^{k-1})$.

In the next proposition, we show that all these sequences are well defined. We start to prove that there are well defined for one step and after in remark 4.6, we extend the result for all k.

Proposition 4.5. *Let* $\boldsymbol{\rho}_0 = (\rho_{1,0}, \dots, \rho_{l,0}) \in \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$, *there exists a unique* $\boldsymbol{\rho}_h^1 = (\rho_{1,h}^1, \dots, \rho_{l,h}^1) \in \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$ *such that, for all* $i \in [\![1, l]\!]$,

$$\rho_{i,h}^1 = \operatorname{argmin} \mathcal{E}_{i,h}(\rho | \boldsymbol{\rho}_h^0). \tag{4.7}$$

Proof. First of all, we distinguish the case $m_i > 1$ from $m_i = 1$.

- If $m_i > 1$, then $\mathcal{E}_{i,h}(\rho | \boldsymbol{\mu}) \geqslant 0$, for all $\rho, \mu_1, \dots, \mu_l \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$. Let ρ_ν be a minimizing sequence. As $\mathcal{E}_{i,h}(\rho_{i,0} | \boldsymbol{\rho}_0) < +\infty$ (according to (4.6)), $(\mathcal{E}_{i,h}(\rho_\nu | \boldsymbol{\rho}_0))_\nu$ is bounded above. So there exists $C > 0$ such that

  $$0 \leqslant \mathcal{F}_i(\rho_\nu) \leqslant C \text{ and } W_2(\rho_\nu, \rho_{i,0}) \leqslant C.$$

  From the second inequality, it follows that the second moment of ρ_ν is bounded.

- Now if $m_i = 1$, following [65], we obtain

  $$\mathcal{E}_{i,h}(\rho | \boldsymbol{\rho}_h^0) \geqslant \frac{1}{4} M(\rho) - C(1 + M(\rho))^\alpha - \frac{1}{2} M(\rho_{i,h}^0), \tag{4.8}$$

  with some $0 < \alpha < 1$. And since $x \mapsto \frac{1}{4} x - C(1 + x)^\alpha$ is bounded below, we see that $\mathcal{E}_{i,h}$ is bounded below.

Let ρ_ν be a minimizing sequence. Then we have $(\mathcal{F}_i(\rho_\nu))_\nu$ bounded above. Indeed, as $\mathcal{E}_{i,h}(\rho_{i,0}|\boldsymbol{\rho}_0) < +\infty$, $(\mathcal{E}_{i,h}(\rho_\nu|\boldsymbol{\rho}_0))_\nu$ is bounded above and from (4.2) we get,

$$\int_{\mathbb{R}} V_i[\boldsymbol{\rho}_0](x)\rho_\nu(x)\,dx \geqslant 0,$$

so $(\mathcal{F}_i(\rho_\nu))_\nu$ is bounded above. According to (4.8), $(M(\rho_\nu))_\nu$ is bounded. Consequently $(\mathcal{F}_i(\rho_\nu))_\nu$ is bounded because $\mathcal{F}_i(\rho) \geqslant -C(1+M(\rho))^\alpha$.

In both cases, using Dunford-Pettis' theorem, we deduce that there exists $\rho^1_{i,h} \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$ such that

$$\rho_\nu \rightharpoonup \rho^1_{i,h} \text{ weakly in } L^1(\mathbb{R}^n).$$

It remains to prove that $\rho^1_{i,h}$ is a solution for the minimization problem. But since $\mathcal{F}_i$ and $W_2^2(\cdot,\rho_{i,0})$ are weakly lower semi-continuous in $L^1(\mathbb{R}^n)$, we have

$$\mathcal{E}_{i,h}(\rho^1_{i,h}|\boldsymbol{\rho}_0) \leqslant \liminf_{\nu\nearrow+\infty} \mathcal{E}_{i,h}(\rho_\nu|\boldsymbol{\rho}_0).$$

To conclude the proof, we show that the minimizer is unique. This follows from the convexity of $\mathcal{V}_i(\cdot|\boldsymbol{\rho}_0)$ and $\rho \in \mathcal{P}_2^{ac}(\mathbb{R}^n) \mapsto W_2^2(\rho,\rho^0_{i,h})$ and the strict convexity of $\mathcal{F}_i$.

□

Remark 4.6. *By induction, proposition 4.5 is still true for all $k \geqslant 1$:*
the proof is similar when we take $k-1$ instead of 0 and if we notice that for all i,

$$\mathcal{F}_i(\rho^1_{i,h}) + \mathcal{V}_i(\rho^1_{i,h}|\boldsymbol{\rho}^1_h) \leqslant \mathcal{F}_i(\rho_{i,0}) + \mathcal{V}_i(\rho_{i,0}|\boldsymbol{\rho}_0) + CW_2(\boldsymbol{\rho}_0,\boldsymbol{\rho}^1_h) \leqslant C.$$

The last inequality is obtained from the minimization scheme and from the assumptions (4.2), (4.4) *and* (4.6). *By induction it becomes, for all $k \geqslant 2$,*

$$\mathcal{F}_i(\rho^{k-1}_{i,h}) + \mathcal{V}_i(\rho^{k-1}_{i,h}|\boldsymbol{\rho}^{k-1}_h) \leqslant \mathcal{F}_i(\rho_{i,0}) + \mathcal{V}_i(\rho_{i,0}|\boldsymbol{\rho}_0) + C\sum_{j=1}^{k-1} W_2(\boldsymbol{\rho}^{j-1}_h,\boldsymbol{\rho}^j_h) \leqslant C.$$

This inequality shows $\mathcal{E}_{i,h}(\rho^{k-1}_{i,h}|\boldsymbol{\rho}^{k-1}_h) < +\infty$ and so we can bound $(\mathcal{F}_i(\rho_\nu))_\nu$ in the previous proof.

Thus we proved that sequences $(\rho^k_{i,h})_{k\geqslant 0}$ are well defined for all $i \in [\![1,l]\!]$. Then we define the interpolation $\rho_{i,h} : \mathbb{R}^+ \to \mathcal{P}_2^{ac}(\mathbb{R}^n)$ by, for all $k \in \mathbb{N}$,

$$\rho_{i,h}(t) = \rho^k_{i,h} \text{ if } t \in ((k-1)h, kh]. \tag{4.9}$$

The following proposition shows that $\rho_{i,h}$ are solutions of a discrete approximation of the system (4.1).

Proposition 4.7. *Let $h > 0$, for all $T > 0$, let N such that $N = \lceil \frac{T}{h} \rceil$. Then for all $(\phi_1,\dots,\phi_l) \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)^l$ and for all $i \in [\![1,l]\!]$,*

$$\begin{aligned}
\int_0^T \int_{\mathbb{R}^n} \rho_{i,h}(t,x)\partial_t\phi_i(t,x)\,dxdt &= -h\sum_{k=0}^{N-1}\int_{\mathbb{R}^n} P_i(\rho^{k+1}_{i,h}(x))\Delta\phi_i(t_k,x)\,dx \\
&+ h\sum_{k=0}^{N-1}\int_{\mathbb{R}^n} \nabla(V_i[\boldsymbol{\rho}^k_h])\cdot\nabla\phi_i(t_k,x)\rho^{k+1}_{i,h}(x)\,dx \\
&+ \sum_{k=0}^{N-1}\int_{\mathbb{R}^n\times\mathbb{R}^n} \mathcal{R}[\phi_i(t_k,\cdot)](x,y)d\gamma^k_{i,h}(x,y) \\
&- \int_{\mathbb{R}^n} \rho_{i,0}(x)\phi_i(0,x)\,dx,
\end{aligned}$$

where $t_k = hk$ ($t_N := T$) and $\gamma_{i,h}^k$ is the optimal transport plan in $W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1})$. Moreover, $\mathcal{R}$ is defined such that, for all $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\mathbb{R}^n)}|x-y|^2.$$

Proof. We split the proof in two steps. We first compute the first variation of $\mathcal{E}_{i,h}(\cdot|\boldsymbol{\rho}_h^k)$ and then we integrate in time. In the following, i is fixed in $[\![1,l]\!]$.

- *First step:* For all $k \geqslant 0$, if $\gamma_{i,h}^k$ is the optimal transport plan in $W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1})$ then

$$\begin{aligned}\int_{\mathbb{R}^n}\varphi_i(x)(\rho_{i,h}^{k+1}(x)-\rho_{i,h}^k(x)) &= h\int_{\mathbb{R}^n} P_i(\rho_{i,h}^{k+1}(x))\Delta\varphi_i(x)\,dx\\ &- h\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}_h^k])(x)\cdot\nabla\varphi_i(x)\rho_{i,h}^{k+1}(x)\,dx\\ &- \int_{\mathbb{R}^n\times\mathbb{R}^n}\mathcal{R}[\varphi_i](x,y)d\gamma_{i,h}^k(x,y),\end{aligned}$$

  for all $\varphi_i \in \mathcal{C}_c^\infty(\mathbb{R}^n)$.

  To obtain this equality, we compute the first variation of $\mathcal{E}_{i,h}(\cdot|\boldsymbol{\rho}_h^k)$. Let $\xi_i \in \mathcal{C}_c^\infty(\mathbb{R}^n,\mathbb{R}^n)$ and $\tau > 0$ and let Ψ_τ defined by

$$\partial_\tau\Psi_\tau = \xi_i\circ\Psi_\tau, \qquad \Psi_0 = Id.$$

  After we perturb $\rho_{i,h}^{k+1}$ by $\rho_\tau = (\Psi_\tau)_\sharp\rho_{i,h}^{k+1}$. According to the definition of $\rho_{i,h}^{k+1}$, we get

$$\frac{1}{\tau}\left(\mathcal{E}_{i,h}(\rho_\tau|\boldsymbol{\rho}_h^k) - \mathcal{E}_{i,h}(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right) \geqslant 0. \tag{4.10}$$

  By standard computations (see for instance [65], [1]) we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\frac{1}{2}W_2^2(\rho_\tau,\rho_{i,h}^k) - \frac{1}{2}W_2^2(\rho_{i,h}^{k+1},\rho_{i,h}^k)) \leqslant \int_{\mathbb{R}^n\times\mathbb{R}^n}(y-x)\cdot\xi_i(y)\,d\gamma_{i,h}^k(x,y), \tag{4.11}$$

  where $\gamma_{i,h}^k$ is the optimal transport plan in $W_2(\rho_{i,h}^k,\rho_{i,h}^{k+1})$,

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{F}_i(\rho_\tau) - \mathcal{F}_i(\rho_{i,h}^{k+1})) \leqslant -\int_{\mathbb{R}^n} P_i(\rho_{i,h}^{k+1}(x))\,\mathrm{div}(\xi_i(x))\,dx, \tag{4.12}$$

  and

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}\left(\mathcal{V}_i(\rho_\tau|\boldsymbol{\rho}_h^k) - \mathcal{V}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right) \leqslant \int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}_h^k])(x)\cdot\xi_i(x)\rho_{i,h}^{k+1}(x)\,dx. \tag{4.13}$$

  If we combine (4.10), (4.11), (4.12) and (4.13), we get

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n}(y-x)\cdot\xi_i(y)\,d\gamma_{i,h}^k(x,y) + h\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}_h^k])(x)\cdot\xi_i(x)\rho_{i,h}^{k+1}(x)\,dx&\\ - h\int_{\mathbb{R}^n} P_i(\rho_{i,h}^{k+1}(x))\,\mathrm{div}(\xi_i(x))\,dx \geqslant 0.&\end{aligned}$$

  And if we replace ξ_i by $-\xi_i$, this inequality becomes an equality.
  To conclude this first part, we choose $\xi_i = \nabla\varphi_i$ and we notice, using Taylor's expansion, that

$$\varphi_i(x) - \varphi_i(y) = \nabla\varphi_i(y)\cdot(x-y) + \mathcal{R}[\varphi_i](x,y),$$

  with $\mathcal{R}[\varphi_i]$ satisfying

$$|\mathcal{R}[\varphi_i](x,y)| \leqslant \frac{1}{2}\|D^2\varphi_i\|_{L^\infty([0,T)\times\mathbb{R}^n)}|x-y|^2.$$

- *Second step:* For all $(\phi_1, \dots, \phi_l) \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^n)^l$, extended, for all $i \in [\![1,l]\!]$, by $\phi_i(0,\cdot)$ on $[-h,0)$, then

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^n} \rho_{i,h}(t,x)\partial_t\phi_i(t,x)\,dxdt &= \sum_{k=0}^{N} \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^n} \rho_{i,h}^k(x)\partial_t\phi_i(t,x)\,dxdt \\ &= \sum_{k=0}^{N} \int_{\mathbb{R}^n} \rho_{i,h}^k(x)(\phi_i(t_k,x) - \phi_i(t_{k-1},x))\,dx \\ &= \sum_{k=0}^{N-1} \int_{\mathbb{R}^n} \phi_i(t_k,x)(\rho_{i,h}^k(x) - \rho_{i,h}^{k+1}(x))\,dx \\ &\quad - \int_{\mathbb{R}^n} \rho_{i,0}(x)\phi_i(0,x)\,dx.\end{aligned}$$

  Using the first part with $\varphi_i = \phi_i(t_k,\cdot)$, we get the desired equality.

□

The last proposition of this section gives usual estimates in gradient flow theory.

Proposition 4.8. *For all $T < +\infty$ and for all $i \in [\![1,l]\!]$, there exists a constant $C < +\infty$ such that for all $k \in \mathbb{N}$ and for all h with $kh \leqslant T$ and let $N = \lceil \frac{T}{h} \rceil$, we have*

$$M(\rho_{i,h}^k) \leqslant C, \tag{4.14}$$

$$\mathcal{F}_i(\rho_{i,h}^k) \leqslant C, \tag{4.15}$$

$$\textstyle\sum_{k=0}^{N-1} W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) \leqslant Ch. \tag{4.16}$$

Proof. The proof combines some techniques used in [65] et [54]. In the following, i is fixed in $[\![1,l]\!]$. As $\rho_{i,h}^{k+1}$ is optimal and $\rho_{i,h}^k$ is admissible, we have

$$\mathcal{E}_{i,h}(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k) \leqslant \mathcal{E}_{i,h}(\rho_{i,h}^k|\boldsymbol{\rho}_h^k).$$

In other words,

$$\frac{1}{2}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) + h\left(\mathcal{F}_i(\rho_{i,h}^{k+1}) + \mathcal{V}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right) \leqslant h\left(\mathcal{F}_i(\rho_{i,h}^k) + \mathcal{V}_i(\rho_{i,h}^k|\boldsymbol{\rho}_h^k)\right).$$

From (4.3), we know that $V_i[\boldsymbol{\rho}]$ is a C-Lipschitz function where C does not depend on the measure. Hence, because of (**??**), we have

$$\mathcal{V}_i(\rho_{i,h}^k|\boldsymbol{\rho}_h^k) - \mathcal{V}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k) \leqslant CW_2(\rho_{i,h}^{k+1}, \rho_{i,h}^k).$$

Using Young's inequality, we obtain

$$\mathcal{V}_i(\rho_{i,h}^k|\boldsymbol{\rho}_h^k) - \mathcal{V}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k) \leqslant C^2h + \frac{1}{4h}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k).$$

It yields

$$\frac{1}{4}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) \leqslant h(\mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\rho_{i,h}^{k+1})) + C^2h^2. \tag{4.17}$$

Summing over k, we can assert that

$$\begin{aligned}\sum_{k=0}^{N-1} \frac{1}{4}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) &\leqslant h\left(\sum_{k=0}^{N-1}\left(\mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\rho_{i,h}^{k+1})\right) + C^2T\right) \\ &\leqslant h\left(\mathcal{F}_i(\rho_{i,0}) - \mathcal{F}_i(\rho_{i,h}^N) + C^2T\right).\end{aligned}$$

But by assumption, $\mathcal{F}_i(\rho_{i,0}) < +\infty$ and $-\mathcal{F}_i(\rho) \leqslant C(1+M(\rho))^\alpha$, with $0<\alpha<1$, then

$$\sum_{k=1}^{N} \frac{1}{4} W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) \leqslant h\left(\mathcal{F}_i(\rho_{i,0}) + C(1+M(\rho_{i,h}^N))^\alpha + C^2T\right). \tag{4.18}$$

Thus we are reduced to prove (4.14). But

$$\begin{aligned} M(\rho_{i,h}^k) &\leqslant 2W_2^2(\rho_{i,h}^k, \rho_{i,0}) + 2M(\rho_{i,0}) \\ &\leqslant 2k \sum_{m=0}^{k-1} W_2^2(\rho_{i,h}^m, \rho_{i,h}^{m+1}) + 2M(\rho_{i,0}) \\ &\leqslant 8kh\left(\mathcal{F}_i(\rho_{i,0}) + C(1+M(\rho_{i,h}^k))^\alpha + C^2T\right) + 2M(\rho_{i,0}) \\ &\leqslant 8T\left(\mathcal{F}_i(\rho_{i,0}) + C(1+M(\rho_{i,h}^k))^\alpha + C^2T\right) + 2M(\rho_{i,0}). \end{aligned}$$

As $\alpha<1$, we get (4.14). The second line is obtained with the triangle inequality and Cauchy-Schwarz inequality while the third line is obtained because of (4.18). So we have poved (4.14) and (4.16).

To have (4.15), we just have to use (4.17) and to sum. This implies

$$\mathcal{F}_i(\rho_{i,h}^k) \leqslant \mathcal{F}_i(\rho_{i,0}) + C^2T,$$

which proves the proposition.

□

4.4 κ-flows and gradient estimate

Estimates of proposition 4.8 permit to obtain weak convergence in L^1 (see proposition 4.13). Unfortunately, it is not enough to pass to the limit in the nonlinear diffusion term $P_i(\rho_{i,h})$. In this section, we follow the general strategy developed in [86] and used in [55] and [35] to get an estimate on the gradient of $\rho_{i,h}^{m_i/2}$. This estimate will be used in proposition 4.14 to have a strong convergence of $\rho_{i,h}$ in $L^{m_i}(]0,T[\times\mathbb{R}^n)$. In the following, we are only interested by the case where $m_i>1$ because if $m_i=1$, $P_i(\rho_{i,h})=\rho_{i,h}$ and the weak convergence is enough to pass to the limit in proposition 4.7. In the first part of this section, we recall the definition of κ-flows (or contractive gradient flow) and some results on the dissipation of $\mathcal{F}_i+\mathcal{V}_i$ then, in the second part, we use these results with the heat flow to find an estimate on the gradient.

4.4.1 κ-flows

Definition 4.9. *A semigroup $\mathfrak{S}_\Psi : \mathbb{R}^+\times\mathcal{P}_2^{ac}(\mathbb{R}^n) \to \mathcal{P}_2^{ac}(\mathbb{R}^n)$ is a κ-flow for the functional $\Psi : \mathcal{P}_2^{ac}(\mathbb{R}^n) \to \mathbb{R}\cup\{+\infty\}$ with respect to W_2 if, for all $\rho\in\mathcal{P}_2^{ac}(\mathbb{R}^n)$, the curve $s\mapsto\mathfrak{S}_\Psi^s[\rho]$ is absolutely continuous on $\mathbb{R}^+$, $\mathfrak{S}_\Psi^0 = Id$ and satisfies the evolution variational inequality (EVI)*

$$\frac{1}{2}\frac{d^+}{d\sigma}\,|_{\sigma=s}\, W_2^2(\mathfrak{S}_\Psi^s[\rho],\tilde{\rho}) + \frac{\kappa}{2}W_2^2(\mathfrak{S}_\Psi^s[\rho],\tilde{\rho}) \leqslant \Psi(\tilde{\rho}) - \Psi(\mathfrak{S}_\Psi^s[\rho]), \tag{4.19}$$

for all $s>0$ and for all $\tilde{\rho}\in\mathcal{P}_2^{ac}(\mathbb{R}^n)$ such that $\Psi(\tilde{\rho})<+\infty$, where

$$\frac{d^+}{dt}f(t) := \limsup_{s\to 0^+}\frac{f(t+s)-f(t)}{s}.$$

In [4], the authors showed that the fact a functional admits a κ-flow is equivalent to λ-displacement convexity (see section 4.7 for definition).

The next two lemmas give results on the variations of $\rho_{i,h}^k$ along specific κ-flows and are extracted from [55]. The goal is to use them with the heat flow.

Lemma 4.10. *Let $\Psi : \mathcal{P}_2^{ac}(\mathbb{R}^n) \to \mathbb{R}\cup\{+\infty\}$ l.s.c on $\mathcal{P}_2^{ac}(\mathbb{R}^n)$ which possesses a κ-flot $\mathfrak{S}_\Psi$. Define the dissipation $\mathcal{D}_{i,\Psi}$ along $\mathfrak{S}_\Psi$ by*

$$\mathcal{D}_{i,\Psi}(\rho|\boldsymbol{\mu}) := \limsup_{s\searrow 0} \frac{1}{s}\left(\mathcal{F}_i(\rho) - \mathcal{F}_i(\mathfrak{S}^s_\Psi[\rho]) + \mathcal{V}_i(\rho|\boldsymbol{\mu}) - \mathcal{V}_i(\mathfrak{S}^s_\Psi[\rho]|\boldsymbol{\mu})\right)$$

for all $\rho \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$ and $\boldsymbol{\mu} \in \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$.
If $\rho_{i,h}^{k-1}$ et $\rho_{i,h}^k$ are two consecutive steps of the semi-implicit JKO scheme, then

$$\Psi(\rho_{i,h}^{k-1}) - \Psi(\rho_{i,h}^k) \geqslant h\mathcal{D}_{i,\Psi}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}) + \frac{\kappa}{2}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k-1}). \tag{4.20}$$

Proof. Since the result is trivial if $\Psi(\rho_{i,h}^{k-1}) = +\infty$, we assume $\Psi(\rho_{i,h}^{k-1}) < +\infty$.
Thus we can use the EVI inequality (4.19) with $\rho := \rho_{i,h}^k$ and $\tilde{\rho} := \rho_{i,h}^{k-1}$. We obtain

$$\Psi(\rho_{i,h}^{k-1}) - \Psi(\mathfrak{S}^s_\Psi[\rho_{i,h}^k]) \geqslant \frac{1}{2}\frac{d^+}{d\sigma}\,|_{\sigma=s}\, W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1}) + \frac{\kappa}{2}W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1}).$$

By lower semi-continuity of Ψ, we have

$$\begin{aligned}
\Psi(\rho_{i,h}^{k-1}) - \Psi(\rho_{i,h}^k) &\geqslant \Psi(\rho_{i,h}^{k-1}) - \liminf_{s\searrow 0}\Psi(\mathfrak{S}^s_\Psi[\rho_{i,h}^k]) \\
&\geqslant \limsup_{s\searrow 0}\left(\Psi(\rho_{i,h}^{k-1}) - \Psi(\mathfrak{S}^s_\Psi(\rho_{i,h}^k))\right) \\
&\geqslant \limsup_{s\searrow 0}\left(\frac{1}{2}\frac{d^+}{d\sigma}\,|_{\sigma=s}\, W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1})\right) \\
&+ \frac{\kappa}{2}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k-1}).
\end{aligned}$$

The last line is obtained thanks to the W_2-continuity of $s \mapsto \mathfrak{S}^s_\Psi[\rho_{i,h}^k]$ in $s = 0$. Moreover, the absolute continuity of $s \mapsto \mathfrak{S}^s_\Psi[\rho_{i,h}^k]$ implies

$$\begin{aligned}
&\limsup_{s\searrow 0}\left(\frac{1}{2}\frac{d^+}{d\sigma}\,|_{\sigma=s}\, W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1})\right) \\
&\qquad\qquad \geqslant \limsup_{s\searrow 0}\frac{1}{2s}\left(W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1}) - W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k-1})\right).
\end{aligned}$$

But since $\rho_{i,h}^k$ minimizes $\mathcal{E}_{i,h}(\cdot|\boldsymbol{\rho}_h^{k-1})$, we get, for all $s \geqslant 0$,

$$\begin{aligned}
W_2^2(\mathfrak{S}^s_\Psi[\rho_{i,h}^k], \rho_{i,h}^{k-1}) - W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k-1}) &\geqslant 2h\left(\mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\mathfrak{S}^s_\Psi[\rho_{i,h}^k])\right) \\
&+ 2h\left(\mathcal{V}_i(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}) - \mathcal{V}_i(\mathfrak{S}^s_\Psi[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right).
\end{aligned}$$

This concludes the proof.

□

Corollary 4.11. *Under the same hypotheses as in lemma 4.10, let $\mathfrak{S}_\Psi$ a κ-flow such that, for all $k \in \mathbb{N}$, the curve $s \mapsto \mathfrak{S}^s_\Psi[\rho_{i,h}^k]$ lies in $L^1(\mathbb{R}^n)$. Moreover, assume that $s \mapsto F_i(\mathfrak{S}^s_\Psi[\rho_{i,h}^k])$ is differentiable for $s > 0$ and is continuous at $s = 0$ in $L^1(\mathbb{R}^n)$.*
In addition, we assume that the family $-\frac{d}{d\sigma}_{|_{\sigma=s}}\left(\mathcal{F}_i(\mathfrak{S}^\sigma_\Psi[\rho_{i,h}^k]) + \mathcal{V}_i(\mathfrak{S}^\sigma_\Psi[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right)$ is bounded from below by an integrable function as s goes to 0 and let $\mathfrak{K}_{i,\Psi} : \mathcal{P}_2^{ac}(\mathbb{R}^n) \mapsto]-\infty, +\infty]$ be a functional such that

$$\liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\left(\mathcal{F}_i(\mathfrak{S}^\sigma_\Psi[\rho_{i,h}^k]) + \mathcal{V}_i(\mathfrak{S}^\sigma_\Psi[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right)\right) \geqslant \mathfrak{K}_{i,\Psi}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}). \tag{4.21}$$

Then, for all $k \in \mathbb{N}$,

$$\Psi(\rho_{i,h}^{k-1}) - \Psi(\rho_{i,h}^k) \geqslant h\mathfrak{K}_{i,\Psi}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}) + \frac{\kappa}{2}W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k-1}). \tag{4.22}$$

Proof. It is sufficient to show that $\mathcal{D}_{i,\Psi}(\cdot|\boldsymbol{\rho}_h^{k-1})$ is bounded below by $\mathfrak{K}_{i,\Psi}(\cdot|\boldsymbol{\rho}_h^{k-1})$. The proof is as in corollary 4.3 of [55]. The hypotheses of L^1-regularity on $F_i(\mathfrak{S}_\Psi^s[\rho_{i,h}^k])$ imply that $s \mapsto \mathcal{F}_i(\mathfrak{S}_\Psi^s[\rho_{i,h}^k])$ is differentiable for $s > 0$ and continuous at $s = 0$. We have the same regularity for $s \mapsto \mathcal{V}_i(\mathfrak{S}_\Psi^s[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})$. By the fundamental theorem of calculus,

$$\begin{aligned}
\mathcal{D}_{i,\Psi}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}) &= \limsup_{s\searrow 0} \frac{1}{s}\left(\mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\mathfrak{S}_\psi^s[\rho_{i,h}^k])\right. \\
&\qquad \left. + \mathcal{V}_i(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}) - \mathcal{V}_i(\mathfrak{S}_\Psi^s[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right) \\
&= \limsup_{s\searrow 0} \int_0^1 \left(-\frac{d}{d\sigma}_{|_{\sigma=sz}} \left(\mathcal{F}_i(\mathfrak{S}_\Psi^\sigma[\rho_{i,h}^k]) + \mathcal{V}_i(\mathfrak{S}_\Psi^\sigma[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right)\right) dz \\
&\geqslant \int_0^1 \liminf_{s\searrow 0} \left(-\frac{d}{d\sigma}_{|_{\sigma=sz}} \left(\mathcal{F}_i(\mathfrak{S}_\Psi^\sigma[\rho_{i,h}^k]) + \mathcal{V}_i(\mathfrak{S}_\Psi^\sigma[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right)\right) dz \\
&\geqslant \mathfrak{K}_{i,\Psi}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}).
\end{aligned}$$

The last line is obtained by Fatou's lemma and assumption (4.21). To conclude we apply lemma 4.10. □

4.4.2 Gradient estimate

Proposition 4.12. *For all $i \in [\![1, l]\!]$ such that $m_i > 1$, there exists a constant C which depends only on $\rho_{i,0}$ such that*

$$\|\rho_{i,h}^{m_i/2}\|_{L^2([0,T];H^1(\mathbb{R}^n))} \leqslant C(1+T)$$

for all $T > 0$.

Before starting the proof of the proposition 4.12, we recall the definition of the Entropy functional ,

$$E(\rho) = \int_{\mathbb{R}^n} \rho \log \rho, \text{ for all } \rho \in \mathcal{P}^{ac}(\mathbb{R}^n).$$

We know that this functional possesses a κ-flow, with $\kappa = 0$ which is given by the heat semigroup (see for instance [49], [65] or [116]). In other words, for a given $\eta_0 \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$, the curve $s \mapsto \eta(s) := \mathfrak{S}_E^s[\eta_0]$ solves

$$\begin{cases} \partial_s \eta = \Delta \eta & \text{on } \mathbb{R}^+ \times \mathbb{R}^n, \\ \eta(0) = \eta_0 & \text{on } \mathbb{R}^n, \end{cases}$$

in the classical sense. $\eta(s)$ is a positive density for all $s > 0$ and is continuously differentiable as a map from $\mathbb{R}^+$ to $\mathcal{C}^\infty \cap L^1(\mathbb{R}^n)$. Moreover, if $\eta_0 \in L^m(\mathbb{R}^n)$, then $\eta(s)$ converges to η_0 in $L^m(\mathbb{R}^n)$ when $s \searrow 0$.

Proof. Based on the facts set out above, $\mathfrak{S}_E$ satisfies the hypotheses of the corollary 4.11. We just have to define a suitable lower bound $\mathfrak{K}_{i,E}$ to use it. The spatial regularity of $\eta(s)$ for all $s > 0$ allows the following calculations. Thus for all $\boldsymbol{\mu} \in \mathcal{P}_2^{ac}(\mathbb{R}^n)^l$, we have

$$\begin{aligned}
\partial_s \left(\mathcal{F}_i(\mathfrak{S}_E^s[\eta_0]) + \mathcal{V}_i(\mathfrak{S}_E^s[\eta_0]|\boldsymbol{\mu})\right) &= \int_{\mathbb{R}^n} \partial_s F_i(\eta)\, dx + \int_{\mathbb{R}^n} V_i[\boldsymbol{\mu}] \partial_s \eta(s,x)\, dx \\
&= \int_{\mathbb{R}^n} F_i'(\eta(s,x)) \Delta \eta(s,x)\, dx + \int_{\mathbb{R}^n} V_i[\boldsymbol{\mu}] \Delta \eta(s,x)\, dx \\
&= -\int_{\mathbb{R}^n} F_i''(\eta(s,x)) |\nabla \eta(s,x)|^2\, dx + \int_{\mathbb{R}^n} \Delta(V_i[\boldsymbol{\mu}]) \eta(s,x)\, dx.
\end{aligned}$$

According to (4.5), $F_i''(x) \geqslant Cx^{m_i-2}$ thus

$$\begin{aligned}\partial_s\left(\mathcal{F}_i(\mathfrak{S}_E^s[\eta_0]) \right. &+ \left. \mathcal{V}_i(\mathfrak{S}_E^s[\eta_0]|\boldsymbol{\mu})\right)\\ &\leqslant -C\int_{\mathbb{R}^n}\eta(s,x)^{m_i-2}|\nabla\eta(s,x)|^2\,dx + \int_{\mathbb{R}^n}\Delta(V_i[\boldsymbol{\mu}])\eta(s,x)\,dx\\ &\leqslant -C\int_{\mathbb{R}^n}|\nabla\eta(s,x)^{m_i/2}|^2\,dx + \int_{\mathbb{R}^n}\Delta(V_i[\boldsymbol{\mu}])\eta(s,x)\,dx.\end{aligned}$$

Since (4.3), we obtain a lower bound on the family $-\partial_s\left(\mathcal{F}_i(\mathfrak{S}_E^s[\eta_0]) + \mathcal{V}_i(\mathfrak{S}_E^s[\eta_0]|\boldsymbol{\mu})\right)$. Indeed,

$$\begin{aligned}-\partial_s\left(\mathcal{F}_i(\mathfrak{S}_E^s[\eta_0]) \right. &+ \left. \mathcal{V}_i(\mathfrak{S}_E^s[\eta_0]|\boldsymbol{\mu})\right)\\ &\geqslant C\int_{\mathbb{R}^n}|\nabla\eta(s,x)^{m_i/2}|^2\,dx - \int_{\mathbb{R}^n}\Delta(V_i[\boldsymbol{\mu}])\eta(s,x)\,dx,\\ &\geqslant -C\end{aligned}$$

because $\|\mathfrak{S}_E^s[\eta_0]\|_{L^1(\mathbb{R}^n)} = 1$.

Then we define

$$\mathfrak{K}_{i,E}(\rho|\boldsymbol{\mu}) := C\int_{\mathbb{R}^n}|\nabla(\rho(x)^{m_i/2})|^2\,dx - \int_{\mathbb{R}^n}\Delta(V[\boldsymbol{\mu}])\rho(x)\,dx.$$

We shall now establish that $\mathfrak{K}_{i,E}$ satisfies (4.21). First of all, we notice that

$$\begin{aligned}\liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\left(\mathcal{F}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k]) + \mathcal{V}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right)\right.&\\ \left.\geqslant \liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\mathcal{F}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k])\right) + \liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\mathcal{V}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right).&\end{aligned}\tag{4.23}$$

Thanks to the proof of lemma 4.4 and with lemma A.1 of [55], we obtain

$$\liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\left(\mathcal{F}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k])\right)\right) \geqslant C\int_{\mathbb{R}^n}|\nabla(\rho_{i,h}^k(x)^{m_i/2})|^2\,dx.\tag{4.24}$$

Moreover, as $\mathfrak{S}_E^s$ is continuous in $L^1(\mathbb{R}^n)$ at $s=0$ and according to (4.3),

$$\liminf_{s\searrow 0}\left(-\frac{d}{d\sigma}_{|_{\sigma=s}}\mathcal{V}_i(\mathfrak{S}_E^\sigma[\rho_{i,h}^k]|\boldsymbol{\rho}_h^{k-1})\right) \geqslant -\int_{\mathbb{R}^n}\Delta(V_i[\boldsymbol{\rho}_h^{k-1}])\rho_{i,h}^k(x)\,dx.\tag{4.25}$$

The combination of (4.23), (4.24) and (4.25) gives (4.21) for $\mathfrak{K}_{i,E}$. We apply corollary 4.11 and we get

$$E(\rho_{i,h}^{k-1}) - E(\rho_{i,h}^k) \geqslant h\mathfrak{K}_{i,E}(\rho_{i,h}^k|\boldsymbol{\rho}_h^{k-1}).\tag{4.26}$$

But since $\Delta(V_i[\boldsymbol{\rho}]) \in L^\infty(\mathbb{R}^n)$ uniformly on $\boldsymbol{\rho}$ (4.3),

$$Ch\int_{\mathbb{R}^n}|\nabla(\rho_{i,h}^k(x)^{m_i/2})|^2\,dx \leqslant E(\rho_{i,h}^{k-1}) - E(\rho_{i,h}^k) + Ch.$$

Now we sum on k from 1 to $N = \lceil\frac{T}{h}\rceil$

$$Ch\sum_{k=1}^N\|\nabla(\rho_{i,h}^k(x)^{m_i/2})\|_{L^2(\mathbb{R}^n)}^2 \leqslant E(\rho_{i,0}) - E(\rho_{i,h}^N) + CT.\tag{4.27}$$

According to [65] and [55], there exists a constant $C>0$ and $0<\alpha<1$ such that for all $\rho\in\mathcal{P}_2^{ac}(\mathbb{R}^n)$,

$$-C(1+M(\rho))^\alpha \leqslant E(\rho) \leqslant C\mathcal{F}_i(\rho).$$

Since for all k,h, $M(\rho_{i,h}^k)$ is bounded, according to (4.14) and the fact that $\mathcal{F}_i(\rho_{i,0}) < +\infty$ by (4.6), we have

$$h\sum_{k=1}^N\|\nabla(\rho_{i,h}^k(x)^{m_i/2})\|_{L^2(\mathbb{R}^n)}^2 \leqslant C(1+T).$$

To conclude the proof, we use (4.5) and (4.15).

□

4.5 Passage to the limit

In this section, we establish weak and strong convergences for sequences $(\rho_{i,h})$, in order to pass to the limit in the discrete system of proposition 4.7.

4.5.1 Weak and strong convergences

The first convergence result is obtained using the estimates on the distance (4.16) and on the energy $\mathcal{F}_i$ (4.15).

Proposition 4.13. *Every sequences $(h_k)_{k\in\mathbb{N}}$ of time steps which tends to 0 contains a subsequence, non-relabelled, such that ρ_{i,h_k} converges, uniformly on compact time intervals, in W_2 to a $\frac{1}{2}$-Hölder function $\rho_i : [0,+\infty[\to \mathcal{P}_2^{ac}(\mathbb{R}^n)$.*

Proof. The estimation on the sum of distances (4.16) gives us for all $t, s \geqslant 0$,

$$W_2(\rho_{i,h}(t,\cdot),\rho_{i,h}(s,\cdot)) \leqslant C(|t-s|+h)^{1/2},$$

with C independ of h.

According to the proposition 3.3.1 of [4] and using a diagonal argument, at least for a subsequence, for all i, ρ_{i,h_k} converges uniformly on compact time intervals in W_2 to a $\frac{1}{2}$-Hölder function $\rho_i : [0,+\infty[\to \mathcal{P}_2(\mathbb{R}^n)$. To conclude we show that for all $t \geqslant 0$, $\rho(t,\cdot) \in \mathcal{P}_2^{ac}(\mathbb{R}^n)$. But as F_i is superlinear, Dunford-Pettis' theorem completes the proof.

□

With the previous proposition, we can pass to the limit in the case $m_i = 1$ because $P_i(\rho_{i,h}) = \rho_{i,h}$ and the term $\nabla(V_i[\boldsymbol{\rho}_h])$ is controlled thanks to the hypothesis (4.4). Unfortunately, it is not enough to pass to the limit in $P_i(\rho_{i,h})$ when $m_i > 1$. In the next proposition, we use proposition 4.12 to get a stronger convergence.

Proposition 4.14. *For all $i \in [\![1,l]\!]$ such that $m_i > 1$, $\rho_{i,h}$ converges to ρ_i in $L^{m_i}(]0,T[\times\mathbb{R}^n)$ and $P_i(\rho_{i,h})$ converges to $P_i(\rho_i)$ in $L^1(]0,T[\times\mathbb{R}^n)$, for all $T > 0$.*

The proof of this proposition is obtained by using an extention of Aubin-Lions lemma given by Rossi and Savaré in [107] (theorem 2) and recalled in [55] (theorem 4.9).

Theorem 4.15 (th. 2 in [107])**.** *On a Banach space X, let be given*

- *a normal coercive integrand $\mathcal{G} : X \to \mathbb{R}^+$, i.e, $\mathcal{G}$ is l.s.c and its sublevels are relatively compact in X,*
- *a pseudo-distance $g : X \times X \to [0,+\infty]$, i.e, g is l.s.c and*
  $[g(\rho,\mu) = 0,\ \rho,\mu \in X \text{ with } \mathcal{G}(\rho),\mathcal{G}(\mu) < \infty] \Rightarrow \rho = \mu$.

Let U be a set of measurable functions $u :]0,T[\to X$ with a fixed $T > 0$. Under the hypotheses that

$$\sup_{u\in U}\int_0^T \mathcal{G}(u(t))\,dt < +\infty \qquad \text{and} \qquad \lim_{h\searrow 0}\sup_{u\in U}\int_0^{T-h} g(u(t+h),u(t))\,dt = 0, \tag{4.28}$$

U contains a subsequence $(u_n)_{n\in\mathbb{N}}$ which converges in measure with respect to $t \in]0,T[$ to a limit $u_\star :]0,T[\to X$.

To apply this theorem, we define on $X := L^{m_i}(\mathbb{R}^n)$, as in [55], g by

$$g(\rho,\mu) := \begin{cases} W_2(\rho,\mu) & \text{if } \rho,\mu \in \mathcal{P}_2(\mathbb{R}^n), \\ +\infty & \text{otherwise,} \end{cases}$$

and $\mathcal{G}_i$ by

$$\mathcal{G}_i(\rho) := \begin{cases} \|\rho^{m_i/2}\|_{H^1(\mathbb{R}^n)} + M(\rho) & \text{if } \rho \in \mathcal{P}_2^{ac}(\mathbb{R}^n) \text{ and } \rho^{m_i/2} \in H^1(\mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases}$$

Now, we show that $\mathcal{G}_i$ satisfies theorem 4.15 conditions.

Lemma 4.16. *For all $i \in [\![1, l]\!]$ such that $m_i > 1$, $\mathcal{G}_i$ is l.s.c and its sublevels are relatively compact in $L^{m_i}(\mathbb{R}^n)$.*

Proof. The l.s.c of $\mathcal{G}_i$ on $L^{m_i}(\mathbb{R}^n)$ follows from lemma A.1 in [55]. To complete the proof we have to show that sublevels $A_c := \{\rho \in L^{m_i}(\mathbb{R}^n) \,|\, \mathcal{G}_i(\rho) \leqslant c\}$ of $\mathcal{G}_i$ are relatively compact in $L^{m_i}(\mathbb{R}^n)$.

To do this, we prove that $B_c := \left\{\eta = \rho^{m_i/2} \,|\, \rho \in A_c\right\}$ is relatively compact in $L^2(\mathbb{R}^n)$ and since the map $j : L^2(\mathbb{R}^n) \to L^{m_i}(\mathbb{R}^n)$, with $j(\eta) = \eta^{2/m_i}$, is continuous, $A_c = j(B_c)$ will be relatively compact in $L^{m_i}(\mathbb{R}^n)$.
We want to apply the Fréchet-Kolmogorov theorem to show that B_c is relatively compact in $L^2(\mathbb{R}^n)$.

- *B_c is bounded in $L^2(\mathbb{R}^n)$:* Since $\eta^2 = \rho^{m_i}$ with $\mathcal{G}_i(\rho) \leqslant c$, it is straightforward to see
$$\int_{\mathbb{R}^n} \eta^2 \leqslant c.$$

- *B_c is tight under translations:* for every $\eta \in B_c$ and $h \in \mathbb{R}^n$ we have that
$$\begin{aligned}\int_{\mathbb{R}^n} |\eta(x+h) - \eta(x)|^2 \, dx &\leqslant |h|^2 \int_{\mathbb{R}^n} \left| \int_0^1 |\nabla \eta(x + zh)| \, dz \right|^2 dx \\ &\leqslant |h|^2 \int_{\mathbb{R}^n} |\nabla \eta(x)|^2 \, dx \leqslant c|h|^2,\end{aligned}$$
thus the left hand side converges to 0 uniformly on B_c as $|h| \searrow 0$.

- *Elements of B_c are unifomly decaying at infinity:* For all $\eta \in B_c$ and $R > 0$, we have
$$\int_{|x|>R} \eta^2 \, dx \leqslant \frac{1}{R^{1/n}} \int_{\mathbb{R}^n} |x|^{1/n} \eta^{1/nm_i} \eta^{2-1/nm_i} \, dx.$$
If we use Hölder inequality with $p = 2n$ and $q = \frac{2n}{2n-1}$, we get
$$\int_{|x|>R} \eta^2 \, dx \leqslant \frac{1}{R^{1/n}} \left(\int_{\mathbb{R}^n} |x|^2 \eta^{2/m_i} \right)^{1/2n} \left(\int_{\mathbb{R}^n} \eta^{2(2m_i - 1/n)/m_i(2-1/n)} \right)^{\frac{2n-1}{2n}}.$$
As $\eta^{2/m_i} = \rho$ with $\mathcal{G}_i(\rho) \leqslant c$, we have
$$\int_{\mathbb{R}^n} |x|^2 \eta^{2/m_i} \leqslant c.$$
To bound the other term we use the Gagliardo-Nirenberg inequality: for $1 \leqslant q, r \leqslant +\infty$, we have
$$\|u\|_{L^p} \leqslant C \|\nabla u\|_{L^r}^{\alpha} \|u\|_{L^q}^{1-\alpha},$$
for all $0 < \alpha < 1$ and for p given by
$$\frac{1}{p} = \alpha \left(\frac{1}{r} - \frac{1}{n} \right) + (1 - \alpha) \frac{1}{q}.$$
We choose $p = \frac{2(2m_i - 1/n)}{m_i(2-1/n)}$, $q = r = 2$ and $\alpha = \frac{m_i - 1}{2(2m_i - 1/n)}$ (since $m_i > 1$ we have $0 < \alpha < 1$) then we obtain:
$$\int_{\mathbb{R}^n} \eta^{2(2m_i - 1/n)/m_i(2-1/n)} \leqslant \left(\int_{\mathbb{R}^n} |\nabla \eta|^2 \right)^{\alpha p/2} \left(\int_{\mathbb{R}^n} \eta^2 \right)^{(1-\alpha)p/2}.$$
but since $\eta = \rho^{m_i/2}$ with $\mathcal{G}_i(\rho) \leqslant c$, the second term is bounded then
$$\int_{|x|>R} \eta^2 \, dx \leqslant \frac{C}{R^{1/n}} \to 0,$$
as R goes to $+\infty$.

We conclude thanks to the Fréchet-Kolmogorov theorem. □

Proof of the proposition 4.14. We want to apply theorem 4.15 with $X := L^{m_i}(\mathbb{R}^n)$, $\mathcal{G} := \mathcal{G}_i$, g and $U := \{\rho_{i,h_k} \,|\, k \in \mathbb{N}\}$. According to lemma 4.16, $\mathcal{G}_i$ satisfies the hypotheses of the theorem. It's obvious that it is the same for g. Thus we only have to check conditions for U. The first condition is satisfied because of (4.14) and proposition 4.12 and the second is satisfied because of (4.16) (the proof is done in [55] proposition 4.8, for example).

According to theorem 4.15 and using a diagonal argument, there exists a subsequence, not-relabeled, such that for all i with $m_i > 1$, there exists $\tilde{\rho}_i :]0,T[\to L^{m_i}(\mathbb{R}^n)$ such that ρ_{i,h_k} converges in measure with respect to t in $L^{m_i}(\mathbb{R}^n)$ to $\tilde{\rho}_i$. Moreover, as $\rho_{i,h_k}(t)$ converges in W_2 for all $t \in [0,T]$ to $\rho_i(t)$ (proposition 4.13) then $\tilde{\rho}_i = \rho_i$. Now since convergence in measure implies a.e convergence up to a subsequence, we may also assume that $\rho_{i,h_k}(t)$ converges strongly in $L^{m_i}(\mathbb{R}^n)$ to $\rho_i(t)$ t-a.e. Now, thanks to (4.15) and (4.5) we have

$$\int_{\mathbb{R}^n} \rho_{i,h}^{m_i}(t,x)\,dx \leqslant C\mathcal{F}_i(\rho_{i,h}(t,\cdot)) \leqslant C,$$

then Lebesgue's dominated convergence theorem implies that ρ_{i,h_k} converges strongly in $L^{m_i}(]0,T[\times\mathbb{R}^n)$ to ρ_i.

To conclude the proof we have to show that $P_i(\rho_{i,h})$ converges to $P_i(\rho_i)$ in $L^1(]0,T[\times\mathbb{R}^n)$. First of all, up to a subsequence, we may assume that there exists $g \in L^{m_i}(]0,T[\times\mathbb{R}^n)$ such that

$$\rho_{i,h_k} \to \rho_i \ (t,x)\text{-a.e and } \rho_{i,h_k} \leqslant g \ (t,x)\text{-a.e.}$$

Thus according to (4.5)

$$P_i(\rho_{i,h_k}) \to P_i(\rho_i) \ (t,x)\text{-a.e and } 0 \leqslant P(\rho_{i,h_k}) \leqslant C(\rho_{i,h_k} + g^{m_i}) \ (t,x)\text{-a.e.}$$

So when we pass to the limit we have (t,x)-a.e

$$0 \leqslant P(\rho_i) \leqslant C(\rho_i + g^{m_i}) \in L^1(]0,T[\times\mathbb{R}^n).$$

Then $C(\rho_{i,h_k} + \rho_i + 2g^{m_i}) - |P_i(\rho_{i,h_k}) - P_i(\rho_i)| \geqslant 0$ and using the a.e convergence of ρ_{i,h_k} and $P_i(\rho_{i,h_k})$,

$$\begin{aligned} 2CT &+ 2C\iint_{]0,T[\times\mathbb{R}^n} g(x)^{m_i}\,dxdt \\ &= \iint_{]0,T[\times\mathbb{R}^n} \liminf \left(C(\rho_{i,h_k} + \rho_i + 2g^{m_i}) - |P_i(\rho_{i,h_k}) - P_i(\rho_i)|\right) \\ &\leqslant 2CT + 2C\iint_{]0,T[\times\mathbb{R}^n} g^{m_i}(x)\,dxdt - \limsup \iint_{]0,T[\times\mathbb{R}^n} |P_i(\rho_{i,h_k}) - P_i(\rho_i)|. \end{aligned}$$

To do these computations, we used that $\|\rho_{i,h_k}\|_{L^1(]0,T[\times\mathbb{R}^n)} = \|\rho_i\|_{L^1(]0,T[\times\mathbb{R}^n)} = T$ and Fatou's lemma. Since $g \in L^{m_i}(]0,T[\times\mathbb{R}^n)$, we obtain

$$\limsup \iint_{]0,T[\times\mathbb{R}^n} |P_i(\rho_{i,h_k}) - P_i(\rho_i)| \leqslant 0,$$

which concludes the proof.

□

4.5.2 Limit of the discrete system

In this section, we pass to the limit in the discrete system of proposition 4.7. In the following, we consider $\phi_i \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$ and $N = \lceil \frac{T}{h} \rceil$.

proof of theorem 4.3. We will pass to the limit in all terms in proposition 4.7.

- *Convergence of the remainder term:* By definition of $\mathcal{R}$, we have

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \mathcal{R}[\phi_i(t_k,\cdot)](x,y)d\gamma_{i,h}^k(x,y) \leqslant \frac{1}{2}\|\nabla^2\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)} W_2^2(\rho_{i,h}^k,\rho_{i,h}^{k+1}).$$

and according to the estimate (4.16), we get

$$\left|\sum_{k=0}^{N-1}\int_{\mathbb{R}^n\times\mathbb{R}^n} \mathcal{R}[\phi_i(t_k,\cdot)](x,y)d\gamma_{i,h}^k(x,y)\right| \leqslant C\sum_{k=0}^{N-1} W_2^2(\rho_{i,h}^k,\rho_{i,h}^{k+1}) \leqslant Ch \to 0.$$

- *Convergence of the linear term:*

$$\left|\int_0^T\int_{\mathbb{R}^n}\rho_{i,h}(t,x)\partial_t\phi_i(t,x)\,dxdt - \int_0^T\int_{\mathbb{R}^n}\rho_i(t,x)\partial_t\phi_i(t,x)\,dxdt\right| \leqslant CT\sup_{t\in[0,T]} W_2(\rho_{i,h}(t,\cdot),\rho_i(t,\cdot)) \to 0,$$

when $h \searrow 0$ because of proposition 4.13.

- *Convergence of the diffusion term:*

$$\left|h\sum_{k=0}^{N-1}\int_{\mathbb{R}^n} P_i(\rho_{i,h}^{k+1}(x))\cdot\Delta\phi_i(t_k,x)\,dx - \int_0^T\int_{\mathbb{R}^n} P_i(\rho_i(t,x))\Delta\phi_i(t,x)\,dxdt\right|$$
$$\leqslant C(1+T)\|D^3\phi_i\|_{L^\infty}h + \left|\int_0^T\int_{\mathbb{R}^n}\left(P_i(\rho_{i,h}(t,x)) - P_i(\rho(t,x))\right)\Delta\phi_i(t,x)\,dxdt\right|.$$

If $m_i = 1$, the right hand side converges to 0 because of proposition 4.13 and otherwise it goes to 0 because of proposition 4.14.

- *Convergence of the interaction term:*

$$\begin{aligned}
&\left|h\sum_{k=0}^{N-1}\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}_h^k])(x)\cdot\nabla\phi_i(t_k,x)\rho_{i,h}^{k+1}(x)\,dx\right.\\
&\left.-\int_0^T\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}(t,\cdot)])(x)\cdot\nabla\phi_i(t,x)\rho_i(t,x)\,dxdt\right|\\
\leqslant\ &\left|h\sum_{k=0}^{N-1}\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}_h^k])(x)\cdot\nabla\phi_i(t_k,x)\rho_{i,h}^{k+1}(x)\,dx\right.\\
&\left.-\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}(t,\cdot)])(x)\cdot\nabla\phi_i(t_k,x)\rho_{i,h}^{k+1}(x)\,dxdt\right|\\
+\ &\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}(t,\cdot)])(x)\cdot(\nabla\phi_i(t_k,x)-\nabla\phi_i(t,x))\rho_{i,h}^{k+1}(x)\,dxdt\right|\\
+\ &\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}^n}\nabla(V_i[\boldsymbol{\rho}(t,\cdot)])(x)\cdot\nabla\phi_i(t,x)(\rho_{i,h}^{k+1}(x)-\rho_i(t,x))\,dxdt\right|\\
\leqslant\ &J_1+J_2+J_3.
\end{aligned}$$

  - As $\rho_{i,h}$ converges weakly $L^1(]0,T[\times\mathbb{R}^n)$ to ρ_i and $(\nabla(V_i[\boldsymbol{\rho}])\cdot\nabla\phi_i) \in L^\infty([0,T]\times\mathbb{R}^n)$, then
$$J_3 \to 0 \text{ as } h \to 0.$$
  - For J_2, we use the fact that $\nabla\phi_i$ is a Lipschitz function and that $\nabla(V_i[\boldsymbol{\rho}])$ is bounded thanks to (4.3), and then,
$$J_2 \leqslant CT\|D^2\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)}h \to 0.$$

– Using assumption (4.4), we have

$$J_1 \leqslant C\|\nabla\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} W_2(\boldsymbol{\rho}_h^k, \boldsymbol{\rho}(t,\cdot))\, dt$$

Then using triangle inequality and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} J_1 &\leqslant C\|\nabla\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} (W_2(\boldsymbol{\rho}_h^k, \boldsymbol{\rho}_h^{k+1}) + W_2(\boldsymbol{\rho}_h^{k+1}, \boldsymbol{\rho}(t,\cdot))) \\ &\leqslant C\|\nabla\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)} \left(h \sum_{k=0}^{N-1} W_2(\boldsymbol{\rho}_h^k, \boldsymbol{\rho}_h^{k+1}) \right. \\ &\qquad \left. + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} W_2(\boldsymbol{\rho}_h^{k+1}, \boldsymbol{\rho}(t,\cdot))\, dt \right) \\ &\leqslant C\|\nabla\phi_i\|_{L^\infty([0,T]\times\mathbb{R}^n)} \left(T \sum_{k=0}^{N-1} W_2^2(\boldsymbol{\rho}_h^k, \boldsymbol{\rho}_h^{k+1}) \right. \\ &\qquad \left. + \int_0^T W_2(\boldsymbol{\rho}_h(t,\cdot), \boldsymbol{\rho}(t,\cdot))\, dt \right) \end{aligned}$$

According to (4.16), we obtain

$$T \sum_{k=0}^{N-1} W_2^2(\boldsymbol{\rho}_h^k, \boldsymbol{\rho}_h^{k+1}) \leqslant CTh \to 0$$

when $h \searrow 0$. Moreover,

$$\int_0^T W_2(\boldsymbol{\rho}_h(t,\cdot), \boldsymbol{\rho}(t,\cdot))\, dt \leqslant T \sup_{t\in[0,T]} W_2(\boldsymbol{\rho}_h(t,\cdot), \boldsymbol{\rho}(t,\cdot)) \to 0,$$

when h goes to 0, which proves that

$$J_1 \to 0 \text{ as } h \to 0.$$

If we combine all these convergences, theorem 4.3 is proved. □

4.6 The case of a bounded domain Ω

In this section, we work on a smooth bounded domain Ω of $\mathbb{R}^n$ and only with one density but, as in the whole space, the result readily extends to systems. Our aim is to solve (4.1). We remark that Ω is not taken convex so we can not use the flow interchange argument anymore because this argument uses the displacement convexity of the Entropy. Moreover since Ω is bounded, the solution has to satisfy some boundary conditions contrary to the periodic case [35] or in $\mathbb{R}^n$. In our case, we study (4.1) with no flux boundary condition, which is the natural boundary condition for gradient flows, i.e we want to solve

$$\begin{cases} \partial_t\rho - \operatorname{div}(\rho\nabla(V[\rho])) - \Delta P(\rho) = 0 & \text{on } \mathbb{R}^+\times\Omega, \\ (\rho\nabla(V[\rho]) + \nabla P(\rho))\cdot\nu = 0 & \text{on } \mathbb{R}^+\times\partial\Omega, \\ \rho(0,\cdot) = \rho_0 & \text{on } \mathbb{R}^n, \end{cases} \tag{4.29}$$

where ν is the outward unit normal to $\partial\Omega$.

We say that $\rho : [0,+\infty[\to \mathcal{P}^{ac}(\Omega)$ is a weak solution of (4.29), with $F \in \mathcal{H}_m$, if $\rho \in \mathcal{C}([0,+\infty[;\mathcal{P}^{ac}(\Omega)) \cap L^m(]0,T[\times\Omega)$, $P(\rho) \in L^1(]0,T[\times\Omega)$, $\nabla P(\rho) \in \mathcal{M}^n([0,T]\times\Omega)$ for all $T < \infty$ and if for all $\varphi \in \mathcal{C}_c^\infty([0,+\infty[\times\mathbb{R}^n)$, we have

$$\int_0^\infty \int_\Omega [(\partial_t\varphi - \nabla\varphi\cdot\nabla(V[\rho]))\rho - \nabla P(\rho)\cdot\nabla\varphi] = -\int_\Omega \varphi(0,x)\rho_0(x).$$

Since test functions are in $\mathcal{C}_c^\infty([0,+\infty[\times\mathbb{R}^n)$, we do not impose that they vanish on the boundary of Ω, which give Neumann boundary condition.

Theorem 4.17. *Let $F\in\mathcal{H}_m$ for $m\geqslant 1$ and let V satisfies (4.2), (4.3), (4.4). If we assume that $\rho_0\in\mathcal{P}^{ac}(\Omega)$ satisfies*

$$\mathcal{F}(\rho_0)+\mathcal{V}(\rho_0|\rho_0)<+\infty, \tag{4.30}$$

with

$$\mathcal{F}(\rho):=\begin{cases}\int_\Omega F(\rho(x))\,dx & \text{if } \rho\ll\mathcal{L}^n_{|\Omega},\\ +\infty & \text{otherwise,}\end{cases}\quad \text{and } \mathcal{V}(\rho|\mu):=\int_\Omega \mathcal{V}[\mu]\rho\,dx.$$

then (4.29) admits at least one weak solution.

The proof of this theorem is different from the one on $\mathbb{R}^n$ because we will not use the flow interchange argument of Matthes, McCann and Savaré to find strong convergence since Ω is not assumed convex. First, we will find an a.e equality using the first variation of energies in order to have a discrete equation, as in proposition 4.7. Then, we will derive a new estimate on the gradient of some power of ρ_h from this a.e equality. To conclude, we will use again the refined version of Aubin-Lions lemma of Rossi and Savaré in [107].

On Ω we can define, with the semi-implicit JKO scheme, the sequence $(\rho_h^k)_k$, where ρ_h^k minimizes

$$\rho\mapsto\mathcal{E}_h(\rho|\rho_h^{k-1}):=\frac{1}{2h}W_2^2(\rho,\rho_h^{k-1})+\mathcal{F}(\rho)+\mathcal{V}(\rho|\rho_h^{k-1})$$

on $\mathcal{P}(\Omega)$. The proof of existence and uniqueness of ρ_h^k is the same as in proposition 4.5. It is even easier because on a bounded domain $\mathcal{F}$ is bounded from below for all $m\geqslant 1$. We find also the same estimates than in the proposition 4.8 on the functional and the distance (see for example [1],[35]).

Now we will establish a discrete equation satisfied by the piecewise interpolation of the sequence $(\rho_h^k)_k$ defined by, for all $k\in\mathbb{N}$,

$$\rho_h(t)=\rho_h^k \text{ if } t\in((k-1)h,kh].$$

Proposition 4.18. *For every $k\geqslant 0$, we have*

$$(y-T_k(y))\rho_h^{k+1}+h\nabla(V[\rho_h^k])\rho_h^{k+1}+h\nabla(P(\rho_h^{k+1}))=0 \ \text{ a.e on } \Omega, \tag{4.31}$$

where T_k is the optimal transport map between ρ_h^{k+1} and ρ_h^k. Then ρ_h satisfies

$$\begin{aligned}\int_0^T\int_\Omega \rho_h(t,x)\partial_t\varphi(t,x)\,dxdt &= h\sum_{k=0}^{N-1}\int_\Omega \nabla(V[\rho_h^k])(x)\cdot\nabla\varphi(t_k,x)\rho_h^{k+1}(x)\,dxdt\\ &\quad +h\sum_{k=0}^{N-1}\int_\Omega \nabla P(\rho_h^{k+1}(x))\cdot\nabla\varphi(t_k,x)\,dx\\ &\quad +\sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\varphi(t_k,\cdot)](x,y)d\gamma_k(x,y)\\ &\quad -\int_\Omega\rho_0(x)\varphi(0,x)\,dx,\end{aligned} \tag{4.32}$$

with $N=\lceil\frac{T}{h}\rceil$, for all $\phi\in\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$, γ_k is the optimal transport plan in $W_2(\rho_h^k,\rho_h^{k+1})$ and

$$|\mathcal{R}[\phi](x,y)|\leqslant\frac{1}{2}\|D^2\phi\|_{L^\infty(\mathbb{R}\times\mathbb{R}^n)}|x-y|^2.$$

Proof. First, we prove the equality (4.31). As in proposition 4.7, taking the first vartiation in the semi-implicit JKO scheme, we find for all $\xi \in \mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega (y - T_k(y)) \cdot \xi(y)\rho_h^{k+1}(y)\,dy + h\int_\Omega \nabla(V[\rho_h^k]) \cdot \xi \rho_h^{k+1} - h\int_\Omega P(\rho_h^{k+1})\,\mathrm{div}(\xi) = 0, \tag{4.33}$$

where T_k is the optimal transport map between ρ_h^{k+1} and ρ_h^k. Now we claim that $P(\rho_h^{k+1}) \in W^{1,1}(\Omega)$. Indeed, since F controls x^m and P is controlled by x^m then (4.15) gives $P(\rho_h^{k+1}) \in L^1(\Omega)$. Moreover, (4.33) gives

$$\begin{aligned}\left|\int_\Omega P(\rho_h^{k+1})\,\mathrm{div}(\xi)\right| &\leqslant \left[\int_\Omega \frac{|y - T_k(y)|}{h}\rho_h^{k+1} + C\right]\|\xi\|_{L^\infty(\Omega)} \\ &\leqslant \left[\frac{W_2(\rho_h^k,\rho_h^{k+1})}{h} + C\right]\|\xi\|_{L^\infty(\Omega)}.\end{aligned}$$

This implies $P(\rho_h^{k+1}) \in BV(\Omega)$ and $\nabla P(\rho_h^{k+1}) = \nabla V[\rho_h^k]\rho_h^{k+1} + \frac{Id - T_k}{h}\rho_h^{k+1}$ in $\mathcal{M}^n(\Omega)$. And, since $\nabla V[\rho_h^k]\rho_h^{k+1} + \frac{Id-T_k}{h}\rho_h^{k+1} \in L^1(\Omega)$, we have $P(\rho_h^{k+1}) \in W^{1,1}(\Omega)$ and (4.31).

Now, we verify that ρ_h statisfies (4.32). We start to take the scalar product between (4.31) and $\nabla\varphi$ with $\varphi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$, and we find, for all $t \in [0,T)$,

$$\begin{aligned}&\int_\Omega (y - T_k(y)) \cdot \nabla\varphi(t,y)\rho_h^{k+1}(y)\,dy \\ &\qquad + h\int_\Omega \nabla(V[\rho_h^k])(y) \cdot \nabla\varphi(t,y)\rho_h^{k+1}(y)\,dy + h\int_\Omega \nabla(P(\rho_h^{k+1}))(y) \cdot \nabla\varphi(t,y)\,dy = 0.\end{aligned} \tag{4.34}$$

Moreover, if we extend φ by $\varphi(0,\cdot)$ on $[-h,0)$, then

$$\begin{aligned}\int_0^T\int_\Omega \rho_h(t,x)\partial_t\varphi(t,x)\,dxdt &= \sum_{k=0}^N \int_{t_{k-1}}^{t_k}\int_\Omega \rho_h^k(x)\partial_t\varphi(t,x)\,dxdt \\ &= \sum_{k=0}^N \int_\Omega \rho_h^k(x)(\varphi(t_k,x) - \varphi(t_{k-1},x))\,dx \\ &= \sum_{k=0}^{N-1}\int_\Omega \varphi(t_k,x)(\rho_h^k(x) - \rho_h^{k+1}(x))\,dx \\ &- \int_{\mathbb{R}^n}\rho_0(x)\varphi(0,x)\,dx.\end{aligned}$$

And using the second order Taylor-Lagrange formula, we find

$$\begin{aligned}&\int_{\Omega\times\Omega}(\varphi(kh,x) - \varphi(kh,y))\,d\gamma_k(x,y) \\ &\qquad = \int_{\Omega\times\Omega}\nabla\varphi(kh,y)\cdot(x-y)\,d\gamma_k(x,y) + \int_{\Omega\times\Omega}\mathcal{R}[\varphi(t_k,\cdot)](x,y)d\gamma_k(x,y).\end{aligned}$$

This concludes the proof if we sum on k and use (4.34).

□

Remark 4.19. *We remark that equality* (4.31) *is still true in* $\mathbb{R}^n$*. Indeed, the first part of the proof does not depend of the domain and we can use this argument on* $\mathbb{R}^n$*. This equality will be used in section 4.7 to obtain uniqueness result.*

In the next proposition, we propose an alternative argument to the flow interchange argument to get an estimate on the gradient of ρ_h. Differences with the flow interchange argument are that we do not need to assume the space convexity and boundary condition on $\nabla V[\rho]$. Moreover we do not obtain exactly the same estimate. Indeed, in proposition 4.12, $\nabla\rho_h^{m/2}$ is bounded in $L^2((0,T)\times\mathbb{R}^n)$ whereas in the following proposition we establish a bound on $\nabla\rho_h^m$ in $L^1((0,T)\times\mathbb{R}^n)$ using (4.31).

Proposition 4.20. *There exists a constant C which does not depend on h such that*

$$\|\rho_h^m\|_{L^1([0,T];W^{1,1}(\Omega))} \leqslant CT$$

for all $T > 0$.

Proof. According to (4.31), we have

$$h\int_\Omega |\nabla(P(\rho_h^{k+1}))|\,dx \quad \leqslant \quad W_2(\rho_h^k, \rho_h^{k+1}) + hC.$$

Then if we sum on k from 0 to $N-1$, we get

$$\begin{aligned}\int_0^T\int_\Omega |\nabla(P(\rho_h))|\,dxdt &\leqslant \sum_{k=0}^{N-1} W_2(\rho_h^k, \rho_h^{k+1}) + TC\\ &\leqslant N\sum_{k=0}^{N-1} W_2^2(\rho_h^k, \rho_h^{k+1}) + TC\\ &\leqslant CT,\end{aligned}$$

because of (4.16).

If $F(x) = x\log(x)$ then $P'(x) = 1$ and if F satisfies (4.5), then $F''(x) \geqslant Cx^{m-2}$ and $P'(x) = xF''(x) \geqslant Cx^{m-1}$. In both cases, we have $P'(x) \geqslant Cx^{m-1}$ (with $m = 1$ for $x\log(x)$). So

$$\begin{aligned}\int_0^T\int_\Omega |\nabla(P(\rho_h))|\,dxdt = \int_0^T\int_\Omega P'(\rho_h)|\nabla\rho_h|\,dxdt&\\ \geqslant C\int_0^T\int_\Omega \rho_h^{m-1}|\nabla\rho_h|\,dxdt = C\int_0^T\int_\Omega |\nabla\rho_h^m|\,dxdt,&\end{aligned}$$

Which proves the proposition. □

Now we introduce $\mathcal{G} : L^m(\Omega) \to [0, +\infty]$ defined by

$$\mathcal{G}(\rho) := \begin{cases} \|\rho^m\|_{BV(\Omega)} & \text{if } \rho \in \mathcal{P}^{ac}(\Omega) \text{ and } \rho^m \in BV(\Omega),\\ +\infty & \text{otherwise.}\end{cases}$$

Proposition 4.21. *$\mathcal{G}$ is lower semi-continuous on $L^m(\Omega)$ and its sublevels are relatively compact in $L^m(\Omega)$.*

Proof. First we show that $\mathcal{G}$ is lower semi-continuous on $L^m(\Omega)$. Let ρ_n be a sequence which converges strongly to ρ in $L^m(\Omega)$ with $\sup_n \mathcal{G}(\rho_n) \leqslant C < +\infty$. Without loss of generality, we assume that ρ_n converges to ρ a.e.
Since $C < +\infty$, the functions ρ_n^m are uniformly bounded in $BV(\Omega)$. So we know that ρ_n^m converges weakly in $BV(\Omega)$ to μ. But since Ω is smooth and bounded, the injection of $BV(\Omega)$ into $L^1(\Omega)$ is compact. We can deduce that $\mu = \rho^m$ and ρ_n^m converges to ρ^m strongly in $L^1(\Omega)$. Then by lower semi-continuity of the BV-norm in L^1, we obtain

$$\mathcal{G}(\rho) \leqslant \liminf_{n\nearrow+\infty} \mathcal{G}(\rho_n).$$

Now, we have to prove that the sublevels, $A_c := \{\rho \in L^m(\Omega) : \mathcal{G}(\rho) \leqslant c\}$, are relatively compact in $L^m(\Omega)$. Since $i : \eta \in L^1(\Omega) \mapsto \eta^{1/m} \in L^m(\Omega)$ is continuous, we just have to prove that $B_c := \{\eta = \rho^m : \rho \in A_c\}$ is relatively compact in $L^1(\Omega)$. So to conclude the proof, it is enough to notice that B_c is a bounded subset of $BV(\Omega)$ and that the injection of $BV(\Omega)$ into $L^1(\Omega)$ is compact. □

Now we can apply Rossi-Savaré theorem (theorem 4.15) to have the strong convergence in $L^m(]0,T[\times\Omega)$ of ρ_h to ρ and then we find the strong convergence in $L^1(]0,T[\times\Omega)$ of $P(\rho_h)$ to $P(\rho)$, for all $T > 0$, using the fact that P is controlled by x^m (4.5) and Krasnoselskii theorem (see [50], chapter 2).

Moreover, since

$$\int_0^T \int_\Omega |\nabla(P(\rho_h))|\, dxdt \leqslant CT,$$

we have

$$\nabla(P(\rho_h))dxdt \rightharpoonup \mu \text{ in } \mathcal{M}^n([0,T]\times\Omega), \tag{4.35}$$

i.e

$$\int_0^T \int_\Omega \xi \cdot \nabla(P(\rho_h))dxdt \to \int_0^T \int_\Omega \xi \cdot d\mu,$$

for all $\xi \in \mathcal{C}_b([0,T]\times\Omega)$ (this means that we do not require ξ to vanish on $\partial\Omega$). But since $P(\rho_h)$ converges strongly to $P(\rho)$ in $L^1([0,T]\times\Omega)$, $\mu = \nabla(P(\rho))$.

To conclude, we pass to the limit in (4.32) and theorem 4.17 follows.

4.7 Uniqueness of solutions

In this section, we assume that Ω is a convex subset of $\mathbb{R}^n$, not necessarily bounded. Let $\mathcal{F}$ an energy defined on $\mathcal{P}_2(\Omega)$. Let $\rho, \mu \in \mathcal{P}_2(\Omega)$, we recall that the geodesics for W_2 are of the form $\rho_t := \pi_{t\#}\gamma$, where γ is an optimal transport plan for $W_2(\rho,\mu)$ and $\pi_t(x,y) := (1-t)x + ty$.

$\mathcal{F}$ is said displacement convex if

$$t \in [0,1] \mapsto \mathcal{F}(\rho_t) \text{ is convex.}$$

It is well known that if $\mathcal{F}$ is displacement convex and $\nabla F'(\rho) \in L^2(\rho)$, then

$$\mathcal{F}(\nu) - \mathcal{F}(\rho) \geqslant \int_\Omega \nabla F'(\rho(x)) \cdot (T_\rho^\nu(x) - x)\, d\rho(x),$$

for all $\nu \in \mathcal{P}_2^{ac}(\Omega)$, where T_ρ^ν is the W_2-optimal transport map between ρ and ν. In particular, if $\nabla F'(\rho) \in L^2(\rho)$ and $\nabla F'(\nu) \in L^2(\nu)$, then

$$\int_\Omega (\nabla F'(\nu(T_\rho^\nu(x)) - \nabla F'(\rho(x))) \cdot (x - T_\rho^\nu(x))\, d\rho(x) \leqslant 0. \tag{4.36}$$

These results can be seen in the caracterization of the subdifferentiel of λ-geodesic convex fonctional of [4]. Now we state a general uniqueness argument based on geodesic convexity.

Theorem 4.22. *Assume that V_i satisfy* (4.3) *and* (4.4) *and F_i such that $\mathcal{F}_i$ is displacement convex. Let $\boldsymbol{\rho}^1 := (\rho_1^1, \dots, \rho_l^1)$ and $\boldsymbol{\rho}^2 := (\rho_1^2, \dots, \rho_l^2)$ two weak solutions of* (4.1) *or* (4.17) *with initial conditions $\rho_i^1(0,\cdot) = \rho_{i,0}^1$ and $\rho_i^2(0,\cdot) = \rho_{i,0}^2$. If for all $T < +\infty$,*

$$\int_0^T \sum_{i=1}^l \|v_{i,t}^1\|_{L^2(\rho_{i,t}^1)}\, dt + \int_0^T \sum_{i=1}^l \|v_{i,t}^2\|_{L^2(\rho_{i,t}^2)}\, dt < +\infty, \tag{4.37}$$

with, for $j \in \{1,2\}$,

$$v_{i,t}^j := -\frac{\nabla P_i(\rho_{i,t}^j)}{\rho_{i,t}^j} - \nabla V_i[\boldsymbol{\rho}_t^j],$$

then for every $t \in [0,T]$,

$$W_2^2(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2) \leqslant e^{4Ct} W_2^2(\boldsymbol{\rho}_0^1, \boldsymbol{\rho}_0^2).$$

In particular, we have uniqueness for the Cauchy problems (4.1) *and* (4.17).

Proof. The proof is a little perturbation of the one of theorem 11.1.4 of [4] and is based on displacement convexity argument and Gronwall's lemma. Using Theorem 5.24 and Corollary 5.25 from [112], assumption (4.37) guarantees that $t \mapsto W_2^2(\rho_{i,t}^1, \rho_{i,t}^2)$ is differentiable for a.e $t \in (0,T)$ with

$$\frac{d}{dt} W_2^2(\rho_{i,t}^1, \rho_{i,t}^2) = \int_\Omega (x - T_1^2(x)) \cdot (v_{i,t}^2(T_1^2(x)) - v_{i,t}^1(x))\rho_{i,t}^1(x)\, dx,$$

where T_1^2 is the optimal map between $\rho_{i,t}^1$ and $\rho_{i,t}^2$. Now, since $\mathcal{F}_i$ is displacement convex and recalling (4.36), we have

$$\int_\Omega (\nabla F'(\rho_{i,t}^2(T_1^2(x))) - \nabla F'(\rho_{i,t}^1(x))) \cdot (x - T_1^2(x))\rho_{i,t}^1(x)\,dx \leqslant 0,$$

so that, using (4.3)-(4.4) yields

$$\begin{aligned}
\frac{d}{dt}W_2^2(\rho_{i,t}^1, \rho_{i,t}^2) &\leqslant \int_\Omega (x - T_1^2(x)) \cdot (\nabla V_i[\boldsymbol{\rho}_t^2](T_1^2(x)) - \nabla V_i[\boldsymbol{\rho}_t^1](x))\rho_{i,t}^1(x)\,dx \\
&\leqslant \int_\Omega (x - T_1^2(x)) \cdot (\nabla V_i[\boldsymbol{\rho}_t^2](T_1^2(x)) - \nabla V_i[\boldsymbol{\rho}_t^2](x))\rho_{i,t}^1(x)\,dx \\
&+ \int_\Omega (x - T_1^2(x)) \cdot (\nabla V_i[\boldsymbol{\rho}_t^2](x) - \nabla V_i[\boldsymbol{\rho}_t^2](x))\rho_{i,t}^1(x)\,dx \\
&\leqslant C\int_\Omega |x - T_1^2(x)|^2 \rho_{i,t}^1(x)\,dx \\
&+ C\left(\int_\Omega |x - T_1^2(x)|^2 \rho_{i,t}^1(x)\,dx\right)^{1/2} \sum_{j=1}^{l} W_2(\rho_{j,t}^1, \rho_{j,t}^2) \\
&\leqslant C\left(W_2^2(\rho_{i,t}^1, \rho_{i,t}^2) + W_2(\rho_{i,t}^1, \rho_{i,t}^2)\sum_{j=1}^{l} W_2(\rho_{j,t}^1, \rho_{j,t}^2)\right).
\end{aligned}$$

Summing over i and using Gronwall's lemma gives the desired inequality. □

Remark 4.23. *We say that $F : [0, +\infty) \to \mathbb{R}$ satisfies McCann's condition if*

$$x \in (0, +\infty) \mapsto x^n F(x^{-n}) \text{ is convex nonincreasing.} \tag{4.38}$$

McCann showed in [93] that if F satisfies (4.38), then $\mathcal{F}$ is displacement convex. Then if for all i, F_i satisfies McCann's condition, we have uniqueness in (4.1). Moreover we remark that $F(x) = x\log(x)$ (linear diffusion) and $F(x) = \frac{1}{m-1}x^m, m > 1$ (porous medium diffusion) satisfy this condition.

In the following proposition, we will prove that assumption (4.37) holds if Ω is a smooth bounded convex subset of $\mathbb{R}^n$ or if $\Omega = \mathbb{R}^n$.

Proposition 4.24. *Let $\boldsymbol{\rho} := (\rho_1, \dots, \rho_l)$ be a weak solution of (4.1) obtained with the previous semi-implicit JKO scheme. Then ρ_i satisfies (4.37) for all $i \in [\![1, l]\!]$.*

Proof. We do not separate the cases where Ω is a bounded set or is $\mathbb{R}^n$. We split the proof in two parts. First, we show that (4.37) is satisfied by $\rho_{i,h}$ defined in (4.9). Then by a l.s.c argument we will conclude the proof.

- In the first step, we show that $\rho_{i,h}$ satisfies

$$\int_0^T \int_\Omega |\nabla F_i'(\rho_{i,h}) + \nabla V_i[\boldsymbol{\rho}_h]|^2 \rho_{i,h}\,dxdt \leqslant C, \tag{4.39}$$

  where C does not depend of h.

  By equality (4.31) and remark 4.19, we have

$$\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^k] = \frac{T_k(y) - y}{h} \qquad \rho_{i,h}^{k+1} - \text{a.e on } \Omega,$$

  where T_k is the optimal transport map between $\rho_{i,h}^{k+1}$ and $\rho_{i,h}^k$. Then if we take the square, multiply by $\rho_{i,h}^{k+1}$ and integrate on Ω, we find

$$\int_\Omega |\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^k]|^2 \rho_{i,h}^{k+1}\,dx = \frac{1}{h^2}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k).$$

Now using (4.4), we get

$$\begin{aligned}|\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^{k+1}]| &\leqslant |\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^{k}]| + |\nabla V_i[\boldsymbol{\rho}_h^{k}] - \nabla V_i[\boldsymbol{\rho}_h^{k+1}]| \\ &\leqslant |\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^{k}]| + CW_2(\boldsymbol{\rho}_h^{k+1}, \boldsymbol{\rho}_h^{k})\end{aligned}$$

So we have

$$\int_\Omega |\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^{k+1}]|^2 \rho_{i,h}^{k+1}\,dx \leqslant C\left(\frac{1}{h^2}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^{k}) + W_2^2(\boldsymbol{\rho}_h^{k+1}, \boldsymbol{\rho}_h^{k})\right).$$

Then using (4.16), we finally get

$$\begin{aligned}\int_0^T \int_\Omega |\nabla F_i'(\rho_{i,h}) &+ \nabla V_i[\boldsymbol{\rho}_h]|^2 \rho_{i,h}\,dxdt \\ &= h\sum_{k=0}^{N-1} \int_\Omega |\nabla F_i'(\rho_{i,h}^{k+1}) + \nabla V_i[\boldsymbol{\rho}_h^{k+1}]|^2 \rho_{i,h}^{k+1}\,dx \\ &\leqslant C\left(\frac{1}{h}\sum_{k=0}^{N-1} W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^{k}) + 1\right) \\ &\leqslant C.\end{aligned}$$

- To conclude, we have to pass to the limit in (4.39). First, we claim that $\nabla P_i(\rho_{i,h})$ converges to $\nabla P_i(\rho_i)$ in $\mathcal{M}^n([0,T]\times\Omega)$. In a bounded set, this has been proved in (4.35). In $\mathbb{R}^n$ thanks to the previous step, we have

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^n} |\nabla P_i(\rho_{i,h})|dt &= \int_0^T \int_{\mathbb{R}^n} |\nabla F_i'(\rho_{i,h})|\rho_{i,h}\,dxdt \\ &\leqslant \int_0^T \int_{\mathbb{R}^n} (|\nabla F_i'(\rho_{i,h})|^2 + 1)\rho_{i,h} \\ &\leqslant C,\end{aligned}$$

which gives the result because $P_i(\rho_{i,h})$ strongly converges in $L^1([0,T]\times\mathbb{R}^n)$ to $P_i(\rho_i)$.

Let $\psi : \mathbb{R}^{n+1} \to \mathbb{R}\cup\{+\infty\}$ defined by

$$\psi(r,m) := \begin{cases} \frac{|m|^2}{r} & \text{if } (r,m) \in]0,+\infty[\times\mathbb{R}^n, \\ 0 & \text{if } (r,m) = (0,0), \\ +\infty & \text{otherwise,}\end{cases}$$

as in [11]. And define $\Psi : \mathcal{M}((0,T)\times\Omega) \times \mathcal{M}^n((0,T)\times\Omega) \to \mathbb{R}\cup\{+\infty\}$, as in [24], by

$$\Psi(\rho,E) := \begin{cases} \int_0^T \int_\Omega \psi(d\rho/d\mathcal{L}, dE/d\mathcal{L})\,dxdt & \text{if } \rho \geqslant 0, \\ +\infty & \text{otherwise,}\end{cases}$$

where $d\sigma/d\mathcal{L}$ is Radon-Nikodym derivative of σ with respect to $\mathcal{L}_{|[0,T]\times\Omega}$. We can remark that since $\psi(0,m) = +\infty$ for any $m \neq 0$, we have

$$\Psi(\rho,E) < +\infty \Rightarrow E \ll \rho.$$

With this definition, we can rewrite (4.39) as

$$\Psi(\rho_{i,h}, \nabla P_i(\rho_{i,h}) + \nabla V_i[\boldsymbol{\rho}_h]\rho_{i,h}) = \int_0^T \int_\Omega |\nabla F_i'(\rho_{i,h}) + \nabla V_i[\boldsymbol{\rho}_h]|^2 \rho_{i,h}\,dxdt \leqslant C,$$

which, in particular, implies that $\nabla P_i(\rho_{i,h}) \ll \rho_{i,h} \ll \mathcal{L}_{|[0,T]\times\Omega}$.

Moreover, according to [19], Ψ is lower semicontinuous on $\mathcal{M}([0,T]\times\Omega)\times\mathcal{M}^n([0,T]\times\Omega)$. So, it holds

$$\Psi(\rho_i, \nabla P_i(\rho_i) + \nabla V_i[\boldsymbol{\rho}]\rho_i) \leqslant \liminf_{h\searrow 0} \Psi(\rho_{i,h}, \nabla P_i(\rho_{i,h}) + \nabla V_i[\boldsymbol{\rho}_h]\rho_{i,h}) \leqslant C,$$

which imply $\nabla P_i(\rho_i) \ll \rho_i \ll \mathcal{L}_{|[0,T]\times\Omega}$ and conclude the proof because

$$\begin{aligned}\int_0^T \int_\Omega \left| \frac{\nabla P_i(\rho_i)}{\rho_i} + \nabla V_i[\boldsymbol{\rho}] \right|^2 \rho_i \, dxdt &= \int_0^T \int_\Omega \frac{|\nabla P_i(\rho_i) + \nabla V_i[\boldsymbol{\rho}]\rho_i|^2}{\rho_i} \, dxdt \\ &= \Psi(\rho_i, \nabla P_i(\rho_i) + \nabla V_i[\boldsymbol{\rho}]\rho_i) \\ &\leqslant C.\end{aligned}$$

□

4.8 Numerical simulations

4.8.1 Nonlocal interactions

If we consider a nonlocal interaction term $\int_{\Omega\times\Omega} W(x,y)\rho(x)\rho(y)\mathrm{d}x\mathrm{d}y$ (with W symmetric and smooth) in the general form of the equation

$$\partial_t \rho = \mathrm{div}(\rho\nabla(F'(\rho) + V + W\star\rho)),\ \rho|_{t=0} = \rho_0, \tag{4.40}$$

the final term $\mathcal{E}$,

$$\mathcal{E}(\rho) := \int_\Omega F(\rho(x))\mathrm{d}x + \int_\Omega V(x)\rho(x)\mathrm{d}x + \int_{\Omega\times\Omega} W(x,y)\rho(x)\rho(y)\mathrm{d}x\mathrm{d}y,$$

becomes nonconvex. Therefore, in order to be able to use our Augmented Lagrangian strategy, we have to modify the JKO scheme in a semi-implicit way by replacing the nonconvex bilinear term $\frac{1}{2}\int_{\Omega\times\Omega} W(x,y)\rho(x)\rho(y)\mathrm{d}x\mathrm{d}y$ by the linear one $\int_\Omega W(x,y)\rho(x)\rho^k(y)\mathrm{d}x\mathrm{d}y$. This leads to the semi-implicit scheme

$$\rho^{k+1} \in \operatorname*{argmin}_{\rho\in\mathcal{P}_2} \left\{ \frac{1}{2\tau} W_2^2(\rho,\rho^k) + \mathcal{E}(\rho|\rho^k) \right\} \tag{4.41}$$

where

$$\mathcal{E}(\rho|\mu) := \int_\Omega F(\rho(x))\mathrm{d}x + \int_\Omega V(x)\rho(x)\mathrm{d}x + \int_{\Omega\times\Omega} W(x,y)\mu(x)\rho(y)\mathrm{d}x\mathrm{d}y. \tag{4.42}$$

The convergence of this scheme to the solution of equation (4.40) has been proved previously. For each time step, we then have to solve exactly the same type of problems as in section 3.3.1 except that the potential has to be updated at each step. In figure 4.1, we see final sate starting from a uniform measure with interaction potential of the form $W(x) = |x|^a - |x|^b$.

4.8.2 Systems

Let us take two species for the sake of simplicity and consider the evolution of the densities of these two species, coupled only through interaction terms:

$$\partial_t\rho_1 = \mathrm{div}(\rho_1(\nabla F_1'(\rho_1) + \nabla U_1[\rho_1,\rho_2])),\ \partial_t\rho_2 = \mathrm{div}(\rho_2(\nabla F_2'(\rho_2) + \nabla U_2[\rho_1,\rho_2])) \tag{4.43}$$

with energies F_1 and F_2 corresponding to independent linear or non linear diffusion terms and the coupling drift terms given (as in Di Francesco and Fagioli [54]) by convolutions with smooth kernels and individual potentials V_1 and V_2:

$$U_1[\rho_1,\rho_2] = V_1 + W_{11}\star\rho_1 + W_{21}\star\rho_2,\ U_2[\rho_1,\rho_2] = V_2 + W_{12}\star\rho_1 + W_{22}\star\rho_2.$$

The nondiffusive case where $F_1 = F_2 = 0$ was studied by Di Francesco and Fagioli [54] (see Zinsl [118] for extensions to more than two species). Note that 4.43 is not a gradient flow except in the

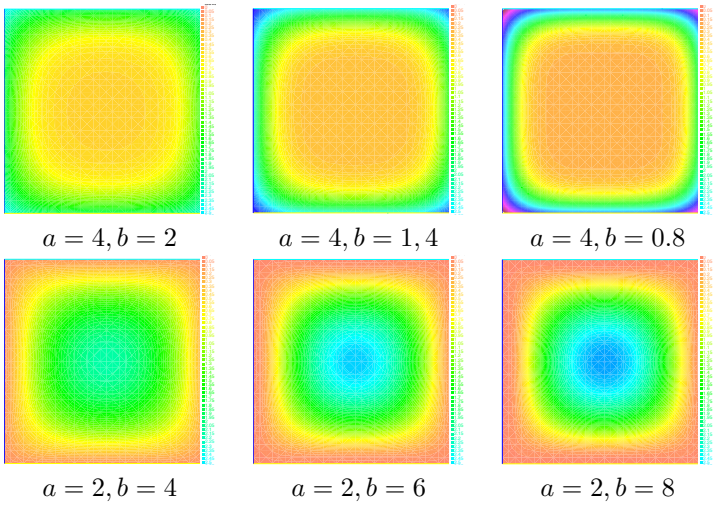


Figure 4.1: $\Omega = (-0.5, 0.5)^2$, *final state, starting from a uniform measure, with potential* $V = 0$, $F(\rho) = 0.01 * (\rho \log(\rho))$ *and interaction potential* $W(x) = |x|^a - |x|^b$.

particular case where interactions are symmetric i.e. $W_{12} = W_{21}$. The semi-implicit JKO scheme first proposed by Di Francesco and Fagioli in the nondiffusive case consists in defining inductively ρ_1^k and ρ_2^k by

$$\rho_1^{k+1} = \operatorname*{argmin}_{\rho} \frac{1}{2\tau} W_2^2(\rho, \rho_1^k) + \int_\Omega F_1(\rho(x))dx + \int_\Omega U_1[\rho_1^k, \rho_2^k](x)\rho(x)dx \tag{4.44}$$

and

$$\rho_2^{k+1} = \operatorname*{argmin}_{\rho} \frac{1}{2\tau} W_2^2(\rho, \rho_2^k) + \int_\Omega F_2(\rho(x))dx + \int_\Omega U_2[\rho_1^k, \rho_2^k](x)\rho(x)dx. \tag{4.45}$$

The convergence of this scheme is established in Di Francesco and Fagioli [54] in the nondiffusive case and in theorem 4.3 and theorem 4.17 (see [73]) in the diffusive case. Clearly, each independent subproblem 4.44 and 4.45 can be solved by the ALG2-JKO scheme described in chapter 3 and the extension to more than two species is direct as shown in the simulations below.

In figure 4.2, we see the evolution of two species which solve (4.43) with $F_1(\rho) = F_2(\rho) = \frac{1}{2}\rho^3$, $V_1 = V_2 = 0$, $W_{11}(x) = W_{21}(x) = W_{22}(x) = \frac{|x|^2}{2}$ and $W_{12}(x) = -\frac{|x|^2}{2}$. In other words, the first species is attracted by the second one but the second species is repelled by the first one. Since we have attractive self-interactions, the two species do not spread too much.

This scheme can treat systems with more than two species. In figure 4.3, we represent the evolution of three species which run after each other and with linear diffusion. More precisely, the interaction potentials we use in this example are of the form

$$U_i[\rho_1, \rho_2, \rho_3] = |x|^2 * \rho_{i+1} - |x|^2 * \rho_{i-1},$$

where $\rho_4 := \rho_1$ and $\rho_0 := \rho_3$.

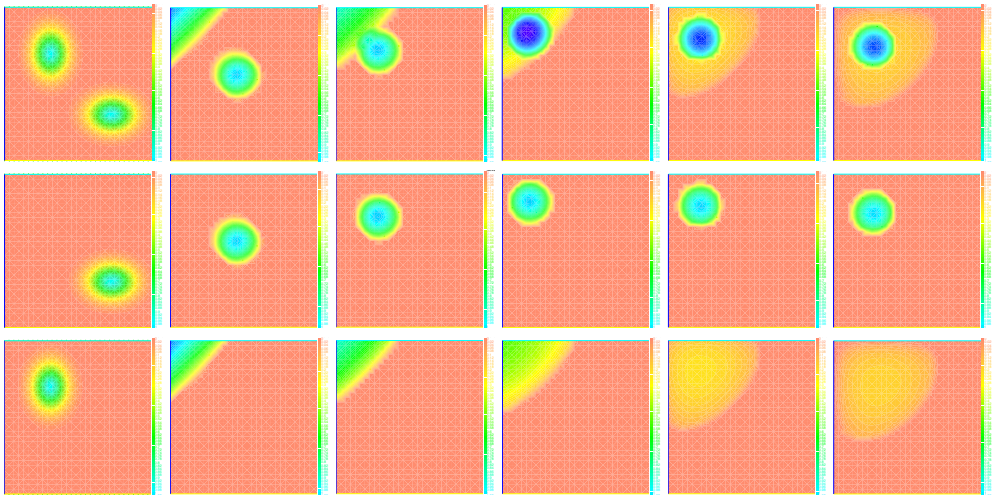


Figure 4.2: *Evolution of two species where the first one is attracted by the other and the second one is repelled by the first one. Top row: display of* $\rho_1 + \rho_2$*. Middle row: display of* ρ_1*. Bottom row: display of* ρ_2*.*

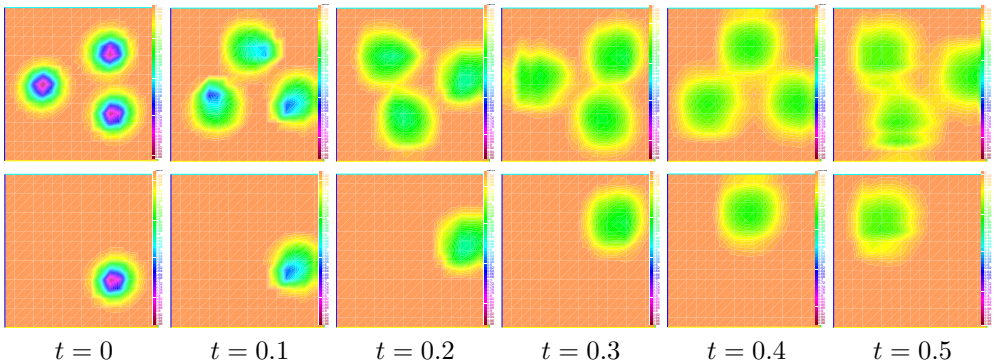


Figure 4.3: *Evolution of three species running after each other with linear diffusion. Top row: display of* $\rho_1 + \rho_2 + \rho_3$*. Bottom row: display of* ρ_1*.*

Chapter 5

Drift interactions: non potential case

The continuity equation with a density-dependent drift

$$\partial_t \rho = \mathrm{div}(\rho v), \quad \text{with } v = V[\rho]$$

is ubiquitous in modeling and arises in a variety of domains such as biology, particle physics, population dynamics, crowd modelling, opinion formation... It should actually come as no surprise since it captures the dynamics of a population of particles following the ODE $\dot{X} = -v(t, X)$ where $v = V[\rho]$ depends itself on the density in a way (local, nonlocal, attractive, repulsive etc..) depending on which phenomena (aggregation, diffusion...) one aims to capture and the type of applications.

Of course, at this level of generality not much can be said on existence and uniqueness. There are however two cases which may be treated in a rather systematic way. The first one, is the *regular case* where $V[\rho]$ is a smooth vector field whatever the probability measure ρ is, with some uniform bounds on some of its derivatives and $\rho \mapsto V[\rho]$ is Lipschitz in the Wasserstein metric. In this regular case, existence and uniqueness can be proved by the method of characteristics and suitable fixed point arguments (see [47] as well as [46], [45] for a different approach with applications to crowd dynamics). This regular case (a typical example being that of a convolution) is however rather restrictive and for instance rules out diffusion. The second case where there is a general theory is the *Wasserstein gradient flow* case. In this case, at least at a formal level, v may be written as $v = \nabla \frac{\delta \mathcal{E}}{\delta \rho}$ that is the gradient of the first variation of a functional $\mathcal{E}$ defined on measures. In their seminal paper [65], Jordan, Kinderlehrer and Otto discovered that the heat flow is the gradient flow of the entropy functional $\mathcal{E}(\rho) = \int \rho \log(\rho)$ which corresponds to the case $v = \frac{\nabla \rho}{\rho}$. The theory of Wasserstein gradient flows has been very succesful in addressing a variety of nonlinear evolution equations such as the porous medium equation [98], aggregation equations [37] or granular media equations [42]. This powerful theory is presented in a complete and detailed way in the reference book of Ambrosio, Gigli and Savaré [4].

The purpose of this chapter is a contribution to the following general question: can one hope for an existence/uniqueness theory in the case where V is the sum of a Wasserstein gradient flow term and a regular term (not necessarily a gradient). Our motivation for this question actually comes from systems. For instance, a simple but natural model, for the evolution of two (say) interacting species is:

$$\partial_t \rho_1 = \nu_1 \Delta \rho_1 + \mathrm{div}(\rho_1 (U_{11} \star \rho_1 + U_{12} \star \rho_2)), \tag{5.1}$$

$$\partial_t \rho_2 = \nu_2 \Delta \rho_2 + \mathrm{div}(\rho_2 (U_{21} \star \rho_1 + U_{22} \star \rho_2)). \tag{5.2}$$

If the vector-fields U_{ij} have potentials i.e. $U_{ij} = \nabla V_{ij}$ and there is no diffusion i.e. when $\nu_1 = \nu_2 = 0$, this is exactly the system studied by Di Francesco and Fagioli [54]. As emphasized in [54], if cross-interactions are symmetric i.e. $V_{12} = V_{21}$ (or more generally V_{12} and V_{21} are proportional), this system has a (product) Wasserstein gradient flow structure but this is certainly a restrictive and often unrealistic assumption in applications. This is why Di Francesco and Fagioli, still taking advantage of the similarity with Wasserstein gradient flows used a semi-implicit scheme à la Jordan-Kinderlehrer-Otto to obtain existence and uniqueness results. In [54], there is no diffusion and we

refer to chapter 4 (or [73]) for an extension of the semi-implicit method to the diffusive case. Clearly the structure of the system (5.1)-(5.2) belongs to the *mixed* case where drifts can be decomposed as the sum of a Wasserstein gradient and a regular term. It is worth noting though that the semi-implicit scheme of [54] cannot cover the case of general (not gradients) vector-fields U_{ij}.

In this chapter, we propose two different methods to overcome this difficulty. The first one is completly parabolic. For general drifts (not gradients) and possibly with nonlinear diffusion, we develop some simple arguments in the sequel which give an alternative to mass transport arguments and that enable us to address regular drifts without a gradient structure. Regularizing the diffusive term, we first prove the result on the approximated equation using a fixed point argument and then we find uniform bound with standard parabolic estimate allowing us to pass to the limit on the parameter. The advantage of this method is to use simple argument from parabolic theory but it is not constructive.

On the other hand, the second section present a splitting transport-JKO scheme to study nonlinear diffusion equations (or more generally, systems) with a general density-dependent drift:

$$\partial_t \rho(t,x) - \Delta P(\rho(t,x)) - \mathrm{div}(\rho(t,x) U[\rho(t,.)](x)) = 0,\ t \geq 0,\ x \in \Omega,$$
$$(\nabla P(\rho) + U[\rho]) \cdot \nu = 0 \text{ on } \partial\Omega,\ \rho_{|t=0} = \rho_0,$$

where for every probability density ρ, $U[\rho]$ is a-not necessarily potential- vector field, for instance, in even dimensions, it can mix a gradient and Hamiltonian structures i.e. be of the form $\nabla V[\rho] + J\nabla H[\rho]$ (where J is the usual symplectic matrix). The potential case where $U[\rho] = \nabla V[\rho]$ can be studied by means of a semi-implicit JKO scheme introduced by Di Francesco and Fagioli [54] in the nondiffusive case and further developed in chapter 4 (or [73]) for the case of a non linear diffusion. The idea of our splitting scheme is natural and consists in performing a Helmholtz decomposition of $U[\rho]$. We then treat the divergence-free part purely by (continuous in time) transport and the potential part by the semi-implicit JKO scheme. For the transport steps of the splitting scheme, we essentially need the divergence-free part to have some Sobolev regularity in x so as to be able to apply DiPerna-Lions theory, we will also need both the potential and divergence-free part of $\rho \mapsto U[\rho]$ to satisfy some Lipschitz continuity condition with respect to the Wasserstein distance. One advantage of the constructive splitting method presented here is that the transport steps by a divergence-free vector field preserve the internal energy, this is one way to overcome some difficulties discussed in [94] (section 5, variant 3). Moreover, we extend the method to obtain the strong convergence developped in the bounded case in the previous chapter 4 to handle the unbounded case which allows us to work in -possibly non convex- unbounded domain.

Another advantage of the splitting method is the extention on more general degenerate diffusion present in section 5.3. In this section, the drift is handle by a pure transport step (we do not split the drift using the Helmotz decomposition) and the diffusion comes from a JKO step replacing the usual 2-Wasserstein distance by a general distance

$$\mathcal{W}_c(\rho,\mu) := \inf \left\{ \iint_{\Omega\times\Omega} c(x,y)\, d\gamma(x,y) \ :\ \gamma \in \Pi(\rho,\mu) \right\}.$$

Unfortunately, during the pure transport phase, we do not have mass conservation anymore and then we have to use a maximum/minimum principle to recover usual estimates in the gradient flow theory.

5.1 Parabolic regularization method

This section is devoted to existence and uniqueness results for some classes of nonlinear diffusion equations in the presence of a *regular* drift term. These equations may be viewed as regular perturbations of Wasserstein gradient flows but the drift terms are not necessarily gradients (which makes it difficult to use Wasserstein gradient flows techniques). We obtain existence by a regularization procedure and parabolic energy estimates and address the uniqueness issue by an elementary H^{-1} contraction argument if the diffusion is nondegenerate. Our arguments directly extend to systems with diagonal nonlinear diffusions which are coupled through regular drifts.

This section is based on a joint work with G. Carlier, see [35], and is organized as follows. In section 5.1.1, we give an existence result by a suitable regularization of the diffusion and energy estimates. In section 5.1.2, we give an elementary H^{-1}-contraction argument from which uniqueness follows. Finally, we observe that these arguments easily extend to the case of systems (without cross-diffusion) in section 5.1.3.

5.1.1 Existence

For the sake of simplicity, we shall work in the periodic in space case so as to avoid boundary issues and thus take the space variable in the flat torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. We then consider the following nonlinear diffusion equation with a nonlocal drift:

$$\partial_t \rho - \Delta(F'(\rho)) + \operatorname{div}(\rho U[\rho]) = 0,\ \rho|_{t=0} = \rho_0, \tag{5.3}$$

on $(0,T)\times\mathbb{T}^d$. Denoting by $H^{-1}(\mathbb{T}^d)$ the dual of $H^1(\mathbb{T}^d)$ and by $\mathcal{P}(\mathbb{T}^d)$ the set of probability measures on $\mathbb{T}^d$, we assume the following regularity on the drift term $U[\rho]$:

$$\forall \rho \in L^2 \cap \mathcal{P}(\mathbb{T}^d), U[\rho] \in L^\infty(\mathbb{T}^d) \text{ and } \operatorname{div}(U[\rho]) \in L^\infty(\mathbb{T}^d)$$

with

$$\sup_{\rho\in L^2\cap\mathcal{P}(\mathbb{T}^d)} \{\|U[\rho]\|_{L^\infty} + \|\operatorname{div}(U[\rho])\|_{L^\infty}\} < +\infty \tag{5.4}$$

and for every $R>0$, there exists a modulus ω_R such that, for every $(\rho,\eta)\in (L^2(\mathbb{T}^d)\cap\mathcal{P}(\mathbb{T}^d))^2$ such that $\|\rho\|_{H^{-1}(\mathbb{T}^d)} \le R$ and $\|\eta\|_{H^{-1}(\mathbb{T}^d)} \le R$, one has

$$\|U[\rho]-U[\eta]\|_{L^2(\mathbb{T}^d)} \le \omega_R(\|\rho-\eta\|_{H^{-1}(\mathbb{T}^d)}). \tag{5.5}$$

Examples: Typical examples of velocity fields $\rho \mapsto U[\rho]$ that satisfy the above assumptions (5.4)-(5.5) are those of the form $U[\rho](x) = \int_{\mathbb{T}^d} B(x,y)\rho(y)\mathrm{d}y$, in which case (5.4) is satisfied as soon as B and $\operatorname{div}_x B$ are bounded. As for (5.5), it holds (with a linear modulus) as soon as

$$\int_{\mathbb{T}^d\times\mathbb{T}^d} |D_y B(x,y)|^2 \mathrm{d}x\mathrm{d}y < +\infty$$

since in this case

$$\|U[\rho]-U[\eta]\|_{L^2(\mathbb{T}^d)} \le \|\rho-\eta\|_{H^{-1}(\mathbb{T}^d)} \Big(\int_{\mathbb{T}^d\times\mathbb{T}^d} |B|^2 + |D_y B|^2\Big)^{\frac12}.$$

One can also consider a velocity of the form $U[\rho](x) = \int_{\mathbb{T}^d\times\mathbb{T}^d} B(x,y,z)\rho(y)\mathrm{d}y\rho(z)\mathrm{d}z$, if $B \in L^\infty(\mathbb{T}^d\times\mathbb{T}^d\times\mathbb{T}^d)$, $\operatorname{div}_x B \in L^\infty(\mathbb{T}^d\times\mathbb{T}^d\times\mathbb{T}^d)$ then (5.4) obviously holds; if, in addition the constant

$$C_B^2 := \int_{\mathbb{T}^d\times\mathbb{T}^d\times\mathbb{T}^d} (|B|^2 + |D_y B(x,y,z)|^2 + |D_z B(x,y,z)|^2 + |D^2_{yz} B(x,y,z)|^2)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

is finite then (5.5) follows from

$$\|U[\rho]-U[\eta]\|_{L^2(\mathbb{T}^d)} \le \|\rho-\eta\|_{H^{-1}(\mathbb{T}^d)}(\|\rho\|_{H^{-1}(\mathbb{T}^d)} + \|\eta\|_{H^{-1}(\mathbb{T}^d)})C_B.$$

Concerning the diffusion term, we make the following assumptions on F (which are satisfied whenever $F'(\rho)=\rho^m$ with $m\ge 1$):

$$F \in C^2(\mathbb{R}^+,\mathbb{R}), F(0) = F'(0) = 0,\ F \text{ is convex}, \tag{5.6}$$

$$F'' \text{ is nondecreasing, and for every } \rho>0,\ F''(\rho)>0 \tag{5.7}$$

and there is a constant $C>0$ such that

$$F'(\rho) \le C(1+\rho^2+F(\rho)),\ \forall \rho\in\mathbb{R}^+. \tag{5.8}$$

Finally, for the initial condition ρ_0 we assume that it is a probability density such that

$$\rho_0 \in L^2(\mathbb{T}^d),\ F(\rho_0) \in L^1(\mathbb{T}^d). \tag{5.9}$$

A nonnegative weak solution of the PDE

$$\partial_t \rho - \Delta(F'(\rho)) + \mathrm{div}(\rho U[\rho]) = 0,\ \rho|_{t=0} = \rho_0. \tag{5.10}$$

is by definition a function $\rho \in L^2((0,T)\times\mathbb{T}^d, \mathbb{R}^+)$ such that

$$F'(\rho) \in L^2((0,T), H^1(\mathbb{T}^d))$$

and

$$\int_0^T \int_{\mathbb{T}^d} (-\partial_t \phi \rho + \nabla F'(\rho)\cdot\nabla\phi - \rho U[\rho]\cdot\nabla\phi)\,\mathrm{d}x\mathrm{d}t = \int_{\mathbb{T}^d} \phi(0,x)\rho_0(x)\mathrm{d}x \tag{5.11}$$

for every $\phi \in C^1([0,T]\times\mathbb{T}^d)$ such that $\phi(T,.) = 0$.

Before we proceed to the existence proof, we need some preliminary results. Let us first study the continuity of the drift term $\rho = \rho(t,x) \mapsto U[\rho(t,.)](x)$. It is easy to see that when (5.4) and (5.5) are satisfied and ρ^n converges strongly in $L^2((0,T)\times\mathbb{T}^d)$ (hence in $L^2((0,T), H^{-1}(\mathbb{T}^d))$) to some ρ then $U[\rho^n]$ converges to $U[\rho]$ in $L^2((0,T)\times\mathbb{T}^d)$, but we wil need a variant in the sequel:

Lemma 5.1. *Assume that* (5.4) *and* (5.5) *are satisfied. Let ρ^n be a sequence in $L^2((0,T)\times\mathbb{T}^d)$ such that $\partial_t\rho^n \in L^2((0,T), H^{-1}(\mathbb{T}^d))$ with*

$$\sup_n \|\partial_t \rho^n\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} < +\infty, \tag{5.12}$$

and $\rho \in L^2((0,T)\times\mathbb{T}^d)$ such that $\rho^n \rightharpoonup \rho$ in $L^2((0,T)\times\mathbb{T}^d)$, then $U[\rho^n]$ converges to $U[\rho]$ strongly in $L^2((0,T)\times\mathbb{T}^d)$.

Proof. First observe that (5.12) implies that for some constant C one has

$$\|\rho^n(t,.) - \rho^n(s,.)\|_{H^{-1}} \le C\sqrt{|t-s|},\ \forall n,\ \forall (t,s)\in(0,T)^2. \tag{5.13}$$

Let $t \in (0,T)$ and for $h \in (0,t)$ define

$$\overline{\rho}^n_{t,h}(x) := \frac{1}{h}\int_{t-h}^t \rho^n(s,x)ds,\ \overline{\rho}_{t,h} := \frac{1}{h}\int_{t-h}^t \rho(s,x)ds$$

thanks to (5.13), we obtain, for every n, t and h:

$$\|\rho^n(t,.) - \overline{\rho}^n_{t,h}\|_{H^{-1}} \le C\sqrt{h},\ \|\rho(t,.) - \overline{\rho}_{t,h}\|_{H^{-1}} \le C\sqrt{h}. \tag{5.14}$$

For fixed $h>0$, $\overline{\rho}^n_{t,h} \rightharpoonup \overline{\rho}_{t,h}$ in $L^2(\mathbb{T}^d)$ as $n\to\infty$, and since the imbedding of $L^2(\mathbb{T}^d)$ into $H^{-1}(\mathbb{T}^d)$ is compact we also have $\|\overline{\rho}^n_{t,h} - \overline{\rho}_{t,h}\|_{H^{-1}(\mathbb{T}^d)} \to 0$ as $n\to\infty$. We then get

$$\|\rho^n(t,.) - \rho(t,.)\|_{H^{-1}} \le 2C\sqrt{h} + \|\overline{\rho}^n_{t,h} - \overline{\rho}_{t,h}\|_{H^{-1}(\mathbb{T}^d)}$$

from which we deduce that $\|\rho^n(t,.) - \rho(t,.)\|_{H^{-1}}$ tends to 0. Thanks to (5.5), this implies that $\|U[\rho^n(t,.)] - U[\rho(t,.)]\|_{L^2(\mathbb{T}^d)}$ tends to 0. The claimed L^2 convergence then follows from (5.4) and Lebesgue's dominated convergence Theorem.

□

We now introduce a regularized nonlinearity to approximate (5.10) by a uniformly parabolic equation as follows. Let $\varepsilon \in (0,1)$, let δ_ε and M_ε be respectively the smallest ρ for which $F''(\rho) \ge \varepsilon$ and the largest ρ for which $F''(\rho) \le \varepsilon^{-1}$. Let then F_ε be defined by

$$F_\varepsilon(\rho) := \begin{cases} F(\delta_\varepsilon) + F'(\delta_\varepsilon)(\rho-\delta_\varepsilon) + \frac{\varepsilon}{2}(\rho-\delta_\varepsilon)^2 \text{ if } \rho\in[0,\delta_\varepsilon]; \\ F(\rho) \text{ if } \rho\in[\delta_\varepsilon, M_\varepsilon], \\ F(M_\varepsilon) + F'(M_\varepsilon)(\rho-M_\varepsilon) + \frac{1}{2\varepsilon}(\rho-M_\varepsilon)^2 \text{ if } \rho \ge M_\varepsilon. \end{cases} \tag{5.15}$$

Clearly, by construction F_ε is convex and C^2 on $\mathbb{R}_+$ with

$$\varepsilon \leq F_\varepsilon'' \leq \frac{1}{\varepsilon} \text{ on } \mathbb{R}^+ \tag{5.16}$$

and F_ε converges pointwise to F since δ_ε and M_ε converge respectively to 0 and $+\infty$. In fact, this approximation also has good Γ-convergence properties:

Lemma 5.2. *Let $\theta \in L^2((0,T)\times\mathbb{T}^d,\mathbb{R}^+)$, then*

$$\lim_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))dxdt = \int_0^T\int_{\mathbb{T}^d} F(\theta(t,x))dxdt \tag{5.17}$$

moreover if $\theta_\varepsilon \in L^2((0,T)\times\mathbb{T}^d),\mathbb{R}^+)$ weakly converges to θ in $\in L^2((0,T)\times\mathbb{T}^d)$, then

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta_\varepsilon(t,x))dxdt \geq \int_0^T\int_{\mathbb{T}^d} F(\theta(t,x))dxdt \tag{5.18}$$

Proof. Fatou's lemma first yields

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))\mathrm{d}x\mathrm{d}t \geq \int_0^T\int_{\mathbb{T}^d} F(\theta(t,x))\mathrm{d}x\mathrm{d}t$$

on the other hand

$$\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))\mathrm{d}x\mathrm{d}t \leq \int_0^T\int_{\mathbb{T}^d} F(\theta(t,x))\mathrm{d}x\mathrm{d}t + \int\int_{\{\theta\leq\delta_\varepsilon\}}(F_\varepsilon(\theta)-F(\theta))\mathrm{d}x\mathrm{d}t$$

since the second term in the right hand side converges to 0, we easily deduce (5.17). Let us now assume that $\theta_\varepsilon \in L^2((0,T)\times\mathbb{T}^d,\mathbb{R}^+)$ weakly converges to θ in $\in L^2((0,T)\times\mathbb{T}^d)$. Let $\gamma>0$ (fixed for the moment) and denote by F^γ the function defined by

$$F^\gamma(\rho) = \begin{cases} F(\rho) \text{ if } \rho\in[0,\gamma], \\ F(\gamma)+F'(\gamma)(\rho-\gamma) \text{ if } \rho\geq\gamma \end{cases}$$

by construction F^γ is convex and below F. For $\varepsilon>0$ small enough so that $\gamma\in[\delta_\varepsilon,M_\varepsilon]$, we similarly define

$$F_\varepsilon^\gamma(\rho) = \begin{cases} F_\varepsilon(\rho) \text{ if } \rho\in[0,\gamma], \\ F(\gamma)+F'(\gamma)(\rho-\gamma) \text{ if } \rho\geq\gamma \end{cases}$$

so that F_ε^γ is convex and coincides with F^γ on $[\delta_\varepsilon,+\infty)$. We then have

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t \geq \liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon^\gamma(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t$$
$$\geq \liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F^\gamma(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t + \liminf_{\varepsilon\to 0^+}\int\int_{\{\theta_\varepsilon\leq\delta_\varepsilon\}}(F_\varepsilon(\theta_\varepsilon)-F(\theta_\varepsilon))$$

the second term converges to 0 whereas by weak lower semi-continuity (thanks to the convexity of F^γ) we have

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F^\gamma(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t \geq \int_0^T\int_{\mathbb{T}^d} F^\gamma(\theta(t,x))\mathrm{d}x\mathrm{d}t,$$

hence

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t \geq \sup_{\gamma>0}\int_0^T\int_{\mathbb{T}^d} F^\gamma(\theta(t,x))\mathrm{d}x\mathrm{d}t$$

and then (5.18) easily follows from the previous inequality, the fact that F^γ converges monotonically to F and Beppo-Levi's monotone convergence Theorem. □

Theorem 5.3. *Assume* (5.4)-(5.5)-(5.6)-(5.7)-(5.8)-(5.9), *then* (5.10) *admits at least one weak nonnegative solution.*

Proof. The proof proceeds in three steps.

Step 1: Regularized equation. We first prove existence of a weak solution to the regularized equation:

$$\partial_t \rho^\varepsilon - \Delta(F_\varepsilon'(\rho^\varepsilon)) + \mathrm{div}(\rho^\varepsilon U[\rho^\varepsilon]) = 0,\ \rho^\varepsilon|_{t=0} = \rho_0. \tag{5.19}$$

Let

$$X := \{\eta \in L^2((0,T)\times\mathbb{T}^d, \mathbb{R}^+)\ :\ \int_{\mathbb{T}^d} \eta(t,x)\mathrm{d}x = 1 \text{ for a.e. } t \in (0,T)\}$$

for fixed $\varepsilon > 0$ and $\eta \in X$, consider the linear parabolic equation in divergence form:

$$\partial_t u - \mathrm{div}(F_\varepsilon''(\eta)\nabla u) + \mathrm{div}(uU[\eta]) = 0,\ u_{t=0} = \rho_0 \tag{5.20}$$

which can be rewritten in nondivergence form as

$$\partial_t u - \mathrm{div}(a_\varepsilon(\eta)\nabla u) + b[\eta]\cdot\nabla u + c[\eta]u = 0 \tag{5.21}$$

where the coefficients $a_\varepsilon(\eta) := F_\varepsilon''(\eta)$, $b[\eta] := U[\eta]$ and $c[\eta] := \mathrm{div}(U[\eta])$ all belong to $L^\infty((0,T)\times\mathbb{T}^d)$ with $a_\varepsilon(\eta) \geq \varepsilon$ by (5.16). It follows from standard linear parabolic theory (see e.g. [82]) that (5.21) admits a unique weak solution which we denote $u := T^\varepsilon(\eta) \in L^2((0,T), H^1(\mathbb{T}^d)) \cap C([0,T], L^2(\mathbb{T}^d))$ with $\partial_t u \in L^2((0,T), H^{-1}(\mathbb{T}^d))$. Obviously, $u(t,.)$ is a probability density for every $t \in [0,T]$: it remains nonnegative by the maximum principle and its integral over $\mathbb{T}^d$ is constant in time, in other words $T^\varepsilon(X) \subset X$. Moreover, multiplying (5.21) by u thanks to (5.4) and (5.16) there is a constant C_ε (independent of η) such that $u := T^\varepsilon(\eta)$ satisfies

$$\int_0^T \int_{\mathbb{T}^d} (|\nabla u|^2 + u^2)\mathrm{d}x\mathrm{d}t + \int_0^T \|\partial_t u\|_{H^{-1}}^2 \leq C_\varepsilon. \tag{5.22}$$

The bound (5.22) and the Aubin-Lions lemma (see [5], [114]) thus imply that $T^\varepsilon(X)$ is relatively compact in $L^2((0,T)\times\mathbb{T}^d)$. Thanks to (5.5), the continuity of F_ε'' and Lebesgue's dominated convergence theorem, it is easy to check that T^ε is continuous with respect to the $L^2((0,T)\times\mathbb{T}^d)$ norm. Schauder's fixed-point Theorem then ensures that T^ε admits at least one fixed point i.e. a solution of (5.19) which we from now denote ρ^ε.

Step 2: A priori estimates. We aim now to derive estimates independent of ε on ρ^ε. Let $\delta > 0$ such that $\delta \in (\delta_\varepsilon, M_\varepsilon)$, we then take $(\rho^\varepsilon - \delta)_+$ as test-function in (5.19) (which is actually licit since this test-function belongs to $L^2((0,T), H^1(\mathbb{T}^d))$) integrating between 0 and $t \in [0,T]$ this yields

$$\int_0^t \langle \partial_t \rho^\varepsilon, (\rho^\varepsilon - \delta)_+\rangle_{H^{-1},H^1} ds + \int_0^t \int_{\{\rho^\varepsilon \geq \delta\}} F_\varepsilon''(\rho^\varepsilon)|\nabla\rho^\varepsilon|^2 = \int_0^t \int_{\{\rho^\varepsilon\geq\delta\}} \rho^\varepsilon U[\rho^\varepsilon]\cdot\nabla\rho^\varepsilon$$

hence, using Young's inequality, for every $\mu > 0$

$$\begin{aligned}\frac{1}{2}\|(\rho^\varepsilon(t,.)-\delta)_+\|_{L^2}^2 - \frac{1}{2}\|(\rho_0-\delta)_+\|_{L^2}^2 + \int_0^t\int_{\{\rho^\varepsilon\geq\delta\}} F_\varepsilon''(\rho^\varepsilon)|\nabla\rho^\varepsilon|^2 \\ \leq C\Big(\frac{\mu}{2}\int_0^t\int_{\{\rho^\varepsilon\geq\delta\}}|\nabla\rho^\varepsilon|^2 + \frac{1}{2\mu}\int_0^t\int_{\{\rho^\varepsilon\geq\delta\}}(\rho^\varepsilon)^2\Big) \\ \leq C\Big(\frac{\mu}{2}\int_0^t\int_{\{\rho^\varepsilon\geq\delta\}}|\nabla\rho^\varepsilon|^2 + \frac{1}{\mu}\int_0^t\int_{\{\rho^\varepsilon\geq\delta\}}[(\rho^\varepsilon-\delta)_+^2 + \delta^2]\Big)\end{aligned}$$

since $F_\varepsilon''(\delta) = F''(\delta) > 0$ and F'' nondecreasing, we can choose μ small enough so that the first term in the right hand side is absorbed by the left hand side of the inequality. Gronwall's lemma then gives

$$\sup_{t\in(0,T)} \|\rho^\varepsilon(t,.)\|_{L^2} \leq C \tag{5.23}$$

for a constant C that does not depend on ε. Next we take $F'_\varepsilon(\rho^\varepsilon)$ as test-function which similarly gives:

$$\int_{\mathbb{T}^d} F_\varepsilon(\rho^\varepsilon(t,.)) - \int_{\mathbb{T}^d} F_\varepsilon(\rho_0) + \int_0^t \int_{\mathbb{T}^d} |\nabla F'_\varepsilon(\rho^\varepsilon)|^2 = \int_0^t \int_{\mathbb{T}^d} \rho^\varepsilon U[\rho^\varepsilon] \cdot \nabla F'_\varepsilon(\rho^\varepsilon)$$
$$\leq C\Big(\frac{\mu}{2} \int_0^t \int_{\mathbb{T}^d} |\nabla F'_\varepsilon(\rho^\varepsilon)|^2 + \frac{1}{2\mu} \int_0^t \int_{\mathbb{T}^d} (\rho^\varepsilon)^2\Big)$$

using (5.23) and chosing μ small enough we thus get

$$\sup_{t\in[0,T]} \int_{\mathbb{T}^d} F_\varepsilon(\rho_\varepsilon(t,.)) + \int_0^T \int_{\mathbb{T}^d} |\nabla F'_\varepsilon(\rho^\varepsilon)|^2 \leq C \tag{5.24}$$

for a constant C not depending on ε. Next we use (5.8) and (5.23)-(5.24) to deduce that

$$\sup_{t\in[0,T]} \int_{\mathbb{T}^d} F'_\varepsilon(\rho^\varepsilon) \leq C \tag{5.25}$$

together with Poincaré-Wirtinger inequality, using again (5.24), this gives

$$\|F'_\varepsilon(\rho^\varepsilon)\|_{L^2((0,T),H^1(\mathbb{T}^d))} \leq C. \tag{5.26}$$

Step 3: Passing to the limit. Let us set

$$u^\varepsilon := F'_\varepsilon(\rho^\varepsilon),\ \sigma^\varepsilon := \nabla u^\varepsilon - \rho^\varepsilon V[\rho^\varepsilon] \tag{5.27}$$

so that (5.19) can be rewritten as

$$\partial_t \rho^\varepsilon = \Delta u^\varepsilon - \mathrm{div}(\rho^\varepsilon U[\rho^\varepsilon]) = \mathrm{div}(\sigma^\varepsilon),\ \rho^\varepsilon|_{t=0} = \rho_0. \tag{5.28}$$

We know from the previous step that

$$\|\rho^\varepsilon\|_{L^\infty((0,T),L^2(\mathbb{T}^d))} + \|\sigma^\varepsilon\|_{L^2((0,T),L^2(\mathbb{T}^d))} + \|u^\varepsilon\|_{L^2((0,T),H^1(\mathbb{T}^d))} \leq C \tag{5.29}$$

as well as

$$\|\partial_t \rho^\varepsilon\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} \leq C. \tag{5.30}$$

Passing to subsequences if necessary, we may therefore assume that

$$\rho^\varepsilon \rightharpoonup \rho \text{ in } L^2((0,T)\times\mathbb{T}^d),\ u^\varepsilon \rightharpoonup u \text{ in } L^2((0,T),H^1(\mathbb{T}^d)) \tag{5.31}$$

and thanks to Lemma 5.1, (5.4) and (5.30), we have

$$\sigma^\varepsilon \rightharpoonup \sigma := \nabla u - \rho U[\rho] \text{ in } L^2((0,T)\times\mathbb{T}^d). \tag{5.32}$$

Obviously one then has:

$$\partial_t \rho = \mathrm{div}(\sigma) = \Delta u - \mathrm{div}(\rho U[\rho]),\ \rho|_{t=0} = \rho_0. \tag{5.33}$$

So to establish that ρ is a weak nonnegative solution of (5.10), it is enough to prove that $u = F'(\rho)$. Thanks to the convexity of F this amounts to prove that

$$\int_0^T \int_{\mathbb{T}^d} F(\theta(t,x))\mathrm{d}x\mathrm{d}t \geq \int_0^T \int_{\mathbb{T}^d} F(\rho(t,x))\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\mathbb{T}^d} u(\theta-\rho)\mathrm{d}x\mathrm{d}t \tag{5.34}$$

for every $\theta \in L^2((0,T)\times\mathbb{T}^d, \mathbb{R}^+)$. By definition of u^ε we know that

$$\int_0^T \int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))\mathrm{d}x\mathrm{d}t \geq \int_0^T \int_{\mathbb{T}^d} F_\varepsilon(\rho^\varepsilon(t,x))\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\mathbb{T}^d} u^\varepsilon(\theta-\rho^\varepsilon)\mathrm{d}x\mathrm{d}t. \tag{5.35}$$

Let us prove that

$$\lim_{\varepsilon} \int_0^T \int_{\mathbb{T}^d} u^\varepsilon \rho^\varepsilon = \int_0^T \int_{\mathbb{T}^d} u\rho. \tag{5.36}$$

For that purpose, let ψ^ε be the potential defined by

$$-\Delta\psi^\varepsilon = \rho^\varepsilon,\ \int_{\mathbb{T}^d} \psi^\varepsilon = 0,\ \psi^\varepsilon \in H^1(\mathbb{T}^d). \tag{5.37}$$

Thanks to (5.23), we have $\psi^\varepsilon \in L^\infty((0,T), H^2(\mathbb{T}^d))$ with a bound independendent of ε:

$$\|\nabla\psi^\varepsilon\|_{L^\infty((0,T),H^1(\mathbb{T}^d))} \leq C. \tag{5.38}$$

As for the time derivative of $\nabla\psi^\varepsilon$ we observe that

$$-\Delta(\partial_t\psi^\varepsilon) = \partial_t\rho^\varepsilon = \mathrm{div}(\sigma^\varepsilon)$$

so that, thanks to (5.29), we have $\partial_t\nabla\psi^\varepsilon \in L^2((0,T)\times\mathbb{T}^d)$ and more precisely

$$\|\partial_t\nabla\psi^\varepsilon\|_{L^2((0,T)\times\mathbb{T}^d)} \leq \|\sigma^\varepsilon\|_{L^2((0,T)\times\mathbb{T}^d)} \leq C$$

this proves that $\nabla\psi^\varepsilon$ is bounded in $H^1((0,T)\times\mathbb{T}^d)$, hence converges in $L^2((0,T)\times\mathbb{T}^d)$, up to an extraction if necessary, to ψ given by

$$-\Delta\psi = \rho,\ \int_{\mathbb{T}^d} \psi = 0,\ \psi \in H^1(\mathbb{T}^d). \tag{5.39}$$

Weak convergence of ∇u^ε and strong convergence of $\nabla\psi^\varepsilon$ in L^2 then give

$$\begin{aligned}\lim_{\varepsilon} \int_0^T \int_{\mathbb{T}^d} u^\varepsilon\rho^\varepsilon &= \lim_{\varepsilon} \int_0^T \int_{\mathbb{T}^d} \nabla u^\varepsilon \nabla\psi^\varepsilon \\ &= \int_0^T \int_{\mathbb{T}^d} \nabla u\nabla\psi = \int_0^T \int_{\mathbb{T}^d} u\rho\end{aligned}$$

which establishes (5.36). Next, we use Lemma 5.2, letting ε tend to 0^+, using (5.36) we obtain inequality (5.34) which proves that $u = F'(\rho)$ and so ρ is a weak solution of (5.10), concluding the proof.

□

5.1.2 H^{-1} contraction and uniqueness

We still assume that the diffusion F and the drift U satisfy (5.4)-(5.5)-(5.6)-(5.7)-(5.8). Given ρ_0 and η_0 in $L^2(\mathbb{T}^d)\cap\mathcal{P}(\mathbb{T}^d)$ such that $F(\rho_0)\in L^1(\mathbb{T}^d)$ and $F(\eta_0)\in L^1(\mathbb{T}^d)$, we have found in the previous section weak solutions ρ and η of the Cauchy problems:

$$\partial_t\rho - \Delta(F'(\rho)) + \mathrm{div}(\rho U[\rho]) = 0,\ \rho|_{t=0} = \rho_0, \tag{5.40}$$

and

$$\partial_t\eta - \Delta(F'(\eta)) + \mathrm{div}(\eta U[\eta]) = 0,\ \eta|_{t=0} = \eta_0 \tag{5.41}$$

such that (recalling (5.23) in the proof of Theorem 5.3) for some $R_0 > 0$

$$\|\rho\|_{L^\infty((0,T),L^2(\mathbb{T}^d))} \leq R_0,\ \|\eta\|_{L^\infty((0,T),L^2(\mathbb{T}^d))} \leq R_0, \tag{5.42}$$

as well as

$$\|\partial_t\rho\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} \leq R_0,\ \|\partial_t\eta\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} \leq R_0. \tag{5.43}$$

We then set $u := \rho - \eta$ (so that $u(t,.)$ has zero mean for a.e. $t\in(0,T)$ and $\partial_t u \in L^2((0,T),H^{-1}(\mathbb{T}^d))$ and our aim is to prove an H^{-1} contraction estimate, that is some decay estimate on

$$\sup\Big\{\int_{\mathbb{T}^d} u(t,.)\varphi,\ \int_{\mathbb{T}^d}|\nabla\varphi|^2 \leq 1,\ \int_{\mathbb{T}^d}\varphi = 0\Big\}$$

slightly abusing notations, we shall denote the previous quantity $\|u(t,.)\|_{H^{-1}}$ (actually since $u(t,.)$ has zero mean, this is indeed an equivalent norm) and thus

$$\|u(t,.)\|_{H^{-1}}^2 = \int_{\mathbb{T}^d} \varphi(t,x)u(t,x)\mathrm{d}x$$

where $\varphi(t,.)$ is the potential of $u(t,.)$ i.e.

$$-\Delta\varphi(t,.) = u(t,.),\ \int_{\mathbb{T}^d} \varphi(t,.) = 0. \tag{5.44}$$

For this contraction estimate, we need two additional assumptions. The first one is a strong ellipiticty condition, namely that there exists $\alpha > 0$ such that

$$(F'(s) - F'(t))(s-t) \geq \alpha(s-t)^2,\ \forall (s,t) \in \mathbb{R}^+ \times \mathbb{R}^+. \tag{5.45}$$

The second assumption is on the drift U and requires that for every $R > 0$, there is some $C_R > 0$ such that for all $(\rho,\eta) \in (L^2(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d))^2$ such that $\|\rho\|_{L^2(\mathbb{T}^d)} \leq R$ and $\|\eta\|_{L^2(\mathbb{T}^d)} \leq R$, one has

$$\|U[\rho] - U[\eta]\|_{L^\infty(\mathbb{T}^d)} \leq C_R\|\rho - \eta\|_{L^2(\mathbb{T}^d)}. \tag{5.46}$$

Note that (5.46) is satisfied for $U[\rho](x) = \int_{\mathbb{T}^d} B(x,y)\rho(y)\mathrm{d}y$ with $B \in L^\infty_x(L^2_y)$ or when $U[\rho](x) = \int_{\mathbb{T}^d} B(x,y,z)\rho(y)\mathrm{d}y\rho(z)\mathrm{d}z$ with $B \in L^\infty_{x,y}(L^2_z)$. Under these extra assumptions, we have the following contraction result:

Theorem 5.4. *Under the assumptions above there exists a constant $\lambda = \lambda(\alpha, R_0, C_{R_0}, \|U[\rho]\|_{L^\infty})$ such that*

$$\|\rho(t,.) - \eta(t,.)\|_{H^{-1}} \leq e^{\lambda t}\|\rho_0 - \eta_0\|_{H^{-1}},\ \forall t \in (0,T), \tag{5.47}$$

so that in particular there is uniqueness for (5.40).

Proof. Again we define $u(t,.) = \rho(t,.) - \eta(t,.)$ and $\varphi(t,.)$ by (5.44), setting $G(t,.) := \rho(t,.)U[\rho(t,.)] - \eta(t,.)U[\eta(t,.)]$, and defining the potential $H(t,.)$ by:

$$\Delta H(t,.) = \mathrm{div}(G(t,.)),\ \int_{\mathbb{T}^d} H(t,.) = 0, \tag{5.48}$$

we then have

$$\begin{aligned}\frac{d}{dt}\|u(t,.)\|_{H^{-1}}^2 &= 2\langle \partial_t u(t,.), \varphi(t,.)\rangle_{H^{-1},H^1} \\ &= 2\langle \Delta(F'(\rho(t,.)) - F'(\eta(t,.))) - \mathrm{div}(G(t,.)), \varphi(t,.)\rangle_{H^{-1},H^1} \\ &= -2\int_{\mathbb{T}^d} (F'(\rho(t,.)) - F'(\eta(t,.)))u(t,.) + 2\int_{\mathbb{T}^d} H(t,.)u(t,.)\end{aligned}$$

so that, thanks to (5.45), we have

$$\frac{d}{dt}\|u(t,.)\|_{H^{-1}}^2 \leq -2\alpha\|u(t,.)\|_{L^2}^2 + 2\|u(t,.)\|_{H^{-1}}\|H(t,.)\|_{H^1}. \tag{5.49}$$

But since $\|H(t,.)\|_{H^1} \leq \|G(t,.)\|_{L^2} \leq \|\rho(t,.) - \eta(t,.)\|_{L^2}\|U[\rho(t,.)]\|_{L^\infty} + \|\eta(t,.)\|_{L^2}\|U[\rho(t,.)] - U[\eta(t,.)]\|_{L^\infty}$, thanks to (5.4), (5.46) and (5.42), we arrive at

$$\|H(t,.)\|_{H^1} \leq C\|u(t,.)\|_{L^2},$$

together with (5.49) and Young's inequality, this gives

$$\frac{d}{dt}\|u(t,.)\|_{H^{-1}}^2 \leq -2\alpha\|u(t,.)\|_{L^2}^2 + 2C\|u(t,.)\|_{H^{-1}}\|u(t,.)\|_{L^2} \leq \frac{C^2}{2\alpha}\|u(t,.)\|_{H^{-1}}^2$$

from which (5.47) directly follows with $\lambda = \frac{C^2}{4\alpha}$.

□

So far, we have only considered probability distributions as initial conditions, and since the evolution equation conserves total mass, one may think that the previous argument enables one to compare two solutions with the same total mass only. In fact, one can also obtain stability, in a similar ways as above but between solutions ρ and η corresponding to nonnegative initial conditions ρ_0 and η_0 not necessarily with the same total mass. Indeed, in this case, set again $u(t,.) := \rho(t,.) - \eta(t,.)$ and define by $\underline{u}$ its integral (note that this is constant in time) as well as the zero-mean function $v = u - \underline{u}$ and its potential φ by $-\Delta\varphi = v$, then computing, as before, the time derivative of $\|v\|^2_{H^{-1}} = \|\nabla\varphi\|_{L^2}$, we just have an extra term to take care of, namely $\underline{u}\int_{\mathbb{T}^d}(F'(\rho(t,.) - F'(\eta(t,.))$ but this term typically can be bounded by a constant times $|\underline{u}|$, thanks to the uniform in time bounds on $\|\rho(t,.)\|_{L^2}$, $\|\eta(t,.)\|_{L^2}$, the energies $\int_{\mathbb{T}^d} F(\rho(t,.))$ and $\int_{\mathbb{T}^d} F(\eta(t,.))$ and inequality (5.8). Doing so, one arrives easily at

$$\frac{d}{dt}\|v(t,.)\|^2_{H^{-1}} \leq C\|v(t,.)\|^2_{H^{-1}} + C|\underline{u}|$$

which, together with Grownwall's Lemma, gives the H^{-1} stability estimate

$$\|u(t,.) - \underline{u}\|^2_{H^{-1}} \leq \|u(0,.) - \underline{u}\|^2_{H^{-1}} e^{Ct} + |\underline{u}|(e^{Ct} - 1).$$

5.1.3 Extension to systems

The previous arguments clearly adapt to systems. More precisely, let us consider the system for the evolution of l densities $\rho := (\rho_1, \dots, \rho_l)$:

$$\partial_t \rho_i - \Delta(F_i'(\rho_i)) + \operatorname{div}(\rho_i U_i[\rho]) = 0,\ \rho_i|_{t=0} = \rho_{i,0} \tag{5.50}$$

on $(0, +\infty) \times \mathbb{T}^d$. Assuming that each function F_i satisfies (5.6)-(5.7)-(5.8), that the initial conditions are probability densities which satisfy

$$\rho_{i,0} \in L^2(\mathbb{T}^d),\ F_i(\rho_{i,0}) \in L^1(\mathbb{T}^d),\ \forall i = 1, \dots, l, \tag{5.51}$$

and for every $i = 1, \dots, l$, the map V_i satisfies

$$\forall \rho \in L^2(\mathbb{T}^d)^l \cap \mathcal{P}(\mathbb{T}^d)^l, U_i[\rho] \in L^\infty(\mathbb{T}^d) \text{ and } \operatorname{div}(U_i[\rho]) \in L^\infty(\mathbb{T}^d)$$

with

$$\sup_{\rho \in L^2(\mathbb{T}^d)^l \cap \mathcal{P}(\mathbb{T}^d)^l} \{\|U_i[\rho]\|_{L^\infty} + \|\operatorname{div}(U_i[\rho])\|_{L^\infty}\} < +\infty \tag{5.52}$$

and for every $R > 0$, there exists a modulus ω_R such that, for every $(\rho, \eta) \in L^2(\mathbb{T}^d)^l \times L^2(\mathbb{T}^d)^l$ such that $\|\rho\|_{H^{-1}(\mathbb{T}^d)^l} \leq R$ and $\|\eta\|_{H^{-1}(\mathbb{T}^d)^l} \leq R$, one has

$$\|V_i[\rho] - V_i[\eta]\|_{L^2(\mathbb{T}^d)} \leq \omega_R\Big(\sum_{j=1}^{l} \|\rho_j - \eta_j\|_{H^{-1}(\mathbb{T}^d)}\Big). \tag{5.53}$$

A direct adaptation of the proof of Theorem 5.3 gives

Theorem 5.5. *Assume that each function F_i satisfies* (5.6)-(5.7)-(5.8)*, and that* (5.51)-(5.52)-(5.53) *are satisfied for $i = 1, \dots, l$, then* (5.50) *admits at least one weak solution $(\rho_1, \dots, \rho_l)$ with each ρ_i nonnegative.*

As for uniqueness, the H^{-1} contraction argument of section 5.1.2 also easily adapts to systems of the form (5.50). Provided that there is an $\alpha > 0$ such that

$$(F_i'(s) - F_i'(t))(s - t) \geq \alpha(s - t)^2,\ \forall (s,t) \in \mathop{\mathbb{R}}_{+} \times \mathop{\mathbb{R}}_{+},\ \forall i = 1, \dots, l, \tag{5.54}$$

and, for every $R > 0$, there is some $C_R > 0$ such that for all $(\rho, \eta) \in (L^2(\mathbb{T}^d)^l \cap \mathcal{P}(\mathbb{T}^d)^l)^2$ such that $\|\rho\|_{L^2(\mathbb{T}^d)^l} \leq R$ and $\|\eta\|_{L^2(\mathbb{T}^d)^l} \leq R$, one has

$$\|U_i[\rho] - U_i[\eta]\|_{L^\infty(\mathbb{T}^d)} \leq C_R\|\rho - \eta\|_{L^2(\mathbb{T}^d)^l},\ \forall i = 1, \dots, l, \tag{5.55}$$

if $\rho = (\rho_1, \dots, \rho_l)$ and $\eta = (\eta_1, \dots \eta_l)$ both solve the system (5.50) then exactly as in the proof of Theorem 5.4 there is some λ such that

$$\|\rho(t,.) - \eta(t,.)\|_{H^{-1}} \leq e^{\lambda t}\|\rho(0,.) - \eta(0,.)\|_{H^{-1}},\ \forall t \in (0,T).$$

In particular, our results apply show well-posedness for systems like (5.1)-(5.2) presented in the introduction.

5.2 Splitting method in Wasserstein space

We prove an existence result for nonlinear diffusion equations in the presence of a nonlocal density-dependent drift which is not necessarily potential. The proof is constructive and based on the Helmholtz decomposition of the drift and a splitting scheme. The splitting scheme combines transport steps by the divergence-free part of the drift and semi-implicit minimization steps à la Jordan-Kinderlherer-Otto to deal with the potential part.

This section is based on a joint work with G. Carlier, see [33], and is organized as follows. Section 5.2.1 recalls some results from DiPerna-Lions theory. Section 5.2.2 lists the various assumptions, explains the splitting scheme and gives the main result. Section 5.2.3 gives estimates on the discrete sequences of measures obtained by the splitting scheme. Convergence of the scheme as the time step goes to 0 to a solution of the PDE is proved in section 5.2.4. In the concluding section 5.2.5, we briefly discuss extension to systems and uniqueness issues.

5.2.1 Flows of weakly differentiable vector fields

We will also need to apply the DiPerna Lions theory [57] in the special case of divergence-free vector fields. Let $W \in W^{1,1}_{\rm loc}(\mathbb{R}^n, \mathbb{R}^n)$ be divergence-free $\mathrm{div}(W) = 0$ and with at most linear growth

$$|W(x)| \leqslant C(1 + |x|).$$

Then there exists a unique flow map $X : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$, $X \in C(\mathbb{R}_+, L^1_{\rm loc}(\mathbb{R}^n))$ such that

- $X(0,.) = \mathrm{id}$ and for a.e. x, $t \in \mathbb{R}_+ \mapsto X(.,x)$ is a solution of the ODE $\dot{X} = W(X)$ i.e.:

$$X(t,x) = x + \int_0^t W(X(s,x))ds,\ t \geq 0,$$

- X satisfies the group property $X(t, X(s,x)) = X(t+s,x)$ for a.e. x and every $t, s \geq 0$,
- for every $X(t,.)$ preserves the n-dimensional Lebesgue measure.

Moreover given $\rho_0 \in \mathcal{P}_2^{\rm ac}(\mathbb{R}^n)$, $\rho(t,.) := X(t,.)_{\#}\rho_0 = \rho_0(X(t,.)^{-1})$ is the unique weak solution of the continuity equation

$$\partial_t \rho + \mathrm{div}(\rho W) = 0, \qquad \rho_{|t=0} = \rho_0, \tag{5.56}$$

which, since W is divergence-free, can also be rewritten as the transport equation $\partial_t\rho + \nabla\rho \cdot W = 0$. If we are given an open subset Ω of $\mathbb{R}^n$ with a smooth boundary and which is tangential (in the sense of traces) to $\partial\Omega$, then the DiPerna-Lions flow X leaves Ω invariant so that if $\rho_0 \in \mathcal{P}_2^{\rm ac}(\Omega)$ (extended by 0 outside Ω, say), the solution $\rho_t = X(t,.)_{\#}\rho_0$ of (5.56) remains supported in Ω hence may be viewed as a curve with values in $\mathcal{P}_2^{\rm ac}(\Omega)$.

5.2.2 Assumptions and main result

Given a suitable convex nonlinearity F and its associated pressure $P(\rho) = \rho F'(\rho) - F(\rho)$ as well as a nonlocal drift $\rho \mapsto U[\rho]$, our goal is to solve

$$\partial_t \rho - \Delta P(\rho) - \mathrm{div}(\rho U[\rho]) = 0, \qquad \rho_{|t=0} = \rho_0, \tag{5.57}$$

on $(0,+\infty)\times\Omega$, where Ω is a smooth domain of $\mathbb{R}^n$ (not necessary bounded), in case Ω has a boundary, the previous equation is supplemented with the no-flux boundary condition (ν denotes the outer unit normal to $\partial\Omega$):

$$(\nabla P(\rho) + U[\rho]\rho)\cdot\nu = 0 \text{ on } \partial\Omega. \tag{5.58}$$

For every $\rho\in\mathcal{P}(\Omega)$, we assume that the Helmholtz decomposition of the vector field $U[\rho]$:

$$U[\rho] = -W[\rho] + \nabla V[\rho], \tag{5.59}$$

with

$$\nabla\cdot W[\rho] = 0,\ W[\rho]\cdot\nu = 0 \text{ on } \partial\Omega,$$

satisfies the following assumptions.

Assumptions on the potential part V:

- $\nabla V[\rho]\in L^\infty_{\text{loc}}$ uniformly in ρ i.e for all $K\subset\subset\Omega$, there exists $C>0$ such that for all $\rho\in\mathcal{P}(\Omega)$,

$$\|\nabla V[\rho]\|_{\infty,K}\leqslant C. \tag{5.60}$$

  Note that by Rademacher's condition, this condition implies that $V[\rho]$ is differentiable a.e.,

- $V[\rho]$ is semi-convex uniformly in ρ i.e there exists C such that for all $\rho\in\mathcal{P}(\Omega)$, for every $y\in\Omega$ and every $x\in\Omega$, point of differentiability of $V[\rho]$:

$$V[\rho](y)\geqslant V[\rho](x) + \langle\nabla V[\rho](x), y-x\rangle - \frac{C}{2}|y-x|^2, \tag{5.61}$$

- there exists $C\geqslant 0$ such that for all $\rho\in\mathcal{P}(\Omega)$, for all $x\in\Omega$:

$$V[\rho](x)\geqslant -C(1+|x|), \tag{5.62}$$

- $\nabla V[\rho]\in L^2(\rho)$ uniformly in ρ, i.e there exists $C>0$ such that for all $\rho\in\mathcal{P}(\Omega)$,

$$\int_\Omega |\nabla V[\rho]|^2\,d\rho\leqslant C, \tag{5.63}$$

- There exists $C>0$ such that for all $\rho,\mu\in\mathcal{P}(\Omega)$,

$$\int_\Omega |\nabla V[\rho]-\nabla V[\mu]|^2\,d\rho\leqslant CW_2^2(\rho,\mu), \tag{5.64}$$

Assumptions on the divergence-free part W:

- there exists $C>0$ such that for all $\rho\in\mathcal{P}(\Omega)$ and

$$W[\rho]\in W^{1,1}_{\text{loc}}(\mathbb{R}^n)\qquad\text{and}\qquad |W[\rho](x)|\leqslant C(1+|x|),\ \text{for all } x\in\mathbb{R}^n \tag{5.65}$$

- There exists $C>0$ such that for all $\rho,\mu\in\mathcal{P}(\Omega)$,

$$\int_\Omega |W[\rho]-W[\mu]|^2\,d\rho\leqslant CW_2^2(\rho,\mu), \tag{5.66}$$

Assumptions on the internal energy F and the associated pressure P:
The nonlinear diffusion term is given by a continuous strictly convex superlinear (i.e. $F(\rho)/\rho\to+\infty$ as $\rho\to+\infty$) function $F:\mathbb{R}^+\to\mathbb{R}$ of class $\mathcal{C}^2((0,+\infty))$ which satisfies

$$F(0)=0,\ \text{and } P(\rho)\leqslant C(\rho+F(\rho)). \tag{5.67}$$

where $P(\rho) := \rho F'(\rho) - F(\rho)$ is the pressure associated to F. Moreover, we define $\mathcal{F} : \mathcal{P}(\Omega) \to \mathbb{R}$ by

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho(x))\,dx & \text{if } \rho \ll \mathcal{L}^n, \\ +\infty & \text{otherwise.} \end{cases}$$

And we assume that

$$\mathcal{F}(\rho) \geqslant -C(1 + M(\rho))^\alpha, \qquad \text{for all } \rho \in \mathcal{P}(\Omega), \tag{5.68}$$

where $\alpha \in (0,1)$ and $M(\rho) := \int_\Omega |x|^2\,d\rho(x)$ is the second moment of ρ.

The typical examples of energies we have in mind are $F(\rho) := \rho\log(\rho)$, which gives a linear diffusion driven by the laplacian, and $F(\rho) := \rho^m$ $(m > 1)$ which corresponds to the porous medium equation.

A weak solution of (5.57)-(5.58) is a curve $\rho : t \in (0,+\infty) \mapsto \rho(t,\cdot) \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ such that $\nabla P(\rho) \in \mathcal{M}^n([0,+\infty)\times\Omega)$ and

$$\int_0^\infty \left(\int_\Omega (\partial_t\phi\rho - \nabla\phi \cdot U[\rho]\rho)dx - \int_\Omega \nabla\phi \cdot d\nabla P(\rho) \right) dt = -\int_\Omega \phi(0,x)\rho_0(x)\,dx, \tag{5.69}$$

for every $\phi \in \mathcal{C}_c^\infty([0,+\infty)\times\mathbb{R}^n)$.

Our main result is the following:

Theorem 5.6. *Assume $\rho_0 \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ such that*

$$\mathcal{F}(\rho_0) < +\infty, \tag{5.70}$$

then (5.57) *admits at least one weak solution.*

The proof of this theorem is given in the next sections and is based on the following splitting scheme that combines pure transport steps by the divergence-free part of the drift U and Wasserstein gradient flow steps taking into account the potential V in a semi-implicit way. More precisely, given a time step $h > 0$, we construct by induction a sequence $\rho_h^k \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ by setting $\rho_h^0 = \rho_0$ and given ρ_h^k we find ρ_h^{k+1} using the following scheme:

- **pure transport phase:** we introduce an intermediate measure, $\tilde\rho_h^{k+1}$ (with $\tilde\rho_h^0 = \rho_0$) defined by

$$\tilde\rho_h^{k+1} = X_h^k(h,\cdot)_\# \rho_h^k, \tag{5.71}$$

where X_h^k is solution of

$$\begin{cases} \partial_t X_h^k = W[\rho_h^k] \circ X_h^k, \\ X_h^k(0,\cdot) = \mathrm{id}\,. \end{cases} \tag{5.72}$$

Since $W[\rho_h^k]$ satisfies (5.65), as recalled in section 5.2.1, DiPerna-Lions theory [57] implies that X_h^k is well defined. Moreover, since $W[\rho_h^k]$ is divergence-free then X_h^k preserves the Lebesgue measure and leaves the domain Ω invariant thanks to the fact that $W[\rho]$ is tangential to $\partial\Omega$. Therefore $\tilde\rho_h^{k+1} = \rho_h^k({X_h^k}^{-1})$ which implies the conservation of the internal energy:

$$\begin{aligned} \mathcal{F}(\tilde\rho_h^{k+1}) &= \int_\Omega F(\rho_h^k({X_h^k}^{-1}(x)))\,dx \\ &= \int_\Omega F(\rho_h^k(x))\,dx = \mathcal{F}(\rho_h^k) \end{aligned} \tag{5.73}$$

In addition, we can see $\tilde\rho_h^{k+1}$ is the value at time h of the solution μ of the continuity equation

$$\begin{cases} \partial_t\mu + \mathrm{div}(\mu W[\rho_h^k]) = 0, \\ \mu_{|t=0} = \rho_h^k. \end{cases} \tag{5.74}$$

Thanks to these observations, we can easily control the W_2-distance between $\tilde{\rho}_h^{k+1}$ and ρ_h^k. Indeed, using Benamou-Brenier formula and (5.65), we obtain

$$\begin{aligned} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^k) &\leqslant h \int_0^h \int_\Omega |W[\rho_h^k]|^2 \, d\mu_t dt \\ &\leqslant Ch \int_0^h \int_\Omega (1 + |x|^2) \, d\mu_t dt \\ &\leqslant Ch \int_0^h (1 + M(\mu_t)) \, dt \end{aligned}$$

Moreover,

$$\begin{aligned} \frac{d}{dt} M(\mu_t) &= \int_\Omega |x|^2 \partial_t \mu_t \\ &= -\int_\Omega |x|^2 \operatorname{div}(W[\rho_h^k]\mu_t) \\ &= 2 \int_\Omega x \cdot W[\rho_h^k] \mu_t \\ &\leqslant C(M(\mu_t) + 1). \end{aligned}$$

We obtain the last line using (5.65), Cauchy-Schwarz inequality and Young's inequality. Then,

$$M(\mu_t) \leqslant C(t+1)e^{Ct} \leqslant 2Ce^C \text{ for } t \leqslant 1,$$

which implies

$$W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^k) \leqslant Ch^2. \tag{5.75}$$

- **semi-implicit JKO scheme:** In the second step we use a semi-implict version of the Jordan-Kinderlehrer-Otto scheme [65], introduced by Di Francesco and Fagioli in [54] and used in [73], with $\tilde{\rho}_h^{k+1}$ being the measure defined in the previous step. More precisely, we select ρ_h^{k+1} as a solution of

$$\inf_{\rho \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)} \mathcal{E}_h(\rho|\tilde{\rho}_h^{k+1}) := W_2^2(\rho, \tilde{\rho}_h^{k+1}) + 2h \left(\mathcal{F}(\rho) + \mathcal{V}(\rho|\tilde{\rho}_h^{k+1}) \right), \tag{5.76}$$

where

$$\mathcal{V}(\rho|\mu) := \int_\Omega V[\mu] \, d\rho.$$

By standard compactness and lower semicontinuity argument, (5.76) admits at least one solution (see for example [65, 73]) so the sequence ρ_h^k is well defined (it is even actually unique by strict convexity of $\mathcal{E}_h(.|\tilde{\rho}_h^{k+1})$).

To summarize, given a time step $h > 0$, we construct by induction two sequences ρ_h^k and $\tilde{\rho}_h^k$ with the following splitting scheme: $\rho_h^0 = \tilde{\rho}_h^0 = \rho_0$ and for all $k \geqslant 0$,

$$\begin{cases} \tilde{\rho}_h^{k+1} = X_h^k(h, \cdot)_\# \rho_h^k, \\ \rho_h^{k+1} \in \operatorname{argmin}_{\rho \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)} \left\{ W_2^2(\rho, \tilde{\rho}_h^{k+1}) + 2h \left(\mathcal{F}(\rho) + \mathcal{V}(\rho|\tilde{\rho}_h^{k+1}) \right) \right\}. \end{cases} \tag{5.77}$$

We finally introduce three different interpolations:

- We denote ρ_h the usual piecewise constant interpolation of the sequence ρ_h^k

$$\rho_h(t, \cdot) := \rho_h^{k+1} \qquad \text{if } t \in (hk, h(k+1)], \tag{5.78}$$

- similarly, we interpolate in a piecewise constant way the sequence $\tilde{\rho}_h^k$:

$$\tilde{\rho}_h^1(t, \cdot) := \tilde{\rho}_h^{k+1} \qquad \text{if } t \in (hk, h(k+1)], \tag{5.79}$$

- finally, we denote by $\tilde{\rho}_h^2$ the *continuous* interpolation of $\tilde{\rho}_h^k$

$$\tilde{\rho}_h^2(t,\cdot) := X_h^k(t-hk,\cdot)_\# \rho_h^k \qquad \text{if } t \in (hk, h(k+1)]. \tag{5.80}$$

We remark that on $(hk, h(k+1)]$, $\tilde{\rho}_h^2$ is the solution on $(0,h)$ of the continuity equation (5.74).

The next two sections are devoted to the proof of theorem 5.57. In section 5.2.3, we derive various estimates on the sequences generated by the splitting scheme above, in particular thanks to the Euler-Lagrange equation of the semi-implicit JKO steps. This enables us to pass to the limit as the time step goes to 0 (the difficult term being of course the nonlinear pressure term) and thus to conclude the existence proof, this is done in section 5.2.4.

5.2.3 Estimates

Basic a priori estimates

Using the semi-implicit JKO scheme we first obtain the following a priori estimates on ρ_h, $\tilde{\rho}_h^1$ and $\tilde{\rho}_h^2$.

Proposition 5.7. *There exists $h_0 > 0$, such that for $T > 0$, there exists $C > 0$ such that, for all h, k, with $h \in (0, h_0)$ and $hk < T$, $N = \lceil \frac{T}{h} \rceil$, we have*

$$M(\rho_h^k) \leqslant C, \tag{5.81}$$
$$\mathcal{F}(\rho_h^k) \leqslant C, \tag{5.82}$$
$$\textstyle\sum_{k=0}^{N-1} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) \leqslant Ch. \tag{5.83}$$

Proof. Using $\tilde{\rho}_h^{k+1}$ as a competitor of ρ_h^{k+1} in (5.76), we obtain

$$\frac{1}{2h} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant \mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^{k+1}) + \int_\Omega V[\tilde{\rho}_h^{k+1}] \left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right). \tag{5.84}$$

Let γ be the optimal transport plan between $\tilde{\rho}_h^{k+1}$ and ρ_h^{k+1}. Then we have

$$\begin{aligned}
\int_\Omega V[\tilde{\rho}_h^{k+1}] \left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right) &= \int_\Omega \left(V[\tilde{\rho}_h^{k+1}](x) - V[\tilde{\rho}_h^{k+1}](y)\right) d\gamma(x,y) \\
&= \int_\Omega \left(V[\tilde{\rho}_h^{k+1}](x) - V[\tilde{\rho}_h^{k+1}](y) + \nabla V[\tilde{\rho}_h^{k+1}](x)\cdot(y-x)\right) d\gamma(x,y) \\
&- \int_\Omega \nabla V[\tilde{\rho}_h^{k+1}](x)\cdot(y-x)\, d\gamma(x,y).
\end{aligned}$$

Using (5.61) for the first part and Cauchy-Schwarz inequality and (5.63) for the second part of the right hand side, we find

$$\begin{aligned}
\int_\Omega V[\tilde{\rho}_h^{k+1}] \left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right) &\leqslant \frac{C}{2}\int_\Omega |x-y|^2\,\gamma(x,y) + C\Big(\int_\Omega |x-y|^2\,\gamma(x,y)\Big)^{\frac{1}{2}} \\
&= \frac{C}{2} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) + C W_2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}).
\end{aligned}$$

Choosing $h \leq h_0 \leq \frac{1}{2C}$ and using Young's inequality

$$W_2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) \leq \frac{1}{8hC} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) + 2Ch,$$

(5.84) becomes

$$\frac{1}{8h} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant \mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^{k+1}) + Ch.$$

Now using (5.73), to recover a telescopic sum, and summing over k, we obtain

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1}) \leqslant 8h\left(\mathcal{F}(\rho_0)-\mathcal{F}(\rho_h^N)+CT\right),$$

this inequality and (5.70) imply (5.82). In addition, since the lower bound of $\mathcal{F}$ is controlled by the second moment,

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1}) \leqslant 8h\left(\mathcal{F}(\rho_0)+C(1+M(\rho_h^N))^\alpha+CT\right). \tag{5.85}$$

But, with (5.85) and by standard arguments (see [65, 73]), we deduce that $M(\rho_h^k)$ satisfies (5.81) and then (5.85), (5.70) and (5.81) give (5.83). □

Remark 5.8. *Using estimate* (5.75) *between* $\tilde\rho_h^{k+1}$ *and* ρ_h^k*, and* (5.83)*, we also have*

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^k,\rho_h^{k+1}) \leqslant Ch \text{ and } \sum_{k=0}^{N-1} W_2^2(\tilde\rho_h^k,\tilde\rho_h^{k+1}) \leqslant Ch.$$

Moreover, using (5.73)*, we have for all* $t\in[0,T]$*,*

$$\mathcal{F}(\tilde\rho_h^1(t)),\mathcal{F}(\tilde\rho_h^2(t)) \leqslant C \qquad \text{and} \qquad M(\tilde\rho_h^1(t)),M(\tilde\rho_h^2(t)) \leqslant C.$$

Discrete Euler-Lagrange equation and stronger estimates

Let us start with the Euler-Lagrange equation of (5.76).

Proposition 5.9. *For all* $k\geqslant 0$*, we have* $P(\rho_h^{k+1})\in W^{1,1}(\Omega)$ *and*

$$h\left(\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1}+\nabla P(\rho_h^{k+1})\right) = -\nabla\varphi_h^{k+1}\rho_h^{k+1} \qquad a.e, \tag{5.86}$$

where φ_h^{k+1} *is a Kantorovich potential from* ρ_h^{k+1} *to* $\tilde\rho_h^{k+1}$ *(so that its gradient is unique* ρ_h^{k+1}*-a.e.) for* W_2*.*

Proof. The proof is the same as in [1, 73] for example. We start by taking the first variation in the semi-implicit JKO scheme along the flow of a smooth vector field. Let $\xi\in\mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$ be given and Φ_τ the corresponding flow defined by

$$\partial_\tau\Phi_\tau = \xi\circ\Phi_\tau,\ \Phi_0=\mathrm{id}\,.$$

We define a pertubation of ρ_h^{k+1} by $\rho_\tau := \Phi_{\tau\#}\rho_h^{k+1}$. Then we get

$$\frac{1}{\tau}\left(\mathcal{E}_h(\rho_\tau|\tilde\rho_h^{k+1})-\mathcal{E}_h(\rho_h^{k+1}|\tilde\rho_h^{k+1})\right)\geqslant 0. \tag{5.87}$$

By standard computations, we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(W_2^2(\rho_\tau,\tilde\rho_h^{k+1})-W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1})) \leqslant \int_{\Omega\times\Omega}(x-y)\cdot\xi(x)\,d\gamma_h^{k+1}(x,y), \tag{5.88}$$

with γ_h^{k+1} is the W_2-optimal transport plan in $\Pi(\rho_h^{k+1},\tilde\rho_h^{k+1})$ and $\gamma_h^{k+1}=(\mathrm{id}\times T_h^{k+1})_\#\rho_h^{k+1}$ with $T_h^{k+1}=\mathrm{id}-\nabla\varphi_h^{k+1}$. Moreover, using (5.67), (5.82) and Lebesgue's dominated convergence theorem, we obtain

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{F}(\rho_\tau)-\mathcal{F}(\rho_h^{k+1})) \leqslant -\int_\Omega P(\rho_h^{k+1}(x))\,\mathrm{div}(\xi(x))\,dx. \tag{5.89}$$

Finally,

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{V}(\rho_\tau|\tilde\rho_h^{k+1})-\mathcal{V}(\rho_{i,h}^{k+1}|\tilde\rho_h^{k+1})) \leqslant \int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\xi\rho_h^{k+1}\,dx, \tag{5.90}$$

If we combine (5.87), (5.88), (5.89) and (5.90), and if we replace ξ by $-\xi$, we find that, for all $\xi \in \mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega \nabla\varphi_h^{k+1}\cdot\xi\rho_h^{k+1} - h\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi) + h\int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\xi\rho_h^{k+1} = 0. \tag{5.91}$$

Now we claim that $P(\rho_h^{k+1}) \in W^{1,1}(\Omega)$. Indeed, since P is controlled by F thanks to assumption (5.67), (5.82) gives $P(\rho_h^{k+1}) \in L^1(\Omega)$. Moreover, using (5.91), we obtain

$$\left|\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi)\right| \leqslant \left[\int_\Omega \frac{|\nabla\varphi_h^k(y)|}{h}\rho_h^{k+1} + \int_\Omega |\nabla V[\tilde\rho_h^{k+1}]|\rho_h^{k+1}\right]\|\xi\|_{L^\infty(\Omega)}.$$

But using (5.64), (5.63), (5.83) and Cauchy-Schwarz inequality we get

$$\begin{aligned}
\int_\Omega |\nabla V[\tilde\rho_h^{k+1}]|\rho_h^{k+1} &\leqslant \left[\int_\Omega |\nabla V[\tilde\rho_h^{k+1}] - \nabla V[\rho_h^{k+1}]|\rho_h^{k+1} + \int_\Omega |\nabla V[\rho_h^{k+1}]|\rho_h^{k+1}\right]\\
&\leqslant \left[\Big(\int_\Omega |\nabla V[\tilde\rho_h^{k+1}] - \nabla V[\rho_h^{k+1}]|^2\rho_h^{k+1}\Big)^{1/2} + \Big(\int_\Omega |\nabla V[\rho_h^{k+1}]|^2\rho_h^{k+1}\Big)^{1/2}\right]\\
&\leqslant C\left[W_2(\tilde\rho_h^{k+1},\rho_h^{k+1}) + 1\right]\\
&\leqslant C.
\end{aligned}$$

We thus have

$$\left|\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi)\right| \leqslant \left[\frac{W_2(\tilde\rho_h^{k+1},\rho_h^{k+1})}{h} + C\right]\|\xi\|_{L^\infty(\Omega)}.$$

This implies $P(\rho_h^{k+1}) \in BV(\Omega)$ and $\nabla P(\rho_h^{k+1}) = \left(-\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1} - \frac{\nabla\varphi_h^{k+1}}{h}\rho_h^{k+1}\right)$ in $\mathcal{M}^n(\Omega)$. In fact, $P(\rho_h^{k+1})$ is in $W^{1,1}(\Omega)$ because $\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1} + \frac{\nabla\varphi_h^k}{h}\rho_h^{k+1} \in L^1(\Omega)$ and we have proved (5.86).

□

We immediately deduce an $L^1((0,T),BV(\Omega))$ estimate for $P(\rho_h)$:

Corollary 5.10. *For all $T > 0$, we have*

$$\|P(\rho_h)\|_{L^1((0,T);W^{1,1}(\Omega))} \leqslant CT. \tag{5.92}$$

Proof. If we integrate (5.86), we obtain

$$h\int_\Omega |\nabla P(\rho_h^{k+1})| \leqslant W_2(\rho_h^{k+1},\tilde\rho_h^{k+1}) + Ch,$$

Then we sum from $k = 0$ to $N-1$ and thanks to (5.83), we have

$$\int_0^T\int_\Omega |\nabla P(\rho_h)| \leqslant CT.$$

We conclude thanks to (5.67) and (5.82). □

Proposition 5.11. *Let $h > 0$, $N \in \mathbb{N}^*$, $T := Nh$, $t_k := hk$, for $k = 0,\cdots,N$, then, for every $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$*

$$\begin{aligned}
\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(\partial_t\phi(t,x) + W[\rho_h(t-h)]\cdot\nabla\phi(t,x))\,dxdt &= h\sum_{k=0}^{N-1}\int_\Omega \nabla P(\rho_h^{k+1}(x))\cdot\nabla\phi(t_k,x)\,dx\\
&+ h\sum_{k=0}^{N-1}\int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\nabla\phi(t_k,x)\rho_h^{k+1}\,dx\\
&+ \sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\phi(t_k,\cdot)](x,y)d\gamma_h^{k+1}(x,y)\\
&- \int_\Omega \rho_0(x)\phi(0,x)\,dx,
\end{aligned}$$

with, for all $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\Omega)}|x-y|^2,$$

and γ_h^{k+1} *is the optimal transport plan in* $\Pi(\rho_h^{k+1},\tilde\rho_h^{k+1})$.

Proof. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, multiplying (5.86) by $\nabla\varphi$ and integrating on Ω, we obtain

$$-\int_\Omega \nabla\varphi_h^{k+1}\cdot\nabla\varphi\rho_h^{k+1} = h\left(\int_\Omega \nabla P(\rho_h^{k+1})\cdot\nabla\varphi + \int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\nabla\varphi\rho_h^{k+1}\right).$$

But, we can rewrite the left hand side by

$$-\int_\Omega \nabla\varphi_h^{k+1}\cdot\nabla\varphi\rho_h^{k+1} = \int_{\Omega\times\Omega}(y-x)\cdot\nabla\varphi(x)\,d\gamma_h^{k+1}(x,y).$$

A second-order Taylor-Lagrange formula then gives

$$\begin{aligned}\int_{\Omega\times\Omega}(y-x)\cdot\nabla\varphi(x)\,d\gamma_h^{k+1} &= \int_{\Omega\times\Omega}(\varphi(y)-\varphi(x))\;d\gamma_h^{k+1}(x,y) - \int_{\Omega\times\Omega}\mathcal{R}[\varphi](x,y)d\gamma_h^{k+1}(x,y)\\ &= \int_\Omega \varphi(\tilde\rho_h^{k+1}-\rho_h^{k+1}) - \int_{\Omega\times\Omega}\mathcal{R}[\varphi](x,y)d\gamma_h^{k+1}(x,y). \qquad (5.93)\end{aligned}$$

Now let $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$, we have

$$\begin{aligned}&\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(\partial_t\phi(t,x)+W[\rho_h(t-h)]\cdot\nabla\phi(t,x))\,dxdt\\ =\;&\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)(\partial_t\phi+W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x))\,dxdt\end{aligned}$$

But, on $[t_k,t_{k+1}]$,

$$\frac{d}{dt}[\phi(t,X_h^k(t-t_k,x))] = (\partial_t\phi+W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x)).$$

Then,

$$\begin{aligned}&\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)(\partial_t\phi+W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x))\,dxdt\\ &\qquad=\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)\frac{d}{dt}[\phi(t,X_h^k(t-t_k,x))]\,dxdt\\ &\qquad=\sum_{k=0}^{N-1}\int_\Omega \rho_h^k(x)\left[\phi(t_{k+1},X_h^k(h,x))-\phi(t_k,x)\right]\,dx\\ &\qquad=\sum_{k=0}^{N-1}\int_\Omega \left[\phi(t_{k+1},x)\tilde\rho_h^{k+1}(x)-\phi(t_k,x)\rho_h^k(x)\right]\,dx\\ &\qquad\qquad=\sum_{k=0}^{N-1}\int_\Omega \phi(t_{k+1},x)\left(\tilde\rho_h^{k+1}(x)-\rho_h^{k+1}(x)\right)-\int_\Omega\phi(0,x)\rho_0(x)\,dx.\end{aligned}$$

Then the proof is complete by applying (5.93) with $\varphi=\phi(t_{k+1},\cdot)$.

□

5.2.4 Convergence and proof of Theorem 5.6

Weak and strong convergences

Using a refined version of Ascoli theorem (see [4]), estimate (5.83) and remark 5.8 and taking subsequences, if necessary, we have that, for every $T < +\infty$, ρ_h, $\tilde{\rho}_h^1$ and $\tilde{\rho}_h^2$ converge in $L^\infty((0,T), W_2)$ to some respective limits ρ, $\tilde{\rho}^1$ and $\tilde{\rho}^2$:

$$\sup_{t\in[0,T]} \max(W_2(\rho_h(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^1(t,.),\tilde{\rho}^1(t,.)), W_2(\tilde{\rho}_h^2(t,.),\tilde{\rho}^2(t,.))) \to 0 \text{ as } h \to 0.$$

In fact, these three sequences have to converge to the same limit, ρ. Indeed, for all $\varphi \in \mathcal{C}_c^\infty((0,T)\times\Omega)$,

$$\begin{aligned}
\int_0^T \int_\Omega \varphi(\rho_h - \tilde{\rho}_h^1) &= \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \int_\Omega \varphi(\rho_h^{k+1} - \tilde{\rho}_h^{k+1}) \\
&\leqslant Ch \sum_{k=0}^{N-1} W_2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \\
&\leqslant ChN^{1/2} \left(\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \right)^{1/2} \\
&\leqslant CT^{1/2} h,
\end{aligned}$$

because of (5.83). With a similar computation, we find that ρ_h and $\tilde{\rho}_h^2$ converge to the same limit. We thus have

$$\sup_{t\in[0,T]} \max(W_2(\rho_h(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^1(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^2(t,.),\rho(t,.))) \to 0 \text{ as } h \to 0. \tag{5.94}$$

Moreover it is classical to deduce from (5.83) and remark 5.8 an Hölder-like estimate of the form
$W_2(\rho_h(t,.),\rho_h(s,.)) \leq C\sqrt{|t-s|+h}$ from which one deduces that the limit curve ρ actually belongs to $\mathcal{C}^{1/2}((0,T),W_2)$. This kind of convergence will be enough to pass to the limit in $\nabla V[\tilde{\rho}_h^1]\rho_h$ and $W[\rho_h]\tilde{\rho}_h^2$, because of assumptions (5.64) and (5.66), but we will need a stronger convergence to deal with the nonlinear diffusion term $P(\rho_h)$. Fo this purpose, we will use an extension of the Aubin-Lions Lemma due to Rossi and Savaré in [107]:

Theorem 5.12 (th. 2 in [107])**.** *On a Banach space B, let be given*

- *a normal coercive integrand $\mathcal{G} : B \to \mathbb{R}^+$, i.e, $\mathcal{G}$ is l.s.c and its sublevels are relatively compact in B,*
- *a pseudo-distance $g : B \times B \to [0,+\infty]$, i.e, g is l.s.c and*

$$[g(\rho,\mu) = 0,\ \rho,\mu \in B \text{ with } \mathcal{G}(\rho),\mathcal{G}(\mu) < \infty] \Rightarrow \rho = \mu.$$

Let $T > 0$ and U be a set of measurable functions $u : (0,T) \to B$. Under the hypotheses that

$$\sup_{u\in U} \int_0^T \mathcal{G}(u(t))\,dt < +\infty \qquad \text{and} \qquad \lim_{h\searrow 0} \sup_{u\in U} \int_0^{T-h} g(u(t+h),u(t))\,dt = 0, \tag{5.95}$$

U contains a subsequence $(u_n)_{n\in\mathbb{N}}$ which converges (strongly in B) in measure with respect to $t \in (0,T)$ to a limit $u_\star : (0,T) \to B$.

We now apply this theorem to $B = L^1(\Omega)$, $U = \{\rho_h\}_h$, g defined by

$$g(\rho,\mu) := \begin{cases} W_2(\rho,\mu) & \text{if } \rho,\mu \in \mathcal{P}_2(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

and $\mathcal{G}$ by

$$\mathcal{G}(\rho) := \begin{cases} \mathcal{F}(\rho) + \|P(\rho)\|_{BV(\Omega)} + M(\rho) & \text{if } \rho \in \mathcal{P}_2^{ac}(\Omega), P(\rho) \in BV(\Omega) \text{ and } F(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Lemma 5.13. *$\mathcal{G}$ is l.s.c on $L^1(\Omega)$ and its sublevels are relatively compact in $L^1(\Omega)$.*

Proof. Let us start by proving that sublevels of $\mathcal{G}$ are relatively compact in $L^1(\Omega)$. Let

$$A_c := \left\{\rho \in L^1(\Omega) \,:\, \mathcal{G}(\rho) \leqslant c\right\}$$

and (ρ_k) be a sequence in A_c then $P(\rho_k)$ is bounded in $BV(\Omega)$ thus, up to a subsequence $P(\rho_k)$ converges to some Φ in $L^1_{\rm loc}(\Omega)$ and a.e.. Since P is continuous, one to one and its inverse is continuous, ρ_k converges to $\rho := P^{-1}(\Phi)$ a.e.; and, since $\mathcal{G}(\rho_k) \leqslant c$ and F is superlinear, ρ_k is uniformly integrable, using Vitali's convergence theorem, we obtain that ρ_k converges to ρ in $L^1(K \cap \Omega)$ for every compact K. To conclude that there is convergence in $L^1(\Omega)$, we use the fact that the second momentum of ρ_k is uniformly bounded:

$$\begin{aligned}\int_\Omega |\rho_k - \rho| &\leqslant \int_{\Omega\setminus B_R} \frac{|x|^2}{R^2}|\rho_k - \rho| + \int_{B_R\cap\Omega} |\rho_k - \rho| \\ &\leqslant \frac{2c}{R^2} + \int_{B_R\cap\Omega} |\rho_k - \rho|.\end{aligned}$$

The first term in the right hand can be made arbitrary small by choosing R large enough and the second term converges to zero by $L^1(B_R \cap \Omega)$-convergence.

Now we have to show the lower semi-continuity of $\mathcal{G}$ on $L^1(\Omega)$. Let (ρ_k) be a sequence which converges strongly to ρ in $L^1(\Omega)$ with (without loss of generality) $\sup_k \mathcal{G}(\rho_k) \leqslant C < +\infty$. Without loss of generality, we can assume that ρ_k converges to ρ a.e. Since $\sup_k \mathcal{G}(\rho_k) \leqslant C$, $P(\rho_k)$ is uniformly bounded in $BV(\Omega)$ so $P(\rho_k)$ converges weakly to μ in $BV(\Omega)$. Moreover, $P(\rho_k)$ converges strongly to μ in $L^1_{\rm loc}(\Omega)$. We can conclude that $\mu = P(\rho)$ and by lower semi-continuity of $\mathcal{F}$, M and the BV-norm we have

$$\mathcal{G}(\rho) \leqslant \liminf_{k\nearrow+\infty} \mathcal{G}(\rho_k).$$

□

Thanks to lemma 5.13, to apply theorem 5.12, it remains to verify (5.95). The first condition of (5.95) is satisfied because of the estimate on the momentum, (5.81), on the internal energy $\mathcal{F}$, (5.82) and on the gradient of $P(\rho_h)$ (5.92). The second condition of (5.95) comes from the estimate on the distance (5.83) and remark 5.8 (see for example [55] for a detailed proof). Then theorem 5.12 implies that ρ_h converges in measure with respect to t in $L^1(\Omega)$ to ρ. Since convergence in measure implies a.e convergence, up to a subsequence, we may also assume that $\rho_h(t,.)$ converges strongly in $L^1(\Omega)$ to $\rho(t,.)$ for a.e. t. Then Lebesgue's dominated convergence theorem implies that ρ_h converges strongly in $L^1((0,T)\times\Omega)$ to ρ.

Thanks to (5.67) and (5.82) $P(\rho_h)$ is uniformly bounded in $L^\infty((0,T), L^1(\Omega))$. In addition, by corollary 5.10, $P(\rho_h)$ is uniformly bounded in $L^1((0,T), W^{1,1}(\Omega))$. Thanks to the Sobolev embedding, we deduce that $P(\rho_h)$ is uniformly bounded in $L^\infty((0,T), L^1(\Omega)) \cap L^1((0,T), L^{n/n-1}(\Omega))$. To have uniform integrability of $P(\rho_h)$ both in the time and space variables, the following will be useful:

Lemma 5.14. *Let $p > 1$, $q := \frac{2p-1}{p}$ and $f \in L^\infty((0,T), L^1(\Omega)) \cap L^1((0,T), L^p(\Omega))$ then $f \in L^q((0,T)\times\Omega)$ and we have*

$$\|f\|^q_{L^q((0,T)\times\Omega)} \leqslant \|f\|^{q-1}_{L^\infty_t(L^1_x)} \|f\|_{L^1_t(L^p_x)}.$$

Proof. Writing

$$\frac{1}{q} = \frac{p}{2p-1} = \theta + \frac{1-\theta}{p}, \quad \text{with } \theta = \frac{p-1}{2p-1} \in (0,1)$$

and observing that $(1-\theta)q = 1$ and $\theta q = q-1$, the interpolation inequality yields that

$$\|f\|^q_{L^q_x} \leqslant \|f\|^{q-1}_{L^1_x} \|f\|_{L^p_x},$$

since $f \in L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^p(\Omega))$ this implies that $f \in L^q((0,T)\times\Omega)$ and

$$\begin{aligned}\|f\|^q_{L^q((0,T)\times\Omega)} = \int_0^T \|f\|^q_{L^q_x} = &\leqslant \int_0^T \|f\|^{q-1}_{L^1_x}\|f\|_{L^p_x} \\ &\leqslant \|f\|^{q-1}_{L^\infty_t(L^1_x)}\|f\|_{L^1_t(L^p_x)}.\end{aligned}$$

□

Applying lemma 5.14 we deduce that $P(\rho_h)$ is uniformly bounded in $L^{(n+1)/n}((0,T)\times\Omega)$. This implies that $P(\rho_h)$ is uniformly integrable and since we know that it converges a.e. to $P(\rho)$, we can deduce from Vitali's convergence theorem that $P(\rho_h)$ converges strongly to $P(\rho)$ in $L^1_{\rm loc}((0,T)\times\Omega)$.

Thanks to Corollary 5.10 we deduce that $\nabla P(\rho_h)$ converges vaguely to $\nabla P(\rho)$ in $\mathcal{M}^n_{\rm loc}((0,T)\times\Omega)$. In fact, we have $\nabla P(\rho)$ in $\mathcal{M}^n((0,T)\times\Omega)$ and narrow convergence of $\nabla P(\rho_h)$ to $\nabla P(\rho)$, thanks to Prokhorov Theorem and the following tightness estimate:

Lemma 5.15. *The family $\nabla P(\rho_h)$, viewed as vector-valued measures on $[0,T]\times\Omega$, is tight, more precisely, for every h and every A measurable, $A\subset\Omega$*

$$\int_0^T\int_A |\nabla P(\rho_h)| \leq C\Big(1+\sqrt{h}\Big)\Big(\int_0^T\int_A \rho_h(t,x)\,dx\,dt\Big)^{1/2}. \tag{5.96}$$

Proof. Integrating (5.86) on $(0,T)\times A$ together with Cauchy Schwarz inequality and (5.83), we get (taking $N = \lceil \frac{T}{h}\rceil + 1$, say)

$$\begin{aligned}\int_0^T\int_A |\nabla P(\rho_h)| &\leqslant \sum_{k=0}^N \int_A |\nabla\varphi_h^{k+1}|\rho_h^{k+1} + \int_0^T\int_A |\nabla V[\tilde\rho_h^1]|\rho_h \\ &\leqslant \sum_{k=0}^N \Big(\int_\Omega |\nabla\varphi_h^{k+1}|^2\rho_h^{k+1}\Big)^{1/2}\Big(\int_A \rho_h^{k+1}\Big)^{1/2} + \int_0^T\int_A |\nabla V[\tilde\rho_h^1]|\rho_h \\ &\leqslant \Big(\sum_{k=0}^N W_2^2(\tilde\rho_h^{k+1},\rho_h^{k+1})\Big)^{1/2}\Big(\int_0^T\int_A \rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2} + \int_0^T\int_A |\nabla V[\tilde\rho_h^1]|\rho_h \\ &\leqslant C\sqrt{h}\Big(\int_0^T\int_A \rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2} + \int_0^T\int_A |\nabla V[\tilde\rho_h^1]|\rho_h.\end{aligned}$$

Moreover, with Cauchy Schwarz inequality, (5.63) and (5.64), we also have

$$\begin{aligned}\int_0^T\int_A |\nabla V[\tilde\rho_h^1]|\rho_h &\leqslant \int_0^T\int_A |\nabla V[\rho_h]|\rho_h + \int_0^T\int_A |\nabla V[\tilde\rho_h^1] - \nabla V[\rho_h]|\rho_h \\ &\leqslant C\Big(1+\sqrt{h}\Big)\Big(\int_0^T\int_A \rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2}\end{aligned}$$

which proves (5.96). The tightness of $\nabla P(\rho_h)$ therefore immediately follows from that of ρ_h and (5.96).

□

We can summarize all of this in the next result:

Theorem 5.16. *Up to a subsequence ρ_h converges strongly in $L^1((0,T)\times\Omega)$, $P(\rho_h)$ converges strongly to $P(\rho)$ in $L^1_{\rm loc}((0,T)\times\Omega)$ and $\nabla P(\rho_h)$ converges to $\nabla P(\rho)$ narrowly in $\mathcal{M}((0,T)\times\Omega)$.*

End of the proof of theorem 5.57

So we just have to check the convergence in the transport terms, in what follows the test-function ϕ belongs again to $\mathcal{C}^\infty_c([0,T)\times\mathbb{R}^n)$.

- *term in W:* We have to show:

$$\int_0^T \int_\Omega \tilde\rho_h^2(t,x) W[\rho_h(t-h)](x)\cdot\nabla\phi(t,x)\,dxdt \to \int_0^T\int_\Omega \rho(t,x)W[\rho(t,\cdot)](x)\cdot\phi(t,x)\,dxdt.$$

We first have

$$\begin{aligned}\int_0^T \int_\Omega \tilde\rho_h^2(t,x) W[\rho_h(t-h)](x)\cdot\nabla\phi(t,x)\,dxdt &= \int_0^T \int_\Omega \tilde\rho_h^2(t,x) W[\tilde\rho_h^2(t)](x)\cdot\nabla\phi(t,x)\,dxdt\\ &+\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt.\end{aligned}$$

The second term in the right hand side goes to zero when h goes to 0. Indeed,

$$\begin{aligned}&\Big|\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt\Big|\\ &\qquad\leqslant C\int_0^T\left(\int_\Omega \tilde\rho_h^2(t,x)|W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)(x)]|^2\,dx\right)^{1/2}dt,\end{aligned}$$

then using (5.66),

$$\begin{aligned}&\Big|\textstyle\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt\Big|\\ &\leqslant C\textstyle\int_0^T W_2(\rho_h(t-h),\tilde\rho_h^2(t))\,dt\\ &\leqslant C\textstyle\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} W_2(\rho_h^k, X_h^k(t-t_k)_\#\rho_h^k)\,dt\\ &\leqslant CTh,\end{aligned}$$

because of (5.75). Moreover, using (5.64), we get

$$\begin{aligned}&\Big|\int_0^T\int_\Omega(\tilde\rho_h^2(t,x)W[\tilde\rho_h^2(t)](x)\cdot\nabla\phi(t,x)-\rho(t,x)W[\rho(t)](x)\cdot\nabla\phi(t,x))\,dxdt\Big|\\ &\qquad\leqslant C\int_0^T\int_\Omega\tilde\rho_h^2(t,x)|W[\tilde\rho_h^2(t)](x)-W[\rho(t)](x))|\,dxdt\\ &\qquad+\Big|\int_0^T\int_\Omega(\rho(t,x)-\tilde\rho_h^2(t,x))W[\rho(t)](x)\cdot\nabla\phi(t,x)\,dxdt\Big|\\ &\leqslant CT\sup_{t\in[0,T]}W_2(\tilde\rho_h^2(t),\rho(t))+\Big|\int_0^T\int_\Omega(\rho(t,x)-\tilde\rho_h^2(t,x))W[\rho(t)](x)\cdot\nabla\phi(t,x)\,dxdt\Big|\end{aligned}$$

the first term in the right hand-side converges to 0 because of (5.94). As for the second one, it also converges to 0, because $W[\rho]\cdot\nabla\phi$ belongs to $L^\infty((0,T)\times\Omega)$ and $\tilde\rho_h^2$ is uniformly integrable by remark 5.8 and the superlinearity of F, hence, up to a subsequence it converges to ρ weakly in $L^1((0,T)\times\Omega)$.

- *term in ∇V:* We claim that

$$h\sum_{k=0}^{N-1}\int_\Omega \nabla V[\tilde\rho_h^{k+1}](x)\cdot\nabla\phi(t_k,x)\rho_h^{k+1}\,dx\to\int_0^T\int_\Omega\nabla V[\rho(t,\cdot)](x)\cdot\nabla\phi(t,x)\rho(t,x)\,dxdt.$$

The proof is the same as the previous one for W, using (5.64), (5.60) and the convergence of ρ_h to ρ in $L^1((0,T)\times\Omega))\cap L^\infty((0,T),W_2)$.

5.2.5 On extension to systems and uniqueness

The splitting transport-JKO scheme described above, can easily be adapted, under suitable assumptions to the case of systems for the evolution of N species coupled by nonlocal drifts:

$$\partial_t \rho_i - \Delta P_i(\rho_i) - \mathrm{div}(\rho_i U_i[\rho_1, \cdots, \rho_N]) = 0,\ \rho_i(0,.) = \rho_{i,0},\ i = 1, \cdots, N, \tag{5.97}$$

where $P_i(s) = sF_i'(s) - F_i(s)$ is the pressure associated to a strictly convex superlinear function F_i with corresponding internal energy $\mathcal{F}_i(\rho_i) := \int F_i(\rho_i(x))dx$. Decomposing each drift $U_i[\rho_1, \cdots, \rho_N] = \nabla V_i[\rho_1, \cdots, \rho_N] - W_i[\rho_1, \cdots, \rho_N]$ with $\mathrm{div}(W_i[\rho_1, \cdots, \rho_N]) = 0$ and under similar assumptions as in paragraph 5.2.2, one can show, by similar arguments as above, convergence as $h \to 0$ to a solution of (5.97) of the following splitting scheme.

Starting form $\rho^0_{i,h} = \rho_{i,0}$ and given $\rho^k_h = (\rho^k_{1,h}, \cdots \rho^k_{N,h})$ we find $\rho^{k+1}_h = (\rho^{k+1}_{1,h}, \cdots \rho^{k+1}_{N,h})$ by:

- setting $\tilde{\rho}^{k+1}_{i,h} = X^k_{i,h}(h,.)_{\#}\rho^k_{i,h}$ where

$$\partial_t X^k_{i,h} = W_i[\rho^k_h] \circ X^k_{i,h},\ X^k_{i,h}(0,\cdot) = \mathrm{id},$$

- defining $\tilde{\rho}^{k+1}_h = (\tilde{\rho}^{k+1}_{1,h}, \cdots, \tilde{\rho}^{k+1}_{N,h})$, $\rho^{k+1}_h = (\rho^{k+1}_{1,h}, \cdots \rho^{k+1}_{N,h})$ is obtained by the semi-implicit JKO scheme:

$$\rho^{k+1}_{i,h} = \underset{\rho_i \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)}{\mathrm{argmin}} \left\{ W_2^2(\rho_i, \tilde{\rho}^{k+1}_{i,h}) + 2h \left(\mathcal{F}_i(\rho_i) + \int_\Omega V_i[\tilde{\rho}^{k+1}_h]\rho_i \right) \right\}.$$

Finally, let us say a few words on uniqueness which we have not addressed here, but which can be obtained at least in two ways: either by assuming some displacement semiconvexity of the internal energy and proving some exponential in time contraction estimate on the W_2 distance between two solutions (see section 4.7 or in [54, 73]), or by assuming some nondegeneracy of the diffusion and establishing some H^{-1} contraction estimate (see section 5.1.2 or in [35]).

5.3 Remarks on general costs

The first two chapters deal with gradient flows in Wasserstein space endowed with the Wasserstein distance of order 2. A natural extention is to study gradient flow in probability space endowed with the Wasserstein distance of order $p > 1$ or more generally an optimal transport distance with a general convex cost c. Several works have already considered this approach. The first one is due to Otto in [97], where he studied the doubly degenerate diffusion equation

$$\partial_t \rho - \mathrm{div}(|\nabla \rho^n|^{p-2} \nabla \rho^n)) = 0,$$

with $p > 1, n > 0$, as the steepest descent of the functional

$$\mathrm{sign}(m-1) \int_\Omega \rho^m,\ \text{with } m := n + \frac{p-2}{p-1},$$

with respect to the Wasserstein distance of order p^*, where p^* is the conjugate of p. In [1], Agueh generalized this result to more general costs. In this section, we give a slight perturbation of their result in order to adapt our splitting method to general costs. We will prove existence and uniqueness of solutions of nonlinear degenerate parabolic equation of the form

$$\partial_t \rho - \mathrm{div}(\rho \nabla c^*(\nabla F'(\rho))) - \mathrm{div}(\rho U[\rho]) = 0. \tag{5.98}$$

In the literature, other generalizations have been proposed. Dolbeault, Nazaret and Savaré introduced in [58] a new class of transport distances between probability measures which are constructed in a dynamical way as a generalization of the Benamou-Brenier formula. They introduced a mobility function $\boldsymbol{m} : \mathbb{R}^+ \to \mathbb{R}^+$ to define

$$W^2_{\boldsymbol{m},2}(\rho_0, \rho_1) := \inf \left\{ \int_0^1 \int_\Omega |v_t(x)|^2 \boldsymbol{m}(\rho_t(x))\, dxdt \ :\ \partial_t \rho + \mathrm{div}(\boldsymbol{m}(\rho)v) = 0,\ \rho_{t=0,1} = \mu_0, \mu_1 \right\}.$$

We remark that when $\boldsymbol{m}(x) = x$, we recover the Benamou-Brenier formula. We refer to [58, 39, 83, 84, 119] for the analysis of this distance and the study of usual gradient flow with respect to $W_{\boldsymbol{m},2}$.

One of the difficulties of problem (5.98) is in the passage to the limit in the very nonlinear term $\nabla c^*(\nabla F'(\rho))$. Indeed, following our previous method in section 5.2, we only have a narrow convergence for the gradient of the pressure associated to F, then stronger estimates will be needed. Actually, in the porous medium case and using the flow interchange argument we will see in chapter 6 and 7 that is possible to obtain a $L^2((0,T), H^1(\Omega))$ estimate for $F'(\rho) := \frac{m}{m-1}\rho^{m-1}$. In addition, since the vector field $U[\rho]$ is not assumed to be divergence free, we do not have conservation of the internal energy during the pure transport step, as in the previous section, and then, we have to find another way to control the energy dissipation during this phase. To deal with these issues, we assume that the initial condition is bounded from below and above and, in section 5.3.2, we start to prove a maximum/minimum principle which is a variant of the one proved by Otto in [97] and generalized to general costs by Agueh in [1]. In this section, we obtain several estimates useful in section 5.3.3 to prove theorem 5.17.

5.3.1 Assumptions and main result

Our goal is to solve (5.98) on Ω a smooth bounded convex subset of $\mathbb{R}^n$ with the usual Neumann boundary condition

$$\rho(\nabla c^*(\nabla F'(\rho)) + U[\rho]) \cdot \nu = 0 \text{ on } \partial\Omega,$$

where ν denotes the outer unit normal to $\partial\Omega$.

Assumption on the cost c: We assume that $c : \mathbb{R}^n \to [0, +\infty)$ is a smooth strictly convex function which satisfies $0 = c(0) < c(x)$, for all $x \neq 0$, and there exists $\alpha, \beta > 0$, $q > 1$ such that

$$\alpha|x|^q \leqslant c(x) \leqslant \beta(1 + |x|^q), \tag{5.99}$$

for all $x \in \mathbb{R}^n$. We denote c^* the Legendre transform of the function c, i.e for all $z \in \mathbb{R}^n$,

$$c^*(z) := \sup_{x \in \mathbb{R}^n} x \cdot z - c(x).$$

As in [1], we denote c_h the function define by

$$c_h(x) := c\left(\frac{x}{h}\right).$$

Assumption on the internal energy F: We assume that $F : [0, +\infty) \to \mathbb{R}$ is a continuous strictly convex superlinear function of class $\mathcal{C}^2((0, +\infty))$ which satisfies $F(0) = 0$ and

$$\text{The map } x \in (0, +\infty) \mapsto x^n F'(x^{-n}) \text{ is convex and nonincreasing.} \tag{5.100}$$

As usual, typical examples are functions of the form $F(\rho) = \rho^m$ $(m > 1)$.

Assumption on the drift $U[\rho]$:

- There exists $C > 0$ such that for all $\rho \in \mathcal{P}(\Omega)$,

$$U[\rho] \in W^{1,\infty}(\mathbb{R}^n),\ U[\rho] \cdot \nu = 0 \text{ on } \partial\Omega \text{ and } |U[\rho]| \leqslant C(1 + |x|), \text{ for all } x \in \mathbb{R}^n, \tag{5.101}$$

- There exists $C > 0$ such that for all $\rho, \mu \in \mathcal{P}(\Omega)$,

$$\int_\Omega |U[\rho] - U[\mu]|^p \, d\rho \leqslant C W_p^p(\rho, \mu), \tag{5.102}$$

for some $p > 1$.

A weak solution of (5.98) with Neumann boundary condition is a curve $\rho : t \in (0,+\infty) \mapsto \mathcal{P}^{ac}(\Omega)$ such that $\rho + \frac{1}{\rho} \in L^\infty((0,T)\times\Omega)$, $\nabla F'(\rho) \in L^{q^*}((0,T)\times\Omega)$, $\nabla c^*(\nabla F'(\rho)) \in L^q((0,T)\times\Omega)$ and for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}\times\mathbb{R}^n)$,

$$\int_{\mathbb{R}\times\mathbb{R}^n} (\rho\partial_t\varphi - \rho\,(\nabla c^*(\nabla(F'(\rho))) - U[\rho])\cdot\nabla\varphi) = -\int_\Omega \rho_0(x)\varphi(0,x)\,dx.$$

Then our main result is the following:

Theorem 5.17. *Assume that $\rho \in \mathcal{P}^{ac}(\Omega)$ satisfies*

$$0 < m \leqslant \rho_0 \leqslant M < +\infty,$$

then (5.98) *admits at least one weak solution.*

The proof of this theorem, given in the next sections, is based on the same kind of splitting scheme using before. We combine pure transport steps by the drift U and Wasserstein gradient flow steps to handle the diffusive term. Given a time step $h > 0$, we construct by induction two sequences ρ_h^k and $\tilde\rho_h^k$ in $\mathcal{P}^{ac}(\Omega)$ by setting $\rho_h^0 = \tilde\rho_h^0 = \rho_0$ and using the following scheme:

$$\left\{\begin{array}{l} \tilde\rho_h^{k+1} = X_h^k(h,\cdot)_\#\rho_h^k, \\ \rho_h^{k+1} = \mathrm{argmin}_{\rho\in\mathcal{P}(\Omega)}\, hW_{c_h}(\tilde\rho_h^{k+1},\rho) + \mathcal{F}(\rho), \end{array}\right. \tag{5.103}$$

where X_h^k is the solution of

$$\left\{\begin{array}{l} \partial_t X_h^k = U[\rho_h^k]\circ X_h^k, \\ X_h^k(0,\cdot) = Id. \end{array}\right. \tag{5.104}$$

Since we assume that $U[\rho]$ satisfies (5.101), the DiPerna-Lions theory implies that X_h^k is well defined and there exists a constant $C > 0$ such that for all $t \in [0,h]$ and $x \in \mathbb{R}^n$,

$$|X_h^k(t,x) - x| \leqslant Ch \text{ and } |\det(\nabla X_h^k(t,x)) - 1| \leqslant Ch. \tag{5.105}$$

We define two interpolations: the piecewise constant interpolation ρ_h defined by

$$\rho_h(t) := \rho_h^{k+1} \text{ if } t \in (hk, h(k+1)],$$

and the *continuous* interpolation during the pure transport phase

$$\tilde\rho_h(t) := X_h^k(t,\cdot)_\#\rho_h^k \text{ if } t \in (0,h].$$

5.3.2 Maximum/minimum principle and standard estimates

We start this section by proving that sequences are bounded from below and above for all $T < +\infty$.

Proposition 5.18. *Let $T < +\infty$ be a fixed time. There exists $C > 0$ such that, for all $k \geqslant 0$ and h such that $kh \leqslant T$,*

$$me^{-CT} \leqslant \tilde\rho_h^k, \rho_h^k \leqslant Me^{CT}.$$

Proof. By definition of $\tilde\rho_h^1$, for all $x \in \Omega$,

$$\rho_0(x) = \tilde\rho_h^1(X_h^0(h,x))\det(\nabla X_h^0(h,x)).$$

Then using (5.105), for h small enough, we have

$$\frac{m}{1+Ch} \leqslant \tilde\rho_h^1(X_h^0(h,x)) \leqslant \frac{M}{1-Ch},$$

which implies that

$$\frac{m}{1+Ch} \leqslant \tilde\rho_h^1(x) \leqslant \frac{M}{1-Ch}.$$

Now we can apply the maximum/minimum principle developped by Otto in [97] and generalized to general costs by Agueh in [1] to obtain

$$\frac{m}{1+Ch} \leqslant \rho_h^1(x) \leqslant \frac{M}{1-Ch}.$$

Then by induction, the following holds

$$\frac{m}{(1+Ch)^k} \leqslant \tilde{\rho}_h^k(x), \rho_h^k(x) \leqslant \frac{M}{(1-Ch)^k},$$

which gives the desired result.

□

In the next proposition we recover usual estimates from the gradient flow theory

Proposition 5.19. *For all T and for all h, k such that $hk \leqslant T$, we have*

$$\mathcal{F}(\rho_h^k) \leqslant CT + \mathcal{F}(\rho_0), \tag{5.106}$$

$$h \sum_{k=0}^{N-1} W_{c_h}(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant C(1+T). \tag{5.107}$$

Proof. To find the first estimate, we use the fact that ρ_h^k is optimal is the minimization problem. Since ρ_h^{k-1} is a competitor we have

$$\mathcal{F}(\rho_h^k) \leqslant hW_{c_h}(\rho_h^k, \tilde{\rho}_h^k) + \mathcal{F}(\rho_h^k) \leqslant hW_{c_h}(\rho_h^{k-1}, \tilde{\rho}_h^k) + \mathcal{F}(\rho_h^{k-1}).$$

But using (5.105),

$$W_{c_h}(\rho_h^{k-1}, \tilde{\rho}_h^k) \leqslant C,$$

then

$$\mathcal{F}(\rho_h^k) \leqslant Ch + \mathcal{F}(\rho_h^{k-1}),$$

and by induction we have (5.106). Now we want to prove (5.107). First we remark that using minimization scheme we have

$$\begin{aligned} h \sum_{k=0}^{N-1} W_{c_h}(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) &\leqslant \sum_{k=0}^{N-1} (\mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^{k+1})) \\ &\leqslant \sum_{k=0}^{N-1} (\mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^k)) + \mathcal{F}(\rho_0) - \mathcal{F}(\rho_h^N). \end{aligned}$$

Using the Monge-Ampère equation, we find

$$\begin{aligned} \mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^k) &\leqslant \int_\Omega F\left(\frac{\rho_h^k(x)}{\det(\nabla X_h^k(h,x))}\right) \det(\nabla X_h^k(h,x)) - \int_\Omega F(\rho_h^k) \\ &\leqslant \int_\Omega \left(F\left(\frac{\rho_h^k(x)}{\det(\nabla X_h^k(h,x))}\right) - F(\rho_h^k)\right) \det(\nabla X_h^k(h,x)) - \int_\Omega F(\rho_h^k)(\det(\nabla X_h^k(h,x)) - 1) \\ &\leqslant I + J. \end{aligned}$$

For J, since $|\det(\nabla X_h^k(h,x)) - 1| \leqslant Ch$, we have

$$J \leqslant Ch \int_\Omega |F(\rho_h^k)| \leqslant Ch,$$

because $\mathcal{F}(\rho_h^k)$ is bounded above and below. And for I we will use that for all k, we have

$$\frac{m}{(1+Ch)^k} \leqslant \rho_h^k \leqslant \frac{M}{(1-Ch)^k},$$

according to the proposition 5.18 and since

$$\frac{1}{1+Ch} \leqslant \frac{1}{\det(\nabla X_h^k(h,x))} \leqslant \frac{1}{1-Ch},$$

we have that $\frac{\rho_h^k}{\det(\nabla X_h^k(h,x))}$ and ρ_h^k are in $[e^{-CT}m, e^{CT}M]$ and since F is $\mathcal{C}^2$ it is locally lipschitz and we get

$$\begin{aligned} I &\leqslant \int_\Omega \left| F\left(\frac{\rho_h^k(x)}{\det(\nabla X_h^k(h,x))}\right) - F(\rho_h^k)\right| |\det(\nabla X_h^k(h,x))| && (5.108)\\ &\leqslant C\int_\Omega \left|\frac{1}{\det(\nabla X_h^k(h,x))} - 1\right| |\det(\nabla X_h^k(h,x))|\rho_h^k(x) && (5.109)\\ &\leqslant C\int_\Omega |\det(\nabla X_h^k(h,x)) - 1|\rho_h^k(x) && (5.110)\\ &\leqslant Ch. && (5.111) \end{aligned}$$

So if we combine I and J we obtain $\mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^k) \leqslant Ch$ and the conclusion follows. □

A direct consequence is

Corollary 5.20. *For all T and for all h, k such that $hk \leqslant T$,*

$$\sum_{k=0}^{N-1} W_q^q(\rho_h^{k+1}, \rho_h^k) \leqslant Ch^{q-1}.$$

Proof. First combining (5.99) and (5.107), we have

$$\sum_{k=0}^{N-1} W_q^q(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant C\sum_{k=0}^{N-1} h^q W_{c_h}(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant Ch^{q-1}. \tag{5.112}$$

By definition of $\tilde{\rho}_h^{k+1}$ and (5.105), we obtain

$$\sum_{k=0}^{N-1} W_q^q(\rho_h^k, \tilde{\rho}_h^{k+1}) \leqslant \sum_{k=0}^{N-1}\int_\Omega |X_h^k(h,x) - x|^q \rho_h^k(x)\,dx \leqslant Ch^{q-1}. \tag{5.113}$$

Then the result is proved combining (5.112) and (5.113).

□

Remark 5.21. *We directly obtain that*

$$\sum_{k=0}^{N-1} W_q^q(\tilde{\rho}_h^{k+1}, \tilde{\rho}_h^k) \leqslant Ch^{q-1},$$

and

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant Ch^\alpha,$$

with

$$\alpha = \begin{cases} q-1 & \text{if } q \leqslant 2,\\ 1 & \text{if } q > 2. \end{cases}$$

In the next proposition, we establish the Euler-Lagrange equation for JKO step. We only state the result and we refer to proposition 2.6 from [1] for the proof.

Proposition 5.22. *For all $k \geqslant 0$, ρ_h^{k+1} satisfies*

$$\int_{\Omega\times\Omega} \nabla c\left(\frac{x-y}{h}\right)\cdot\psi(y)\,d\gamma_h^{k+1}(x,y) + \int_\Omega P(\rho_h^{k+1}(y))\,\mathrm{div}(\psi(y))\,dy = 0,$$

for all $\psi \in \mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$, with $P(x) := xF'(x) - F(x)$ is the pressure associated to F and γ_h^{k+1} is the W_{c_h}-optimal transport plan between $\tilde{\rho}_h^{k+1}$ and ρ_h^{k+1}. Moreover,

1. $P(\rho_h^{k+1}) \in W^{1,\infty}(\Omega)$,

2. *If* T_h^{k+1} *is the* W_{c_h}*-optimal transport map between* ρ_h^{k+1} *and* $\tilde{\rho}_h^{k+1}$*, then*

$$\frac{T_h^{k+1}(y) - y}{h} = \nabla c^*(\nabla F'(\rho_h^{k+1})) \text{ a.e in } \Omega. \tag{5.114}$$

We immediately deduce estimates on $\nabla c^*(\nabla F'(\rho_h))$ and $\nabla F'(\rho_h)$.

Corollary 5.23. *For all* $T > 0$, $\nabla c^*(\nabla F'(\rho_h))$ *is bounded in* $L^q((0,T)\times\Omega)$ *and* $F'(\rho_h)$ *is bounded in* $L^{q^*}((0,T)\times\Omega)$.

Proof. Using (5.114) and proposition 5.18, we obtain

$$\begin{aligned}\int_\Omega |\nabla c^*(\nabla F'(\rho_h^{k+1}))|^q &\leqslant \frac{1}{me^{-CT}} \int_\Omega \left|\frac{T_h^{k+1}(y)-y}{h}\right|^q \rho_h^{k+1}(y)\,dy \\ &\leqslant \frac{1}{\alpha m e^{-CT}} W_{c_h}(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}),\end{aligned}$$

when we used assumption (5.99). Then summing over k and (5.107), we obtain

$$\int_0^T \int_\Omega |\nabla c^*(\nabla F'(\rho_h^{k+1}))|^q \leqslant Ch \sum_{N=0}^{N-1} W_{c_h}(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant C.$$

To prove that $F'(\rho_h)$ is bounded in $L^{q^*}((0,T)\times\Omega)$, we just have to remark that using (5.99), we have

$$C|z|^{q^*} \leqslant \beta + c^*(z) \leqslant \nabla c^*(z)\cdot z \leqslant \frac{1}{\varepsilon q}|\nabla c^*(z)|^q + \frac{\varepsilon}{q^*}|z|^{q^*},$$

which concludes the proof for ε small enough. □

5.3.3 Convergences and conclusion

Convergences Corollary 5.20 and remark 5.21 give $\frac{1}{q^*}$-Hölder estimates on ρ_h and $\tilde{\rho}_h$ then using a refined version of Ascoli theorem, we have that ρ_h and $\tilde{\rho}_h$ converge in $L^\infty((0,T), W_q)$ to the same limit ρ.

Moreover, using corollary 5.20, proposition 5.19 and Rossi-Savaré theorem 5.12, we obtain that ρ_h strongly converges in $L^1((0,T)\times\Omega)$ to ρ. Then since ρ_h is bounded in $L^\infty((0,T)\times\Omega)$, ρ_h converges strongly in $L^p((0,T)\times\Omega)$ for all $p \in [1,+\infty)$, $F'(\rho_h)$ converges strongly in $L^p((0,T)\times\Omega)$ for all $p \in [1,+\infty)$ to $F'(\rho)$. In addition, proposition 5.19 gives that $\nabla F'(\rho_h)$ converges weakly to $\nabla F'(\rho)$ in $L^{q^*}((0,T)\times\Omega)$ and $\nabla c^*(\nabla F'(\rho_h))$ converges weakly to σ in $L^q((0,T)\times\Omega)$.

The next theorem proves that $\sigma = \nabla c^*(\nabla F'(\rho))$ and is due to Agueh in [1].

Theorem 5.24 (Theorem 3.10 from [1])**.** *Assume that* F *satisfies* (5.100)*, then for all nonnegative function* $u \in \mathcal{C}_c^2(\mathbb{R})$,

$$\lim_{h\searrow 0} \int_{\mathbb{R}\times\Omega} \rho_h \nabla c^*(\nabla F'(\rho_h))\cdot\nabla F'(\rho_h) u(t)\,dxdt = \int_{\mathbb{R}\times\Omega} \rho \nabla c^*(\nabla F'(\rho))\cdot\nabla F'(\rho) u(t).$$

Therefore, $(\mathrm{div}(\rho_h \nabla c^*(\nabla F'(\rho_h))))$ *converges weakly to* $\mathrm{div}(\rho\sigma)$ *in* $\mathcal{C}_c^2(\mathbb{R}\times\mathbb{R}^n)'$ *and*

$$\mathrm{div}(\rho\sigma) = \mathrm{div}(\rho\nabla c^*(\nabla F'(\rho))).$$

Proof. The proof is based on an argument of geodesic convexity and the Minty's trick using the convexity of c^*. We refer to [1] for more details. □

Conclusion For all $\varphi \in \mathcal{C}_c^\infty([0,+\infty)\times\Omega)$, with $\operatorname{supp}\varphi(\cdot,x)\subset[0,T)$,

$$\begin{aligned}\int_{hk}^{h(k+1)}\int_\Omega \tilde\rho_h(t,x)(\partial_t\varphi - U[\rho_h^k]\cdot\nabla\varphi)(t,x)\,dxdt&\\ &= \int_\Omega \tilde\rho_h^{k+1}(x)\int_{hk}^{h(k+1)}\frac{d}{dt}[\varphi(t,X_h^k(t,x))]\,dtdx\\ &= \int_\Omega \rho_h^k(x)\left(\varphi(h(k+1),X_h^k(h(k+1),x)) - \varphi(hk,x)\right)\,dx.\end{aligned}$$

Summing over k, $(N := \lceil\frac{T}{h}\rceil)$,

$$\begin{aligned}\int_0^T\int_\Omega \tilde\rho_h(t,x)(\partial_t\varphi(t,x) - U[\rho_h(t-h)](x)\cdot\nabla\varphi(t,x))\,dxdt&\\ &= \sum_{k=0}^{N-1}\int_{hk}^{h(k+1)}\int_\Omega \tilde\rho_h(t,x)(\partial_t\varphi - U[\rho_h^k]\cdot\nabla\varphi)(t,x)\,dxdt\\ &= \sum_{k=0}^{N-1}\int_\Omega \rho_h^k(x)\left(\varphi(h(k+1),X_h^k(h(k+1),x)) - \varphi(hk,x)\right)\,dx\\ &= \sum_{k=0}^{N-1}\int_\Omega \tilde\rho_h^{k+1}(x)\,dx\varphi(h(k+1),x)\,dx - \sum_{k=0}^{N-1}\int_\Omega \rho_h^k(x)\,dx\varphi(hk,x)\,dx\\ &= \sum_{k=0}^{N-1}\int_\Omega (\tilde\rho_h^{k+1}(x) - \rho_h^{k+1}(x))\varphi(hk,x)\,dx - \int_\Omega \rho_0(x)\varphi(0,x)\,dx.\end{aligned}$$

Combining with (5.22), we conclude that

$$\begin{aligned}\int_0^T\int_\Omega \tilde\rho_h(t,x)(\partial_t\varphi(t,x) - U[\rho_h(t-h)](x)\cdot\nabla\varphi(t,x))\,dxdt&\\ &= \sum_{k=0}^{N-1}\int_\Omega \nabla c^*(\nabla F'(\rho_h^{k+1}(x)))\cdot\nabla\varphi(hk,x)\,dx - \int_\Omega \rho_0(x)\varphi(0,x)\,dx\\ &+ \sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\varphi(hk,\cdot)](x,y)\,d\gamma_h^{k+1}(x,y),\end{aligned}$$

where

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\Omega)}|x-y|^2,$$

and γ_h^{k+1} is the optimal transport plan in $\Pi(\rho_h^{k+1},\tilde\rho_h^{k+1})$.

Remark 5.21 implies directly that

$$\sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\varphi(hk,\cdot)](x,y)\,d\gamma_h^{k+1}(x,y)\to 0,$$

as $h\searrow 0$. Then we can pass to the limit in the discreet equation and theorem 5.17 follows.

Chapter 6

Systems with cross-diffusion

The modelling of crowd behaviour has become a very active field of applied mathematics in recent years. These models permit to understand many phenomena as cell migration, tumor growth, etc. Several models already exist to tackle this problem. The first one, microscopic, consists to seeing a population as a high number of individuals which satisfy ODEs, see for instance [92] and the second is macroscopic and consists in describing a population by a density ρ satisfying one PDE, where $\rho(t,x)$ represents the number of individuals in x at time t. In this last framework, different methods to handle the congestion effect have been proposed. The first one consists in saying that the motion has to be slower when the density is very high, see for example [47, 46, 45] for a different approach with applications to crowd dynamics. Another way to modelling the congestion effect is to use a threshold: the density evolves as we would expect until it touches a maximal level and then the motion has to be adapted in these regions, see for example [89] for crowd motion model and [90] for application to dendritic growth. In [91], we can see a comparison between microscopic and macroscopic models. Recently in [94], Meszaros and Santambrogio proposed a model of hard congestion where individuals are subject to a Brownian diffusion. This corresponds to modified Fokker-Planck equation with a constraint on the density.

Since in macroscopic models, we have mass conservation, the theory of optimal transportation is a very natural tool to attack it. In [89], the authors investigated a model of room evacuation. They showed that if the velocity field is given by a gradient, say $V = \nabla D$, where D is the distance to a given target, then the problem has a gradient flow structure in the Wasserstein space and the velocity field has to be adapted by a pressure field to handle congestion effect. More recently in [94], a splitting scheme has been introduced to handle velocity fields which are -not necessarily potential- vector field. The scheme consists in combining steps where the density follows Fokker-Planck equation and Wasserstein projections over the set of densities which cannot exceed 1.

A natural variant of the model of [94], consists in considering two (or more) populations, each of whom having its own potential but coupled through the constraint that the total density cannot exceed 1 and then subject to a common pressure field. Note that variant problems with total density *equal* to 1 were treated in [48, 7, 12] and for more general cross-diffusion systems, we refer, for example, to [79, 53, 66, 67, 72]. For a linear diffusion (corresponding to a Brownian noise on each species), the two-species crowd dynamic is expressed by the PDEs

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \operatorname{div}(\rho_1(\nabla V_1 + \nabla p)) = 0, \\ \partial_t \rho_2 - \Delta \rho_2 - \operatorname{div}(\rho_2(\nabla V_2 + \nabla p)) = 0, \\ p \geq 0, \rho_1 + \rho_2 \leq 1,\ p(1 - \rho_1 - \rho_2) = 0. \end{cases} \tag{6.1}$$

In this chapter, we prove that this system is the gradient flow (for the product Wasserstein distance) of the energy

$$\mathcal{E}(\rho_1, \rho_2) := \sum_{i=1}^{2} \int_\Omega (\rho_i \log(\rho_i) + V_i \rho_i) + \int_\Omega \chi_{[0,1]}(\rho_1(x) + \rho_2(x)) \mathrm{d}x,$$

and for a more general energy of the form

$$\mathcal{E}(\rho_1,\rho_2) := \mathcal{E}_1(\rho_1) + \mathcal{E}_2(\rho_2) + \mathcal{F}_m(\rho_1+\rho_2)$$

with

$$\mathcal{E}_i(\rho_i) := \int_\Omega (\rho_i \log(\rho_i) + V_i\rho_i),\ \mathcal{F}_m(\rho) := \int_\Omega F_m(\rho(x))\mathrm{d}x,$$

with F_m define as in (6.4). The gradient flow of $\mathcal{E}$ is the following system with a coupling in the diffusion:

$$\partial_t\rho_i = \Delta\rho_i + \mathrm{div}(\rho_i\nabla(V_i + F_m'(\rho_1+\rho_2))),\ i=1,\ 2. \tag{6.2}$$

The JKO scheme for this energy then reads

$$(\rho_1^{k+1},\rho_2^{k+1}) = \underset{(\rho_1,\rho_2)}{\mathrm{argmin}} \Big\{ \sum_{i=1}^{2} \frac{1}{2h} W_2^2(\rho_i,\rho_i^k) + \mathcal{E}(\rho_1,\rho_2) \Big\} \tag{6.3}$$

which, in the particular case of the linear diffusion crowd motion problem with two species, takes the form

$$(\rho_1^{k+1},\rho_2^{k+1}) = \underset{\rho_1+\rho_2\le 1}{\mathrm{argmin}} \Big\{ \sum_{i=1}^{2} \left(\frac{1}{2h} W_2^2(\rho_i,\rho_i^k) + \int_\Omega (\rho_i\log(\rho_i) + V_i\rho_i) \right) \Big\}.$$

This chapter is composed of five sections. In section 6.1, we introduce our assumptions and we state our main results. In section 6.2, we prove existence of solution of system (6.2). The key ingredient is the flow interchange argument (see [86, 55, 73] for example or chapter 4) which gives an estimate on the gradient of $\rho_1+\rho_2$ and on the gradient of ρ_i. Section 6.3 gives the proof of existence of solution for system with hard congestion (6.1). In this section we use again the flow interchange argument to obtain stronger estimates. In section 6.4, we focus on the particular case where $V_1 = V_2$. In this case, we state the convergence when $m \to +\infty$ of solution of (6.2) to solution of (6.1) and we prove a L^1-contraction theorem. In the final section 6.5, we present numerical simulations.

6.1 Assumptions and main results:

Let Ω a convex bounded subset of $\mathbb{R}^n$. For $i \in \{1,2\}$, we define $\mathcal{V}_i : \mathcal{P}(\Omega) \to \mathbb{R}$ the potential energy associated to $V_i \in W^{1,\infty}(\Omega)$ by

$$\mathcal{V}_i(\rho) := \int_\Omega V_i(x)\, d\rho(x).$$

Let $m \in [1,+\infty)$, we define $F_m : \mathbb{R}^+ \to \mathbb{R}$ by

$$F_m(x) = \begin{cases} x\log x & \text{if } m=1, \\ \frac{x^m}{m-1} & \text{if } m>1. \end{cases} \tag{6.4}$$

We remark that F_m is a convex superlinear continuous function. We denote $\mathcal{F}_m : \mathcal{M}(\Omega) \to \mathbb{R}$ the internal energy defined by

$$\mathcal{F}_m(\rho) := \begin{cases} \int_\Omega F_m(\rho(x))\, dx & \text{if } \rho \ll \mathcal{L}_{|\Omega}, \\ +\infty & \text{otherwise,} \end{cases}$$

and the functional $\mathcal{F}_\infty$ is defined by

$$\mathcal{F}_\infty(\rho) := \begin{cases} 0 & \text{if } \|\rho\|_\infty \leqslant 1, \\ +\infty & \text{otherwise.} \end{cases}$$

In the following, we focus on the case where the individual diffusion is linear but the proof can be adpated to deal with individual nonlinear diffusions.

We say that $(\rho_1, \rho_2) : [0, +\infty) \to \mathcal{P}^{\rm ac}(\Omega)^2$ is a weak solution of (6.2) if for all $i \in \{1, 2\}$ and for all $T < +\infty$, $\rho_i \in \mathcal{C}([0, T], \mathcal{P}^{\rm ac}(\Omega))$, $\rho_i \in L^\infty([0, T], BV(\Omega))$, $\rho_i \nabla F_m'(\rho_1 + \rho_2) \in L^1((0, T) \times \Omega)$ and for all $\phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$,

$$\int_0^{+\infty} \int_\Omega \left([\rho_i \partial_t \phi - (\rho_i \nabla V_i + \rho_i \nabla F_m'(\rho_1 + \rho_2)) \cdot D\phi] \, dx - \int_\Omega \nabla \phi \cdot \, d\nabla \rho_i(x) \right) = - \int_\Omega \phi(0, x) \rho_{i,0}(x) \, dx.$$

and we say that $(\rho_1, \rho_2, p) : [0, +\infty) \to \mathcal{P}^{\rm ac}(\Omega)^2 \times H^1(\Omega)$ is a weak solution of (6.1) if for all $i \in \{1, 2\}$ and for all $T < +\infty$, $\rho_i \in \mathcal{C}([0, T], \mathcal{P}^{\rm ac}(\Omega))$, $\rho_i \in L^\infty([0, T], BV(\Omega))$, $p \in L^2([0, T], H^1(\Omega))$ with $p \geqslant 0$, $\rho_1 + \rho_2 \leqslant 1$ and $p(1 - \rho_1 - \rho_2) = 0$ a.e. In addition, for all $\phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$,

$$\int_0^{+\infty} \int_\Omega \left([\rho_i \partial_t \phi - (\rho_i \nabla V_i + \rho_i \nabla p) \cdot \nabla \phi] \, dx - \int_\Omega \nabla \phi \cdot \, dD\rho_i(x) \right) = - \int_\Omega \phi(0, x) \rho_{i,0}(x) \, dx.$$

Main results of this chapter are

Theorem 6.1. *Assume that* $(\rho_{1,0}, \rho_{2,0}) \in \mathcal{P}^{\rm ac}(\Omega)^2$ *satisfy*

$$\mathcal{F}_1(\rho_{1,0}) + \mathcal{F}_1(\rho_{2,0}) + \mathcal{F}_m(\rho_{1,0} + \rho_{2,0}) < +\infty, \tag{6.5}$$

then (6.2) *admits at least one weak solution.*

and

Theorem 6.2. *Assume that* $|\Omega| > 2$. *If* $(\rho_{1,0}, \rho_{2,0}) \in \mathcal{K} := \left\{ (\rho_1, \rho_2) \in \mathcal{P}^{\rm ac}(\Omega)^2 \, : \, \rho_1 + \rho_2 \leqslant 1 \right\}$ *satisfies*

$$\mathcal{F}_1(\rho_{1,0}) + \mathcal{F}_1(\rho_{2,0}) < +\infty,$$

then there exists at least one weak solution of (6.1).

We assume that $|\Omega| > 2$ to be sure that $\mathcal{K}$ is not empty or trivial.

Remarks on possible extensions:

1. These models can be generalized to more than two species. Moreover, instead of assumed that all the densities take the same space, we can generalize to densities evolving under the constraints on $\alpha_1 \rho_1 + \alpha_2 \rho_2$. Then system (6.2) becomes

   $$\partial_t \rho_i = \operatorname{div}(\rho_i \nabla V_i) + \Delta \rho_i + \alpha_i \operatorname{div}(\rho_i \nabla F_m'(\alpha_1 \rho_1 + \alpha_2 \rho_2)), \ i = 1, \ 2.$$

   and system with hard congestion becomes

   $$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \operatorname{div}(\rho_1(\nabla V_1 + \nabla p)) = 0, \\ \partial_t \rho_2 - \Delta \rho_2 - \operatorname{div}(\rho_2(\nabla V_2 + \nabla p)) = 0, \\ p \geq 0, \alpha_1 \rho_1 + \alpha_2 \rho_2 \leq 1, \ p(1 - \alpha_1 \rho_1 - \alpha_2 \rho_2) = 0. \end{cases}$$

2. We can deal with more general velocities. Indeed, using the semi-implicit scheme introduced in chapter 4 or the splitting method introduced in chapter 5, we can treat vector fields which depend of the densities and which come not necessarily from a potential. These extensions allow to treat nonlocal interactions between different species which are subject to a common congection effect.

6.2 Entropy of the sum

In this section, we prove theorem 6.1 using the implicit JKO scheme, first introduced by Jordan, Kinderlherer and Otto in [65]. Given a time step $h > 0$, we construct by induction two sequences $\rho_{1,h}^k$ and $\rho_{2,h}^k$ with the following scheme: $\rho_{i,h}^0 = \rho_{i,0}$ and for all $k \geqslant 0$,

$$(\rho_{1,h}^{k+1}, \rho_{2,h}^{k+1}) \in \underset{(\rho_1, \rho_2) \in \mathcal{P}^{\rm ac}(\Omega)^2}{\operatorname{argmin}} \left\{ \sum_{i=1}^{2} \left(W_2^2(\rho_i, \rho_{i,h}^k) + 2h \left(\mathcal{F}_1(\rho_i) + \mathcal{V}_i(\rho_i) \right) \right) + 2h \mathcal{F}_m(\alpha_1 \rho_1 + \alpha_2 \rho_2) \right\}. \tag{6.6}$$

These sequences are well-defined by compactness and l.s.c standard argument. Then we define the piecewise constant interpolations $\rho_{i,h} : \mathbb{R}^+ \to \mathcal{P}^{\rm ac}(\Omega)$ by

$$\rho_{i,h}(t) := \rho_{i,h}^{k+1}, \qquad \text{if } t \in (kh, (k+1)h].$$

In the first part of this section, we prove the convergence of these sequences and then we give the proof of theorem 6.1.

6.2.1 Estimates and convergences

In this section, we present basic estimates in JKO scheme ([65]) and a stronger one using the flow interchange argument ([86, 55]).

First, using the minimization scheme, we prove

Proposition 6.3. *For all $T < +\infty$ and for all $i \in [\![1, 2]\!]$, there exists a constant $C < +\infty$ such that for all $k \in \mathbb{N}$ and for all h with $kh \leqslant T$ and let $N = \lceil \frac{T}{h} \rceil$, we have*

$$\mathcal{F}_1(\rho_{i,h}^k) \leqslant C, \tag{6.7}$$

$$\mathcal{F}_m(\rho_{1,h}^k + \rho_{2,h}^k) \leqslant C, \tag{6.8}$$

$$\textstyle\sum_{k=0}^{N-1} W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) \leqslant Ch. \tag{6.9}$$

Proof. These results are obtained easily taking $\rho_i = \rho_{i,h}^k$ as competitors in (6.6), see [65]. □

Using the refined version of Ascoli-Arzelà's theorem (see [4], proposition 3.3.1), $\rho_{i,h}$ converges to $\rho_i \in \mathcal{C}^{1/2}([0,T], \mathcal{P}^{\rm ac}(\Omega))$ in $L^\infty([0,T], \mathcal{P}^{\rm ac}(\Omega))$.

Remark 6.4. *We notice that estimate* (6.9) *does not depend of m, then the refined version of Ascoli-Arzelà's theorem gives*

$$W_2(\rho_i(t), \rho_i(s)) \leqslant C|t-s|^{1/2},$$

for all $t, s \leqslant T$ and C which does not depend of m.

But to pass to the limit in the nonlinear diffusive term we have to obtain a strong convergence so we have to find stronger estimates.

Proposition 6.5. *There exists a constant $C > 0$ such that, for all $T > 0$,*

$$\|\rho_{1,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))} + \|\rho_{2,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))} + \|(\rho_{1,h} + \rho_{2,h})^{m/2}\|_{L^2((0,T),H^1(\Omega))} \leqslant C(1+T). \tag{6.10}$$

Proof. We use the flow interchange argument, introduced in [86], to find a stronger estimate as in [55, 73]. In other words, we perturb $\rho_{1,h}^k$ and $\rho_{2,h}^k$ by the heat flow. Let η_i be the solution of

$$\begin{cases} \partial_t \eta_i = \Delta \eta_i & \text{in } (0,T)\times\Omega, \\ \nabla \eta_i \cdot \nu = 0 & \text{in } (0,T)\times\partial\Omega, \\ \eta_{i|t=0} = \rho_{i,h}^k. \end{cases} \tag{6.11}$$

The entropy is geodesically convex then the heat flow is a 0-flow of the Entropy $\mathcal{F}_1$, [65, 116, 4, 49, 112], i.e

$$\frac{1}{2}\frac{d^+}{d\sigma}_{|\sigma=s} W_2^2(\eta(s), \rho) \leqslant \mathcal{F}_1(\rho) - \mathcal{F}_1(\eta(s)), \tag{6.12}$$

for all $s > 0$ and $\rho \in \mathcal{P}_2^{\rm ac}(\Omega)$, where

$$\frac{d^+}{dt} f(t) := \limsup_{s \to 0^+} \frac{f(t+s) - f(t)}{s}.$$

Moreover, using the scheme (6.6), we get

$$\begin{aligned}\sum_{i=1}^{2}\frac{1}{2}\frac{d^+}{d\sigma}_{|\sigma=s} W_2^2(\eta_i(s),\rho_{i,h}^{k-1}) \\ \geqslant -2h\frac{d^+}{d\sigma}_{|\sigma=s}\left(\sum_{i=1}^{2}\frac{1}{2}\left(\mathcal{F}_1(\eta_i(s))+\mathcal{V}_i(\eta_i(s))\right)+\mathcal{F}_m(\eta_1(s)+\eta_2(s))\right)\end{aligned} \tag{6.13}$$

Since η_i is smooth, we can easily compute for $s>0$

$$\begin{aligned}&\partial_s\Big(\sum_{i=1}^{2}\frac{1}{2}\big(\mathcal{F}_1(\eta_i(s))+\mathcal{V}_i(\eta_i(s))\big)+\mathcal{F}_m(\eta_1(s)+\eta_2(s))\Big)\\ &\quad=\sum_{i=1}^{2}\left(\int_\Omega \Delta\eta_i(s)((1+\log(\eta_i))+V_i\right)+\int_\Omega\Delta(\eta_1+\eta_2)F_m'(\eta_1(s)+\eta_2(s)) \\ &\quad=-\sum_{i=1}^{2}\left(\int_\Omega\frac{|\nabla\eta_i(s)|^2}{\eta_i(s)}+\int_\Omega\nabla V_i\cdot\nabla\eta_i\right)-\int_\Omega|\nabla(\eta_1(s)+\eta_2(s))|^2F_m''(\eta_1(s)+\eta_2(s)),\end{aligned} \tag{6.14}$$

But using Young's inequality,

$$\begin{aligned}-\int_\Omega\nabla V_i\cdot\nabla\eta_i &\leqslant \int_\Omega|\nabla V_i||\nabla\eta_i(s)| \\ &\leqslant \int_\Omega|\nabla V_i|\eta_i^{1/2}(s)\frac{|\nabla\eta_i(s)|}{\eta_i^{1/2}(s)} \\ &\leqslant \frac{1}{2}\int_\Omega|\nabla V_i|^2\eta_i(s)+\frac{1}{2}\frac{|\nabla\eta_i(s)|^2}{\eta_i(s)}\end{aligned}$$

Then, we have

$$\begin{aligned}&\partial_s\Big(\sum_{i=1}^{2}\frac{1}{2}\left(\mathcal{F}_1(\eta_i(s))+\mathcal{V}_i(\eta_i(s))\right)+\mathcal{F}_m(\eta_1(s)+\eta_2(s))\Big)\\ &\qquad=\sum_{i=1}^{2}\left(-\frac{1}{2}\int_\Omega\frac{|\nabla\eta_i(s)|^2}{\eta_i(s)}+\int_\Omega|\nabla V_i|^2\eta_i(s)\right)\\ &\qquad-\int_\Omega|\nabla(\eta_1(s)+\eta_2(s))|^2F_m''(\eta_1(s)+\eta_2(s)).\end{aligned}$$

By definition of F_m, for $m\geqslant 1$, $F_m''=mx^{m-2}$ for all $x>0$. And since $V_i\in W^{1,\infty}(\Omega)$,

$$\begin{aligned}&\partial_s\Big(\sum_{i=1}^{2}\frac{1}{2}\left(\mathcal{F}_1(\eta_i(s))+\mathcal{V}_i(\eta_i(s))\right)+\mathcal{F}_m(\eta_1(s)+\eta_2(s))\Big)\\ &\qquad\leqslant C-\frac{1}{2}\sum_{i=1}^{2}\int_\Omega|\nabla\eta_i(s)^{1/2}|^2-\frac{4}{m}\int_\Omega|\nabla(\eta_1(s)+\eta_2(s))^{m/2}|^2.\end{aligned}$$

Using (6.12) and a lower semi-continuity argument,

$$h\sum_{i=1}^{2}\int_\Omega|\nabla(\rho_{i,h}^k)^{1/2}|^2+\frac{4h}{m}\int_\Omega|\nabla(\rho_{1,h}^k+\rho_{2,h}^k)^{m/2}|^2\leqslant\sum_{i=1}^{2}\Big(\mathcal{F}_1(\rho_{i,h}^{k-1})-\mathcal{F}_1(\rho_{i,h}^k)\Big)+Ch.$$

Then if we sum over k, we obtain

$$\|\rho_{1,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))}+\|\rho_{2,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))}+\|(\rho_{1,h}+\rho_{2,h})^{m/2}\|_{L^2((0,T),H^1(\Omega))}\leqslant C(1+T).$$

□

Remark 6.6. *The bound on* $\|\rho_{i,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))}$ *does not depend of* m.

Now, let $\mathcal{G} : L^1(\Omega) \to (-\infty, +\infty]$ and $g : L^1(\Omega) \times L^1(\Omega) \to [0, +\infty]$ defined by

$$\mathcal{G}(\rho) := \begin{cases} \|\rho^{1/2}\|_{H^1(\Omega)} & \text{if } \rho \in \mathcal{P}^{\mathrm{ac}}(\Omega) \text{ and } \rho^{1/2} \in H^1(\Omega) \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$g(\rho, \mu) := \begin{cases} W_2(\rho, \mu) & \text{if } \rho, \mu \in \mathcal{P}^{\mathrm{ac}}(\Omega) \\ +\infty & \text{otherwise,} \end{cases}$$

$\mathcal{G}$ is l.s.c and its sublevels are relatively compact in $L^1(\Omega)$ (see [55, 73] or lemma 4.16 from chapter 4) and g is a pseudo-distance. According to (6.9) and (6.10), we have

$$\sup_{h \leqslant 1} \int_0^T \mathcal{G}(\rho_{i,h}(t))\, dt < +\infty, \text{ and } \lim_{\tau \searrow 0} \sup_{h \leqslant 1} \int_0^{T-\tau} g(\rho_{i,h}(t+\tau), \rho_{i,h}(t))\, dt = 0,$$

then applying an extension of the Aubin-Lions lemma proved by Rossi and Savaré in [107] (theorem 2), there exists a subsequence, not-relabeled, such that for $i = 1, 2$, $\rho_{i,h}$ converges in measure with respect to t in $L^1(\Omega)$ to ρ_i. Moreover Lebesgue's dominated convergence theorem implies that $\rho_{i,h}$ converges strongly in $L^1((0,T) \times \Omega)$ to ρ_i.

Now we use the same argument to obtain a strong convergence on a nonlinear quantity of $\rho_{1,h} + \rho_{2,h}$. We define $\mathcal{G}$ by

$$\mathcal{G}(\rho) := \begin{cases} \|\rho^{m/2}\|_{H^1(\Omega)} & \text{if } \rho \in \mathcal{P}^{\mathrm{ac}}(\Omega) \text{ and } \rho^{m/2} \in H^1(\Omega) \\ +\infty & \text{otherwise,} \end{cases}$$

and g is defined as before. We want to apply theorem 2 of [107] in $L^m(\Omega)$ over the sequence $\frac{\rho_{1,h}+\rho_{2,h}}{2}$. Using (6.10), we obtain

$$\sup_{h \leqslant 1} \int_0^T \mathcal{G}(\rho_{i,h}(t))\, dt < +\infty.$$

Moreover, we know that for all $\rho_1, \rho_2, \mu_1, \mu_2 \in \mathcal{P}^{\mathrm{ac}}(\Omega)$,

$$W_2^2\left(\frac{\rho_1 + \rho_2}{2}, \frac{\mu_1 + \mu_2}{2}\right) \leqslant \frac{1}{2} W_2^2(\rho_1, \mu_1) + \frac{1}{2} W_2^2(\rho_2, \mu_2).$$

Indeed, if we note γ_i an optimal transport plan for $W_2(\rho_i, \mu_i)$, $\gamma = \frac{\gamma_1+\gamma_2}{2}$ defines a transport plan in $\Pi\left(\frac{\rho_1+\rho_2}{2}, \frac{\mu_1+\mu_2}{2}\right)$ then

$$\begin{aligned} W_2^2\left(\frac{\rho_1 + \rho_2}{2}, \frac{\mu_1 + \mu_2}{2}\right) &\leqslant \int_{\Omega\times\Omega} |x-y|^2\, d\gamma(x,y) \\ &\leqslant \frac{1}{2}\left(\int_{\Omega\times\Omega} |x-y|^2\, d\gamma_1(x,y) + \int_{\Omega\times\Omega} |x-y|^2\, d\gamma_2(x,y)\right) \\ &\leqslant \frac{1}{2} W_2^2(\rho_1, \mu_1) + \frac{1}{2} W_2^2(\rho_2, \mu_2). \end{aligned}$$

Using (6.9), we obtain

$$\lim_{\tau \searrow 0} \sup_{h \leqslant 1} \int_0^{T-\tau} g\left(\frac{\rho_{1,h} + \rho_{2,h}}{2}(t+\tau), \frac{\rho_{1,h} + \rho_{2,h}}{2}(t)\right) dt = 0.$$

Theorem 2 of [107] and Lebesgue's dominated convergence theorem imply that $\rho_{1,h}+\rho_{2,h}$ converges strongly to $\rho_1+\rho_2$ in $L^m((0,T)\times\Omega)$. In addition, Krasnoselskii theorem (see [50], chapter 2) implies that $(\rho_{1,h} + \rho_{2,h})^{m/2}$ converges to $(\rho_1 + \rho_2)^{m/2}$ in $L^2((0,T) \times \Omega)$.

6.2.2 Euler-Lagrange equation and proof

In this section, we first give the optimality conditions of (6.6). Instead of using *horizontal perturbations*, $\rho_\varepsilon = \Phi_{\varepsilon\#}\rho_h^{k+1}$, as introduced in by Jordan, Kinderlherer and Otto in [65] and used in previous chapters, we will perturb ρ_h^{k+1} with *vertical perturbations* introduced in [25, 36, 110, 112] which consist to take $\rho_\varepsilon = (1-\varepsilon)\rho_h^{k+1} + \varepsilon\tilde{\rho}$, for all $\tilde{\rho} \in L^\infty(\Omega)$. Before giving the optimality conditions of (6.6), we state the following lemma.

Lemma 6.7. *For all $k \geqslant 1$, $\rho_{i,h}^k > 0$ a.e and $\log(\rho_{i,h}^k) \in L^1(\Omega)$.*

Proof. The proof is the same as lemma 8.6 from [112]. Let $\tilde{\rho} = c = \frac{1}{|\Omega|}$ the uniform density on Ω. We define $\rho_{i,\varepsilon}$ as the vertical perturbation of $\rho_{i,h}^{k+1}$ by $\tilde{\rho}$,

$$\rho_{i,\varepsilon} := (1-\varepsilon)\rho_{i,h}^{k+1} + \varepsilon\tilde{\rho}.$$

Using $(\rho_{1,\varepsilon}, \rho_{2,\varepsilon})$ as a competitor in (6.6), we obtain

$$\begin{aligned}\mathcal{F}_1(\rho_{1,h}^{k+1}) - \mathcal{F}_1(\rho_{1,\varepsilon}) + \mathcal{F}_1(\rho_{2,h}^{k+1}) - \mathcal{F}_1(\rho_{2,\varepsilon}) \\ \leqslant \sum_{i=1}^{2}\left(\int_\Omega V_i(\rho_{i,\varepsilon} - \rho_{i,h}^{k+1}) + \frac{1}{2h}W_2^2(\rho_{i,\varepsilon}, \rho_{i,h}^k) - \frac{1}{2h}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k)\right) \\ + \mathcal{F}_m(\rho_{1,\varepsilon} + \rho_{2,\varepsilon}) - \mathcal{F}_m(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}).\end{aligned}$$

We remark that

$$\int_\Omega V_i(\rho_{i,\varepsilon} - \rho_{i,h}^{k+1}) \leqslant C\varepsilon,$$

and using the convexity of $\frac{1}{2h}W_2^2(\cdot, \rho_{i,h}^k)$ we obtain

$$\frac{1}{2h}W_2^2(\rho_{i,\varepsilon}, \rho_{i,h}^k) - \frac{1}{2h}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k) \leqslant \varepsilon\left(\frac{1}{2h}W_2^2(\tilde{\rho}, \rho_{i,h}^k) - \frac{1}{2h}W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k)\right) \leqslant C\varepsilon.$$

If $m > 1$, then since $\rho_{1,\varepsilon} + \rho_{2,\varepsilon} \in L^m(\Omega) \cap L^1(\Omega)$, by convexity of F_m,

$$\mathcal{F}_m(\rho_{1,\varepsilon} + \rho_{2,\varepsilon}) - \mathcal{F}_m(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}) \leqslant C\varepsilon.$$

Now we denote A_i and B_i the sets defined by

$$A_i := \{\rho_{i,h}^{k+1} > 0\} \text{ and } B_i := \{\rho_{i,h}^{k+1} = 0\}.$$

On A_i, since $F_1(x) = x\log(x)$ is convex, $F_1(\rho_{i,h}^{k+1}) - F_1(\rho_{i,\varepsilon}) \geqslant (\rho_{i,h}^{k+1} - \rho_{i,\varepsilon})F_1'(\rho_{i,\varepsilon}) \geqslant \varepsilon(\rho_{i,h}^{k+1} - c)(1+\log(c))$. And on B_i, $F_1(\rho_{i,h}^{k+1}) - F_1(\rho_{i,\varepsilon}) = -\varepsilon c\log(\varepsilon c)$. This implies

$$-c\log(\varepsilon c)(|B_1| + |B_2|) + \sum_{i=1}^{2}\int_{A_i}(\rho_{i,h}^{k+1} - c)(1+\log(\rho_{i,\varepsilon})) \leqslant C. \tag{6.15}$$

Since $-c\log(\varepsilon c) \to +\infty$, when $\varepsilon \searrow 0$, we conclude that $|B_1| = |B_2| = 0$.

If $m = 1$, the proof is the same introducing

$$A := \{\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1} > 0\} \text{ and } B := \{\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1} = 0\}.$$

Then we prove in both case that $\rho_{i,h}^{k+1} > 0$ a.e.

Now using (6.15) and the fact that $(\rho_{i,h}^{k+1} - c)(1+\log(\rho_{i,\varepsilon}))$ is bounded from below by $(\rho_{i,h}^{k+1} - c)(1+\log(c))$ which is in L^1 and applying Fatou's lemma, we obtain

$$\int_\Omega (\rho_{i,h}^{k+1} - c)(1+\log(\rho_{i,h}^{k+1})) \leqslant C.$$

This implies that $(\rho_{i,h}^{k+1} - c)(1+\log(\rho_{i,h}^{k+1}))$ is L^1 and since $\rho_{i,h}^{k+1}$ and $\rho_{i,h}^{k+1}\log(\rho_{i,h}^{k+1})$ are in L^1, we conclude that $\log(\rho_{i,h}^{k+1})$ is in L^1.

□

This lemma is needed to ensure the uniqueness (up to a constant) of the Kantorovich potential in the transport from $\rho_{i,h}^{k+1}$ and $\rho_{i,h}^{k}$ and then we can easily compute the first variation of $W_2(\cdot,\rho_{i,h}^{k})$ according to proposition 7.17 of [112]. However, this lemma can be skip allowing to deal with nonlinear diffusion (see example 7.22 from [112]).

Proposition 6.8. *For $i\in\{1,2\}$, $\rho_{i,h}^{k+1}$ satisfies*

$$\nabla V_i\rho_{i,h}^{k+1}+\nabla\rho_{i,h}^{k+1}+\nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\rho_{i,h}^{k+1}+\frac{\nabla\varphi_{i,h}^{k+1}}{h}\rho_{i,h}^{k+1}=0\qquad a.e, \tag{6.16}$$

where $\varphi_{i,h}^{k+1}$ is the (unique) Kantorovich potential from $\rho_{i,h}^{k+1}$ to $\rho_{i,h}^{k}$.

Proof. We prove the result for $i=1$ and the proof is the same for $i=2$. Define $\rho_{1,\varepsilon}=\rho_{1,h}^{k+1}+\varepsilon(\tilde\rho-\rho_{1,h}^{k+1})$, for $\tilde\rho\in L^\infty(\Omega)$. Using optimality of $\rho_{1,h}^{k+1}$ in (6.6), we obtain

$$\begin{aligned}&\frac{1}{\varepsilon}\left(\frac{1}{2}W_2^2(\rho_{1,\varepsilon},\rho_{1,h}^{k})-\frac{1}{2}W_2^2(\rho_{1,h}^{k+1},\rho_{1,h}^{k})\right.\\&\left.+h\left(\mathcal{V}_i(\rho_{1,\varepsilon})-\mathcal{V}_i(\rho_{1,h}^{k+1})+\mathcal{F}_1(\rho_{1,\varepsilon})-\mathcal{F}_1(\rho_{1,h}^{k+1})+\mathcal{F}_m(\rho_{1,\varepsilon}+\rho_{2,h}^{k+1})-\mathcal{F}_m(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\right)\right)\geqslant 0.\end{aligned} \tag{6.17}$$

Lemma 6.7 ensures uniqueness (up to a constant) of the Kantorovich potentiel between $\rho_{1,h}^{k+1}$ and $\rho_{1,h}^{k}$ (proposition 7.18 of [112]). Then applying the proposition 7.17 of [112], the first variation of the Wasserstein distance exists and

$$\limsup_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\left(\frac{1}{2}W_2^2(\rho_{1,\varepsilon},\rho_{1,h}^{k})-\frac{1}{2}W_2^2(\rho_{1,h}^{k+1},\rho_{1,h}^{k})\right)\leqslant\int_\Omega\varphi_{1,h}^{k+1}\,d(\tilde\rho-\rho_{1,h}^{k+1}), \tag{6.18}$$

where $\varphi_{i,h}^{k+1}$ is the (unique) Kantorovich potential from $\rho_{i,h}^{k+1}$ to $\rho_{i,h}^{k}$. It is clear that

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\left(\mathcal{V}_1(\rho_{1,\varepsilon})-\mathcal{V}_1(\rho_{1,h}^{k+1})\right)=\int_\Omega V_1\,d(\tilde\rho-\rho_{1,h}^{k+1}). \tag{6.19}$$

Arguing as in the proof of lemma 3.1 of [25], since $F_1:x\mapsto x\log(x)$ is convex, the monotonicity of the incremental ratio gives for $\varepsilon<1$,

$$\frac{|F_1(\rho_{1,\varepsilon})-F_1(\rho_{1,h}^{k+1})|}{\varepsilon}\leqslant|F_1(\rho_{1,h}^{k+1})-F_1(\tilde\rho)|.$$

Since $F_1(\rho_{1,h}^{k+1})-F_1(\tilde\rho)\in L^1(\Omega)$, Lebesgue's dominated convergence theorem implies

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\left(\mathcal{F}_1(\rho_{1,\varepsilon})-\mathcal{F}_1(\rho_{1,h}^{k+1})\right)=\int_\Omega(1+\log(\rho_{1,h}^{k+1}))\,d(\tilde\rho-\rho_{1,h}^{k+1}). \tag{6.20}$$

With the same argument since F_m is convex,

$$\frac{|F_m(\rho_{1,\varepsilon}+\rho_{2,h}^{k+1})-F_m(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})|}{\varepsilon}\leqslant|F_m(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})-F_m(\tilde\rho+\rho_{2,h}^{k+1})|.$$

Then we obtain

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\left(\mathcal{F}_m(\rho_{1,\varepsilon}+\rho_{2,h}^{k+1})-\mathcal{F}_m(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\right)=\int_\Omega F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\,d(\tilde\rho-\rho_{1,h}^{k+1}). \tag{6.21}$$

Combining (6.17), (6.18), (6.19), (6.20), (6.21), we obtain, for all $\tilde\rho\in L^\infty(\Omega)$,

$$\begin{aligned}&\int_\Omega\left(\varphi_{1,h}^{k+1}+hV_1+h(1+\log(\rho_{1,h}^{k+1}))+hF_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\right)d\tilde\rho\\&\qquad\geqslant\int_\Omega\left(\varphi_{1,h}^{k+1}+hV_1+h(1+\log(\rho_{1,h}^{k+1}))+hF_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\right)d\rho_{1,h}^{k+1}.\end{aligned}$$

Applying proposition 7.20 from [112], there exists a constant C such that

$$\varphi_{1,h}^{k+1} + hV_1 + h(1+\log(\rho_{1,h}^{k+1})) + hF_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1}) = C, \qquad \rho_{1,h}^{k+1} - a.e. \tag{6.22}$$

Since $\varphi_{1,h}^{k+1} + hV_1$ is a Lipschitz function, it is differentiable a.e (Rademacher's theorem). Moreover $\rho_{1,h}^{k+1}$ is absolutly continuous with repsect to Lebesgue measure then $\varphi_{1,h}^{k+1}$ is differentiable $\rho_{1,h}^{k+1}$-a.e and it is the same for $h(1+\log(\rho_{1,h}^{k+1})) + hF_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})$ because (6.22). Then

$$\nabla\varphi_{1,h}^{k+1} + h\nabla V_1 + h\nabla\log(\rho_{1,h}^{k+1}) + h\nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1}) = 0, \qquad \rho_{1,h}^{k+1} - a.e,$$

which concludes the proof.

□

A consequence of previous proposition is that $\rho_{1,h}$ and $\rho_{2,h}$ are solutions of a discrete approximation of system (6.2), see for example proposition 4.18 from chapter 4.

Proposition 6.9. *Let $h > 0$, for all $T > 0$, let N such that $N = \lfloor \frac{T}{h} \rfloor$. Then for all $(\phi_1, \phi_2) \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)^l$ and for all $i \in \{1,2\}$,*

$$\int_0^T\int_\Omega \rho_{i,h}(t,x)\partial_t\phi_i(t,x)\,dxdt + \int_\Omega \rho_{i,0}(x)\phi_i(0,x)\,dx = h\sum_{k=0}^{N-1}\int_\Omega \nabla V_i(x)\cdot\nabla\phi_i(t_k,x)\rho_{i,h}^{k+1}(x)\,dx$$
$$+ h\sum_{k=0}^{N-1}\int_\Omega \nabla\rho_{i,h}^{k+1}(x)\cdot\nabla\phi_i(t_k,x)\,dx + h\sum_{k=0}^{N-1}\int_\Omega \nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\cdot\nabla\phi_i(t_k,x)\rho_{i,h}^{k+1}(x)\,dx$$
$$+ \sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\phi_i(t_k,\cdot)](x,y)d\gamma_{i,h}^k(x,y)$$

where $t_k = hk$ ($t_N := T$) and $\gamma_{i,h}^k$ is the optimal transport plan in $W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1})$. Moreover, $\mathcal{R}$ is defined such that, for all $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\mathbb{R}^n)}|x-y|^2.$$

Now we prove theorem 6.1.

Proof of theorem 6.1. We have to pass to the limit in all terms in proposition 6.9. The remainder term converges to 0 using estimate (6.9) and the linear term converges to

$$\int_0^T\int_\Omega \rho_i\partial_t\phi_i - \int_0^T\int_\Omega \nabla V_i\cdot\nabla\phi_i\rho_i,$$

when h goes to 0 because $\rho_{i,h}$ converges to $\rho_i \in \mathcal{C}([0,T],\mathcal{P}^{\mathrm{ac}}(\Omega))$ in $L^\infty([0,T],\mathcal{P}^{\mathrm{ac}}(\Omega))$.

Moreover since $\nabla\rho_{i,h} = 2\rho_{i,h}^{1/2}\nabla\rho_{i,h}^{1/2}$, then $\nabla\rho_{i,h}$ is bounded in $\mathcal{M}^n((0,T)\times\Omega)$ because proposition (6.10). We conclude that $\nabla\rho_{i,h}$ converges in $\mathcal{M}^n((0,T)\times\Omega)$ to $D\rho_i$ because $\rho_{i,h}$ strongly converges to ρ_i in $L^1((0,T)\times\Omega)$. This implies that the diffusive linear term converges to

$$\int_0^T\int_\Omega \nabla\phi_i\cdot\,dD\rho_i.$$

It remains to study the convergence of the cross diffusion term. First we remark that we can rewrite $\nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})$ by

$$\nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1}) = 2\frac{(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})^{m/2}}{\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1}}\nabla(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})^{m/2},$$

then

$$\nabla F_m'(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})\rho_{i,h}^{k+1} = 2G_{1-m/2}(\rho_{1,h}^{k+1},\rho_{2,h}^{k+1})\nabla(\rho_{1,h}^{k+1}+\rho_{2,h}^{k+1})^{m/2},$$

with $G_\alpha : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ is the continuous function (for $\alpha < 1$) defined by

$$G_\alpha(x,y) := \begin{cases} \frac{x}{(x+y)^\alpha} & \text{if } x > 0, y \geqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

As $m \geqslant 1$, $1 - \frac{m}{2} < 1$ so $G_{1-m/2}$ is continuous and since, up to a subsequence, $\rho_{i,h}$ converges to ρ_i a.e, we obtain that $G_{1-m/2}(\rho_{1,h}, \rho_{2,h})$ converges to $G_{1-m/2}(\rho_1, \rho_2)$ a.e. Moreover,

$$\left|G_{1-m/2}(\rho_{1,h}, \rho_{2,h})\right| = \left|(\rho_{1,h} + \rho_{2,h})^{m/2} \frac{\rho_{1,h}}{\rho_{1,h} + \rho_{2,h}}\right| \leqslant (\rho_{1,h} + \rho_{2,h})^{m/2}. \tag{6.23}$$

We complete the proof if we prove

$$h \sum_{k=0}^{N-1} \int_\Omega \nabla F_m'(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}) \cdot \nabla \phi_i(t_k, x) \rho_{i,h}^{k+1}(x)\, dx \to \int_0^T \int_\Omega 2G_{1-m/2}(\rho_1, \rho_2) \nabla(\rho_1 + \rho_2)^{m/2} \cdot \nabla \phi_i \, dxdt.$$

$$\left| h \sum_{k=0}^{N-1} \int_\Omega \nabla F_m'(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}) \cdot \nabla \phi_i(t_k, x) \rho_{i,h}^{k+1}(x)\, dx \right.$$
$$\left. - \int_0^T \int_\Omega 2G_{1-m/2}(\rho_1, \rho_2) \nabla(\rho_1 + \rho_2)^{m/2} \cdot \nabla \phi_i \, dxdt \right|$$

$$\leqslant \left| h \sum_{k=0}^{N-1} \int_\Omega \nabla F_m'(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}) \cdot \nabla \phi_i(t_k, x) \rho_{i,h}^{k+1}(x)\, dx \right.$$
$$\left. - \int_0^T \int_\Omega \nabla F_m'(\rho_{1,h} + \rho_{2,h}) \cdot \nabla \phi_i(t, x) \rho_{i,h}(t, x)\, dxdt \right|$$

$$+ \left| \int_0^T \int_\Omega 2G_{1-m/2}(\rho_{1,h}, \rho_{2,h}) \nabla(\rho_{1,h} + \rho_{2,h})^{m/2} \cdot \nabla \phi_i(t, x)\, dxdt \right.$$
$$\left. - \int_0^T \int_\Omega 2G_{1-m/2}(\rho_1, \rho_2) \nabla(\rho_{1,h} + \rho_{2,h})^{m/2} \cdot \nabla \phi_i(t, x)\, dxdt \right|$$

$$+ \left| \int_0^T \int_\Omega 2G_{1-m/2}(\rho_1, \rho_2) \nabla(\rho_{1,h} + \rho_{2,h})^{m/2} \cdot \nabla \phi_i(t, x)\, dxdt \right.$$
$$\left. - \int_0^T \int_\Omega 2G_{1-m/2}(\rho_1, \rho_2) \nabla(\rho_1 + \rho_2)^{m/2} \cdot \nabla \phi_i(t, x)\, dxdt \right|$$

The first term on the right hand side converges to 0 because $\nabla \phi_i$ is Lipschitz and

$$\int_0^T \int_\Omega |\nabla F_m'(\rho_{1,h} + \rho_{2,h})| \rho_{i,h} \leqslant C,$$

where C is a constant that does not depend of h. Indeed since $\rho_{i,h}^{1/2}$ is bounded in $L^2((0,T), H^1(\Omega))$, $\nabla \rho_{i,h} = 2\rho_{i,h}^{1/2} \nabla \rho_{i,h}^{1/2}$ is bounded in $L^1((0,T) \times \Omega)$, then using (6.16) and (6.9) we prove the statement.

The second term is handled using (6.10) and Lebesgue's dominated convergence theorem.

$$\left| \int_0^T \int_\Omega (G_{1-m/2}(\rho_{1,h}, \rho_{2,h}) - G_{1-m/2}(\rho_1, \rho_2)) \nabla(\rho_{1,h} + \rho_{2,h})^{m/2} \cdot \nabla \phi_i \right|$$
$$\leqslant \|\nabla(\rho_{1,h} + \rho_{2,h})^{m/2}\|_{L^2((0,T)\times\Omega)} \left(\int_0^T \int_\Omega |G_{1-m/2}(\rho_{1,h}, \rho_{2,h}) - G_{1-m/2}(\rho_1, \rho_2)|^2 |\nabla \phi_i|^2 \right)^{1/2}.$$

$\|\nabla(\rho_{1,h}+\rho_{2,h})^{m/2}\|_{L^2((0,T)\times\Omega)}$ is bounded using (6.10) and $G_{1-m/2}(\rho_{1,h},\rho_{2,h})$ converges to $G_{1-m/2}(\rho_1,\rho_2)$ a.e. Since $(\rho_{1,h}+\rho_{2,h})^{m/2}$ converges to $(\rho_1+\rho_2)^{m/2}$ in $L^2((0,T)\times\Omega)$, there exists a function $g\in L^2((0,T)\times\Omega)$ such that, up to a subsequence,

$$|(\rho_{1,h}+\rho_{2,h})^{m/2}|\leqslant g.$$

Then using (6.23),

$$|G_{1-m/2}(\rho_{1,h},\rho_{2,h})-G_{1-m/2}(\rho_1,\rho_2)|^2|\nabla\phi_i|^2\leqslant 4g^2\|\nabla\phi_i\|_{L^\infty}\in L^1((0,T)\times\Omega).$$

Lebesgue's dominated convergence theorem implies that the second term converges to 0 when h goes to 0.

The third term goes to 0 because $\nabla(\rho_{1,h}+\rho_{2,h})^{m/2}$ converges weakly to $\nabla(\rho_1+\rho_2)^{m/2}$ in $L^2((0,T)\times\Omega)$ and $G_{1-m/2}(\rho_1,\rho_2)\nabla\phi_i\in L^2((0,T)\times\Omega)$.

□

6.3 Crowd motion with diffusion for two species

In this section we will prove an existence result for system of Fokker-Planck equations coupled by hard congestion on the sum. In other word we prove the existence of weak solution in (6.1). This system can be seen as gradient flow in a Wasserstein product space. Using the Jordan-Kinderlherer-Otto scheme ([65]), we construct two sequences defined in the following way: Let $h>0$ be a time step, we construct a sequence $(\rho^k_{1,h},\rho^k_{2,h})$ with $(\rho^0_{1,h},\rho^0_{2,h})=(\rho_{1,0},\rho_{2,0})$ and $(\rho^{k+1}_{1,h},\rho^{k+1}_{2,h})$ is a solution of

$$\inf_{(\rho_1,\rho_2)\in\mathcal{K}}\sum_{i=1}^{2}\left[\frac{1}{2h}W_2^2(\rho_i,\rho^k_{i,h})+\mathcal{F}_1(\rho_i)+\mathcal{V}_i(\rho_i)\right], \tag{6.24}$$

where $\mathcal{K}:=\left\{(\rho_1,\rho_2)\in\mathcal{P}^{\mathrm{ac}}(\Omega)^2\,:\,\rho_1+\rho_2\leqslant 1\right\}$ and $|\Omega|>2$. The existence of these sequences is obvious by standard compactness and l.s.c argument. As before, we define the piecewise constant interpolations $\rho_{i,h}:\mathbb{R}^+\to\mathcal{P}^{\mathrm{ac}}(\Omega)$ by

$$\rho_{i,h}(t):=\rho^{k+1}_{i,h},\qquad \text{if } t\in(kh,(k+1)h].$$

6.3.1 Properties of sequences

Estimates and convergences

Lemma 6.10. *Minimizers of* (6.24) *satisfy* $\rho^k_{i,h}>0$ *a.e. and* $\log(\rho^k_{i,h})\in L^1(\Omega)$.

Proof. The proof is the same as in lemma 8.5 from [112] (see also lemma 6.7). Indeed we can use a constant perturbation $\tilde\rho$ because $(\tilde\rho,\tilde\rho)$ is admissible in (6.24) ($\tilde\rho+\tilde\rho=2/|\Omega|\leqslant 1$). □

This lemma will help us because it implies the uniqueness of the pair of Kantorovich potentials from $\rho^{k+1}_{i,h}$ to $\rho^k_{i,h}$ and then the existence of the first variation of $W_2^2(\cdot,\rho^k_{i,h})$ (propositions 7.18 and 7.17 from [112]).

Moreover, standard gradient flow theory implies that for $i\in\{1,2\}$ and for all $k\geqslant 0$,

$$\rho^k_{1,h}+\rho^k_{2,h}\leqslant 1,\qquad \mathcal{F}_1(\rho^k_{i,h})\leqslant C,\qquad \sum_{k=0}^{N-1}W_2^2(\rho^k_{i,h},\rho^{k+1}_{i,h})\leqslant Ch, \tag{6.25}$$

as in proposition 6.3. Using (6.25) and the refined version of Ascoli-Arzelà's theorem (see [4], proposition 3.3.1), $\rho_{i,h}$ converges to $\rho_i\in\mathcal{C}^{1/2}([0,T],\mathcal{P}^{\mathrm{ac}}(\Omega))$ in $L^\infty([0,T],\mathcal{P}^{\mathrm{ac}}(\Omega))$.

In order to obtain an estimate on the gradient, we apply the flow interchange argument as previously. We keep same notations than in the previous section. We note η_i the heat flow with initial condition $\rho^k_{i,h}$. Since the heat flow decreases the L^∞-norm, $(\eta_1(t),\eta_2(t))$, defined in (6.11), is

admissible for problem (6.24), for all $t \geqslant 0$. Then using the same computations than in proposition 6.5, we obtain

$$\|\rho_{1,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))} + \|\rho_{2,h}^{1/2}\|_{L^2((0,T),H^1(\Omega))} \leqslant C. \tag{6.26}$$

With the same argument as in the previous section, we obtain the strong convergence of $\rho_{i,h}$ to ρ_i in $L^1((0,T)\times\Omega)$.

First variation and pressure associated to $\mathcal{K}$

Lemma 6.11. *Let $(\rho_{1,h}^{k+1}, \rho_{2,h}^{k+1})$ be the unique solution of (6.24). Then for all $(\rho_1,\rho_2)\in\mathcal{K}$,*

$$\int_\Omega \psi_1^{k+1}(\rho_1 - \rho_{1,h}^{k+1}) + \int_\Omega \psi_2^{k+1}(\rho_2 - \rho_{2,h}^{k+1}) \geqslant 0, \tag{6.27}$$

where $\psi_i^{k+1} = \frac{\varphi_{i,h}^{k+1}}{h} + V_i + 1 + \log(\rho_{i,h}^{k+1})$ and $\varphi_{i,h}^{k+1}$ is the optimal (up to a constant) Kantorovich potential in $W_2(\rho_{i,h}^{k+1}, \rho_{i,h}^k)$.

Proof. We refer to the proof of lemma 3.1 from [89]. □

In the next proposition, we introduce the pressure associated to the constraint $\mathcal{K}$.

Proposition 6.12. *There exists $p_h^k \geqslant 0$ such that for all, $k \geqslant 1$,*

$$\nabla p_h^k = -\frac{\nabla\varphi_{i,h}^k}{h} - \nabla V_i - \nabla\log(\rho_{i,h}^k), \tag{6.28}$$

for $i = 1,2$ and $p_h^k(1-\rho_{1,h}^k - \rho_{2,h}^k) = 0$ a.e.

Proof. Before starting, we remark that (6.27) can be rewrite as

$$\int_\Omega \psi_1^k h_1 + \int_\Omega \psi_2^k h_2 \geqslant 0,$$

for all functions h_1, h_2 such that

$$h_1 + h_2 \leqslant \frac{1-\rho_{1,h}^k - \rho_{2,h}^k}{\varepsilon}, \quad h_i \geqslant \frac{-\rho_{i,h}^k}{\varepsilon} \text{ and } \int_\Omega h_i = 0, \tag{6.29}$$

for all $\varepsilon > 0$. Let $S := \{\rho_{1,h}^k + \rho_{2,h}^k = 1\}$ be the set where the constraint is saturated.

First, we choose $h_2 = 0$ on Ω and $h_1 = 0$ on S. Then we have

$$\int_{S^c} \psi_1^k h_1 \geqslant 0.$$

Since ε is not fixed and $\rho_{1,h}^k > 0$ a.e., this implies that there exists a constant C_1 such that $\psi_1^k = C_1$ a.e. on S^c. And using the same argument, with $h_1 = 0$ on Ω and $h_2 = 0$ on S, we find a constant C_2 such that $\psi_2^k = C_2$ a.e. on S^c.

Since h_i satisfy (6.29), we have

$$\int_\Omega (\psi_1^k - C_1)h_1 + \int_\Omega (\psi_2^k - C_2)h_2 \geqslant 0.$$

Now if we choose $h_1 = h$ and $h_2 = -h$ on S and we interchange after, we find

$$\int_S ((\psi_1^k - C_1) - (\psi_2^k - C_2))h = 0,$$

for all $h \in L^\infty(\Omega)$. We conclude that $(\psi_1^k - C_1) = (\psi_2^k - C_2) = \psi$ a.e on S. This implies that

$$\int_S \psi(h_1 + h_2) \geqslant 0.$$

But, since $h_1 + h_2 \leqslant 0$ on S, $\psi \leqslant 0$ a.e on S. Then we define p_h^k by

$$p_h^k := (C_1 - \psi_1^k)_+ = (C_2 - \psi_2^k)_+.$$

□

We define the piecewise interpolation $p_h : \mathbb{R}^+ \to W^{1,1}(\Omega)$ by

$$p_h(t) := p_h^{k+1}, \qquad \text{if } t \in (kh, (k+1)h].$$

We remark that $p_h(t) \geqslant 0$ and for all $t \geqslant 0$, $p_h(t)(1 - \rho_{1,h}(t) - \rho_{2,h}(t)) = 0$ a.e. Moreover, we can immediately deduce the following estimate on the pressure.

Proposition 6.13. *For all $T > 0$, p_h is bounded in $L^2((0,T), H^1(\Omega))$.*

Proof. First, we prove that p_h is bounded in $L^2((0,T) \times \Omega)$ and then we will conclude using Poincaré's inequality. By definition of p_h^{k+1}, we have

$$\begin{aligned}
\int_\Omega |\nabla p_h^{k+1}|^2(\rho_{1,h}^{k+1} + \rho_{2,h}^{k+1}) &= \sum_{i=1}^2 \int_\Omega |\nabla \psi_{i,h} h^{k+1}|^2 \rho_{i,h}^{k+1} \\
&\leqslant C \sum_{i=1}^2 \left(\int_\Omega \left| \frac{\nabla \phi_{i,h}^{k+1}}{h} \right|^2 \rho_{i,h}^{k+1} + \int_\Omega |\nabla V_i|^2 \rho_{i,h}^{k+1} + \int_\Omega \frac{|\nabla \rho_{i,h}^{k+1}|^2}{\rho_{i,h}^{k+1}} \right) \\
&\leqslant C \sum_{i=1}^2 \left(\frac{1}{h^2} W_2^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) + C + \|(\rho_{i,h}^{k+1})^{1/2}\|_{H^1(\Omega)} \right),
\end{aligned}$$

where the last line is obtained using the fact that $\nabla V_i \in L^\infty(\Omega)$. Summing the previous inequalities over k and using (6.25) and (6.26), we obtain that

$$\int_0^T \int_\Omega |\nabla p_h(t)|^2 (\rho_{1,h}(t) + \rho_{2,h}(t)) \leqslant C.$$

And since $p_h(t) = 0$ a.e on $\{\rho_{1,h}(t) + \rho_{2,h}(t) < 1\}$,

$$\int_0^T \int_\Omega |\nabla p_h(t)|^2 = \int_0^T \int_\Omega |\nabla p_h(t)|^2 (\rho_{1,h}(t) + \rho_{2,h}(t)) \leqslant C.$$

We conclude using the same argument as [94]. Using Poincaré's inequality, since $|\{p_h(t) = 0\}| \geqslant |\{\rho_{1,h}(t) + \rho_{2,h}(t) < 1\}| \geqslant |\Omega| - 2 > 0$, we obtain that p_h is bounded in $L^2((0,T), H^1(\Omega))$. □

Proposition 6.14. *Let $h > 0$, for all $T > 0$, let N such that $N = \lfloor \frac{T}{h} \rfloor$. Then for all $(\phi_1, \phi_2) \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^n)^l$ and for all $i \in \{1,2\}$,*

$$\begin{aligned}
\int_0^T \int_\Omega \rho_{i,h}(t,x) \partial_t \phi_i(t,x)\, dxdt + \int_\Omega \rho_{i,0}(x) \phi_i(0,x)\, dx &= h \sum_{k=0}^{N-1} \int_\Omega \nabla V_i(x) \cdot \nabla \phi_i(t_k, x) \rho_{i,h}^{k+1}(x)\, dx \\
&+ h \sum_{k=0}^{N-1} \int_\Omega \nabla \rho_{i,h}^{k+1}(x) \cdot \nabla \phi_i(t_k, x)\, dx + h \sum_{k=0}^{N-1} \int_\Omega \nabla p_h^{k+1} \cdot \nabla \phi_i(t_k, x) \rho_{i,h}^{k+1}(x)\, dx \\
&+ \sum_{k=0}^{N-1} \int_{\Omega \times \Omega} \mathcal{R}[\phi_i(t_k, \cdot)](x,y) d\gamma_{i,h}^k(x,y)
\end{aligned}$$

where $t_k = hk$ ($t_N := T$) and $\gamma_{i,h}^k$ is the optimal transport plan in $W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1})$. Moreover, $\mathcal{R}$ is defined such that, for all $\phi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2} \|D^2 \phi\|_{L^\infty([0,T) \times \mathbb{R}^n)} |x - y|^2.$$

Proof. We multiply (6.28) by $\rho_{i,h}^{k+1}$ and the proof is the same as in proposition 6.9. □

6.3.2 Proof of theorem 6.2

First, we recall the lemma given in [89, 94] useful to analyse the pressure field.

Lemma 6.15. *Let $(p_h)_{h>0}$ be a bounded sequence in $L^2([0,T],H^1(\Omega))$ and $(\rho_h)_{h>0}$ a sequence of piecewise constant curves valued in $\mathcal{P}(\Omega)$ which satisfiy $W_2(\rho_h(t),\rho_h(s)) \leqslant C\sqrt{t-s-h}$ for all $s<t\in[0,T]$ and $\rho_h \leqslant C$ for a fixed constant C. Suppose that*

$$p_h \geqslant 0, \quad p_h(1-\rho_h)=0, \quad \rho_h \leqslant 1,$$

and that

$$p_h \rightharpoonup p \text{ weakly in } L^2([0,T],H^1(\Omega)) \text{ and } \rho_h \to \rho \text{ uniformly in } \mathcal{P}(\Omega).$$

Then $p(1-\rho)=0$.

We apply this lemma to $\rho_h := \rho_{1,h}+\rho_{2,h}$ and p_h. According to proposition 6.13, p_h weakly converges in $L^2((0,T)\times\Omega)$ to p such that

$$p \geqslant 0, \quad p(1-\rho_1-\rho_2)=0, \quad \rho_1+\rho_2 \leqslant 1. \tag{6.30}$$

Moreover, using the estimate on p_h, we know that ∇p_h weakly converges to ∇p in $L^2((0,T)\times\Omega)$. Then since $\rho_{i,h}$ strongly converges to ρ_i in $L^1((0,T)\times\Omega)$ with $\rho_{i,h},\rho_i \leqslant 1$, we have $\rho_{i,h}\nabla p_h$ narrowly converges to $\rho_i\nabla p$.

Using the same argument as in the previous section, $\nabla\rho_{i,h}$ converges in $\mathcal{M}^n((0,T)\times\Omega)$ and the other terms, which are linear, in proposition 6.14 converges because of the convergence of $\rho_{i,h}$ in $L^\infty((0,T),\mathcal{P}(\Omega))$.

Remark 6.16. *With a finer analysis, we can prove that $\nabla\rho_{i,h}$ converges weakly to $\nabla\rho_i$ in $L^2((0,T)\times\Omega)$. Indeed, if we use again* (6.28) *and $\rho_{i,h}^{k+1}\leqslant 1$, we obtain that*

$$|\nabla\rho_{i,h}^{k+1}|^2 \leqslant C\left(\frac{|\nabla\varphi_{i,h}^{k+1}|^2}{h^2}\rho_{i,h}^{k+1}+|\nabla V_i|^2\rho_{i,h}^{k+1}+|\nabla p_h^{k+1}|^2\right).$$

Since ∇p_h is bounded in $L^2((0,T)\times\Omega)$ and

$$h\sum_{k=0}^{N-1}\int_\Omega \frac{|\nabla\varphi_{i,h}^{k+1}|^2}{h^2}\rho_{i,h}^{k+1} \leqslant C,$$

because of (6.25)*, then*

$$\|\nabla\rho_{i,h}\|_{L^2((0,T)\times\Omega)} \leqslant C.$$

6.4 Links between these systems and remarks on uniqueness

In this section, we focus on the special case where $V_1=V_2=V\in W^{1,\infty}(\Omega)$. We denote $(\rho_{1,m},\rho_{2,m})$ and $(\rho_{1,\infty},\rho_{2,\infty})$ solutions of (6.1) and (6.2), respectively. In the following, we prove that $(\rho_{1,m},\rho_{2,m})$ and $(\rho_{1,\infty},\rho_{2,\infty})$ are unique and that $(\rho_{1,m},\rho_{2,m})$ converges to $(\rho_{1,\infty},\rho_{2,\infty})$ in L^1 when m goes to $+\infty$.

These results are based on the remark that $\rho_{1,m}+\rho_{2,m}$ solves

$$\partial_t\mu-\Delta\mu-\operatorname{div}(\mu\nabla V)-\operatorname{div}(\mu\nabla F_m'(\mu))=0, \tag{6.31}$$

with initial condition $\mu_{|t=0}=\rho_{1,0}+\rho_{2,0}$. By geodesic convexity of $\mathcal{E}$ and $\mathcal{F}_m$, we know that solution of (6.31) is unique (see, for example, [35] theorem 6.1 and [73] theorem 7.1). In the next proposition, we give an estimate $L^2((0,T),H^1(\Omega))$ of $F_m'(\mu)$ independant of m.

Proposition 6.17. *Let μ_m be the solution of* (6.31)*. Then $F_m'(\mu_m)$ is bounded independently of m in $L^2((0,T),H^1(\Omega))$, for all $T<+\infty$.*

Proof. We use the flow interchange argument with $\partial_t \nu = \Delta \nu^{m-1} + \varepsilon \Delta \nu$ to obtain a L^2 bound of $\nabla F_m'(\mu_m)$. We refer to lemma 7.30 from chapter 7 for the proof of this result. Using the L^1-bound of $F_m'(\mu_m)$ and the Poincaré-Wirtinger inequality, we conclude the proof. □

Proposition 6.17 and an L^1-contraction argument yield the following:

Theorem 6.18. *For $1 \leqslant m \leqslant +\infty$, there is a unique solution of (6.1) ((6.2) if $m = +\infty$). Moreover, if $(\rho^1_{1,m}, \rho^1_{2,m})$ and $(\rho^2_{1,m}, \rho^2_{2,m})$ are two solutions of (6.1) ((6.2) if $m = +\infty$),*

$$\|\rho^1_{i,m}(t,\cdot) - \rho^2_{i,m}(t,\cdot)\|_{L^1(\Omega)} \leqslant \|\rho^1_{i,m}(0,\cdot) - \rho^2_{i,m}(0,\cdot)\|_{L^1(\Omega)}.$$

Proof. First if $m < +\infty$, since $\rho_{1,m} + \rho_{2,m}$ solves (6.31), then it is unique and according to proposition 6.17, $p_m := F_m'(\rho_{1,m} + \rho_{2,m})$ is in $L^2((0,T), H^1(\Omega))$. Moreover, if $m = +\infty$, we have already shown that the pressure p_∞ associated to the constraint $\rho_{1,\infty} + \rho_{2,\infty} \leqslant 1$ in theorem 6.2 is in $L^2((0,T), H^1(\Omega))$. According to [85], we know that $(\rho_{1,\infty} + \rho_{2,\infty}, p_\infty)$ is unique.

Now, we will give the proof of the L^1-contraction. We follow the line of [1]. We start to remark that $\rho^i_{1,m}$ solves

$$\partial_t \rho^i_{1,m} - \Delta \rho^i_{1,m} - \operatorname{div}(\rho^i_{1,m}(\nabla V + \nabla p_m)) = 0,$$

for all $1 \leqslant m \leqslant +\infty$.

We note $\Omega_T := (0,T) \times \Omega$. For $\delta > 0$, define

$$\zeta_\delta := \phi_\delta(\rho^1_{1,m} - \rho^2_{1,m}),$$

where

$$\phi_\delta(z) := \begin{cases} 0 & \text{if } z \leqslant 0, \\ \frac{z}{\delta} & \text{if } 0 \leqslant z \leqslant \delta, \\ 1 & \text{if } z \geqslant \delta. \end{cases}$$

Using ζ_δ, or a smooth approximation of ζ_δ in the equation satisfies by $\rho^1_{1,m}$ and $\rho^2_{1,m}$, we obtain

$$\iint_{\Omega_T} \partial_t(\rho^1_{1,m} - \rho^2_{1,m})\zeta_\delta = -\iint_{\Omega_T} \left((\rho^1_{1,m} - \rho^2_{1,m})(\nabla V + \nabla p_m) \cdot \nabla \zeta_\delta + \nabla(\rho^1_{1,m} - \rho^2_{1,m}) \cdot \nabla \zeta_\delta\right) dxdt.$$

We introduce $\Omega_T^\delta := \Omega_T \cap \{0 < \rho^1_{1,m} - \rho^2_{1,m} < \delta\}$. Then by definition of ζ_δ and using Young's inequality

$$\begin{aligned}
&\iint_{\Omega_T} \partial_t(\rho^1_{1,m} - \rho^2_{1,m})\zeta_\delta \\
&\quad = -\frac{1}{\delta} \iint_{\Omega_T^\delta} \left((\rho^1_{1,m} - \rho^2_{1,m})(\nabla V + \nabla p_m) \cdot \nabla(\rho^1_{1,m} - \rho^2_{1,m}) + |\nabla(\rho^1_{1,m} - \rho^2_{1,m})|^2\right) dxdt \\
&\quad \leqslant \frac{1}{2\delta} \iint_{\Omega_T^\delta} (\rho^1_{1,m} - \rho^2_{1,m})^2 |\nabla V + \nabla p_m|^2 \, dxdt - \frac{1}{2\delta} \iint_{\Omega_T^\delta} |\nabla(\rho^1_{1,m} - \rho^2_{1,m})|^2 \, dxdt \\
&\quad \leqslant \frac{1}{2} \|\nabla V + \nabla p_m\|^2_{L^2(\Omega_T)} \delta \to 0,
\end{aligned}$$

when $\delta \searrow 0$.

If we reverse the roles of $\rho^1_{1,m}$ and $\rho^2_{1,m}$, we have

$$\iint_{\Omega_T} \partial_t(|\rho^1_{1,m} - \rho^2_{1,m}|) \leqslant 0,$$

which concludes the proof.

□

In the end of this section, we prove that $(\rho_{1,m}, \rho_{2,m})$ converges to the solution $(\rho_{1,\infty}, \rho_{2,\infty})$ of (6.2) when $m \nearrow +\infty$.

Theorem 6.19. *Up to a subsequence, as $m \to +\infty$, the solution of (6.1) $(\rho_{1,m}, \rho_{2,m})$ converges strongly $L^1((0,T)\times\Omega)$ to $(\rho_{1,\infty}, \rho_{2,\infty})$ and $p_m := F_m'(\rho_{1,m}+\rho_{2,m})$ converges weakly in $L^2((0,T), H^1(\Omega))$ to p_∞, where $(\rho_{1,\infty}, \rho_{2,\infty}, p_\infty)$ is the solution of (6.2).*

Proof. First we prove the convergence of $\rho_{i,m}$. We start noticing that the estimate (6.10) does not depend of m and then using remark 6.4, we have

$$\|\rho_{i,m}^{1/2}\|_{L^2((0,T),H^1(\Omega))} \leqslant C_T \text{ and } W_2(\rho_{i,m}(t), \rho_{i,m}(s)) \leqslant C|t-s|^{1/2},$$

for all $t, s \leqslant T$ and with C independant of m. Then using the Rossi-Savaré theorem we obtain that $\rho_{i,m}$ converges to $\rho_{i,\infty}$ in $L^1((0,T)\times\Omega)$. Moreover, proposition 6.17 gives that p_m converges weakly in $L^2((0,T), H^1(\Omega))$ to p_∞ and obviously $p_m \geqslant 0$.

To conclude the proof, by uniqueness of the solution of (6.2), we only have to prove that

$$\rho_{1,\infty} + \rho_{2,\infty} \leqslant 1 \text{ and } p_\infty(1-\rho_{1,\infty}-\rho_{2,\infty}) = 0 \qquad a.e.$$

We start to show that $\rho_{1,\infty}+\rho_{2,\infty} \leqslant 1$. To do that we use the argument of lemma 4.4 from [2]. The estimate (6.8) does not depend of m so we have

$$\int_0^T \int_\Omega (\rho_{1,m}+\rho_{2,m})^m \leqslant (m-1)C,$$

and for $m > 2$,

$$\iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} (\rho_{1,m}+\rho_{2,m})^m \, dxdt \geqslant \iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} \frac{m(m-1)}{2}(\rho_{1,m}+\rho_{2,m}-1)^2 \, dxdt.$$

Then we obtain

$$\int_0^T \int_\Omega (\rho_{1,m}+\rho_{2,m}-1)_+^2 \, dxdt \leqslant \frac{2C}{m} \to 0, \tag{6.32}$$

when $m \to +\infty$, which implies that $\rho_{1,\infty}+\rho_{2,\infty} \leqslant 1$.

To obtain the second part of the claim, we start to prove

$$\int_0^T \int_\Omega p_m(1-\rho_{1,m}-\rho_{2,m}) \, dxdt \to \int_0^T \int_\Omega p_\infty(1-\rho_{1,\infty}-\rho_{2,\infty}) \, dxdt.$$

Indeed we have

$$\begin{aligned}\left| \int_0^T \int_\Omega \big(p_m(1-\rho_{1,m}-\rho_{2,m}) - p_\infty(1-\rho_{1,\infty}-\rho_{2,\infty})\big) \, dxdt \right| & \\ \leqslant \left| \int_0^T \int_\Omega (p_m - p_\infty)(1-\rho_{1,\infty}-\rho_{2,\infty}) \, dxdt \right| & \\ + \left| \int_0^T \int_\Omega p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m}) \, dxdt \right|. &\end{aligned}$$

The first term on the right hand side goes to 0 because p_m converges weakly to p_∞ in $L^2((0,T)\times\Omega)$ and $1-\rho_{1,\infty}-\rho_{2,\infty} \in L^\infty((0,T)\times\Omega) \subset L^2((0,T)\times\Omega)$. To handle the second term, we split the integral in two parts.

$$\begin{aligned}\left| \int_0^T \int_\Omega p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m}) \, dxdt \right| & \\ \leqslant \left| \iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m}) \, dxdt \right| & \\ + \left| \iint_{\{\rho_{1,m}+\rho_{2,m}\leqslant 1\}} p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m}) \, dxdt \right| &\end{aligned}$$

Since $\rho_{1,\infty}+\rho_{2,\infty}\leqslant 1$ and $p_m\geqslant 0$,

$$\begin{aligned}\left|\iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m})\,dxdt\right| &\\ \leqslant \left|\iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} p_m(1-\rho_{1,m}-\rho_{2,m})\,dxdt\right| &\\ \leqslant \|p_m\|^{1/2}_{L^2((0,T)\times\Omega)}\left(\int_0^T\int_\Omega(\rho_{1,m}+\rho_{2,m}-1)_+^2\,dxdt\right)^{1/2} &\\ \leqslant \frac{C}{m^{1/2}}\to 0,&\end{aligned}$$

because of proposition 6.17 and (6.32). Moreover, on $\{\rho_{1,m}+\rho_{2,m}\leqslant 1\}$, $p_m\leqslant 1$, so we obtain

$$\begin{aligned}\left|\iint_{\{\rho_{1,m}+\rho_{2,m}\leqslant 1\}} p_m(\rho_{1,\infty}+\rho_{2,\infty}-\rho_{1,m}-\rho_{2,m})\,dxdt\right| &\\ \leqslant \|\rho_{1,\infty}-\rho_{1,m}\|_{L^1((0,T)\times\Omega)}+\|\rho_{2,\infty}-\rho_{2,m}\|_{L^1((0,T)\times\Omega)}\to 0,&\end{aligned}$$

by L^1-convergence. Now, to conclude the proof, since $p_\infty(1-\rho_{1,\infty}-\rho_{2,\infty})\geqslant 0$, we just have to prove that

$$\int_0^T\int_\Omega p_m(1-\rho_{1,m}-\rho_{2,m})\,dxdt\to 0.$$

And since we already know that

$$\iint_{\{\rho_{1,m}+\rho_{2,m}\geqslant 1\}} p_m(1-\rho_{1,m}-\rho_{2,m})\,dxdt\to 0,$$

we just have to prove that

$$\iint_{\{\rho_{1,m}+\rho_{2,m}\leqslant 1\}} p_m(1-\rho_{1,m}-\rho_{2,m})\,dxdt\to 0.$$

Since $p_m:=\frac{m}{m-1}(\rho_{1,m}+\rho_{2,m})^{m-1}$, we have

$$0\leqslant\iint_{\{\rho_{1,m}+\rho_{2,m}\leqslant 1\}} p_m(1-\rho_{1,m}-\rho_{2,m})\,dxdt\leqslant\frac{m}{m-1}\left(|\Omega|^{1/m}\|\rho_{1,m}+\rho_{2,m}\|_{L^m}^{m-1}-\|\rho_{1,m}+\rho_{2,m}\|_{L^m}^m\right),$$

which concludes the proof.

□

6.5 Simulations

To end this chapter, we propose to use the algorithm presented in chapter 3 to give numerical simulations. The first system we study is transport equation with porous media congestion,

$$\partial_t\rho_i-\operatorname{div}(\rho_i\nabla F_m'(\alpha_1\rho_1+\alpha_2\rho_2))-\operatorname{div}(\rho_i\nabla V_i)=0,\ i=1,2, \tag{6.33}$$

which, as we saw, is the gradient flow in Wasserstein space for the energy

$$\mathcal{E}(\rho_1,\rho_2):=\int_\Omega V_1\rho_1+\int_\Omega V_2\rho_2+\int_\Omega F_m(\alpha_1\rho_1+\alpha_2\rho_2).$$

Arguing as in section 3.2, setting $\phi=(\phi_1,\phi_2)$, $(D\phi_1,D\phi_2):=(\partial_t\phi_1,\nabla\phi_1,\partial_t\phi_2,\nabla\phi_2)$, $q=(q_1,q_2)=(a_1,b_1,c_1,a_2,b_2,c_2)$, $\sigma=(\sigma_1,\sigma_2)=((\mu_1,m_1,\widetilde{\mu}_1),(\mu_2,m_2,\widetilde{\mu}_2))$ and defining the convex set K by

3.12, one can rewrite the discretization of the JKO scheme with $\mathcal{E}$ as a saddle-point problem for the augmented Lagrangian

$$\begin{aligned}L_r(\phi,q,\sigma) &= \sum_{i=1}^{2}\int_\Omega \phi_i(0,x)\rho_i^k(x)\mathrm{d}x + \sum_{i=1}^{2}\int_\Omega \chi_K(a_i(t,x),b_i(t,x))\mathrm{d}x\mathrm{d}t\\ &+\sum_{i=1}^{2}\int_\Omega \Big((\mu_i,m_i)\cdot(D\phi_i-(a_i,b_i))+\frac{r}{2}|D\phi_i-(a_i,b_i)|^2\Big)\mathrm{d}x\mathrm{d}t\\ &+\sum_{i=1}^{2}\int_\Omega \Big(\frac{r}{2}|\phi_i(1,x)+c_i(x)|^2\mathrm{d}x-(\phi_i(1,x)+c_i(x))\widetilde{\mu}_i(x)\Big)\mathrm{d}x\\ &+\tau\mathcal{E}^*\Big(\frac{c_1}{\tau},\frac{c_2}{\tau}\Big).\end{aligned}$$

One can then again use ALG2 for this Lagrangian, note that since the coupling between the two species only appears in term $\tau\mathcal{E}^*\Big(\frac{c_1}{\tau},\frac{c_2}{\tau}\Big)$, the only significant difference is in the proximal problem in the variables $c=(c_1,c_2)$ of the second step of the algorithm. Taking $r=1$ for simplicity and using remark 3.1, we see that the two-populations analogue of 3.26 simply takes the form

$$\begin{aligned}(c_1^{n+1}(x),c_2^{n+1}(x)) &= \mathrm{prox}_{\tau\mathcal{E}^*(x,\frac{\cdot}{\tau})}\Big(-\phi_1^{n+1}(1,x)+\widetilde{\mu}_1^n(x),-\phi_2^{n+1}(1,x)+\widetilde{\mu}_2^n(x)\Big)\\ &= \Big(-\phi_1^{n+1}(1,x)+\widetilde{\mu}_1^n(x),-\phi_2^{n+1}(1,x)+\widetilde{\mu}_2^n(x)\Big)\\ &-\mathrm{prox}_{\tau\mathcal{E}(x,.)}\Big(-\phi_1^{n+1}(1,x)+\widetilde{\mu}_1^n(x),-\phi_2^{n+1}(1,x)+\widetilde{\mu}_2^n(x)\Big)\end{aligned}$$

Then setting $V_i=V_i(x)$ and $\overline{c}_i=-\phi_i^{n+1}(1,x)+\widetilde{\mu}_i^n(x)$, $\mathrm{prox}_{\tau\mathcal{E}(x,.)}(\overline{c}_1,\overline{c}_2)$ is obtained by solving

$$\inf\Big\{\sum_{i=1}^{2}\frac{1}{2}(c_i-\overline{c}_i)^2+\tau V_ic_i+\tau F_m(\alpha_1c_1+\alpha_2c_2)\ :\ c_i\geq 0\Big\}.$$

and the solution is given by

$$\mathrm{prox}_{\tau\mathcal{E}(x,.)}(\overline{c}_1,\overline{c}_2)=\left(\overline{c}_1-(\overline{c}_1-\tau V_1-\tau\alpha_1F_m'(u))_+,\overline{c}_2-(\overline{c}_2-\tau V_2-\tau\alpha_2F_m'(u))_+\right),$$

where u is the nonnegative root of

$$c-\alpha_1(\overline{c}_1-\tau V_1)-\alpha_2(\overline{c}_2-\tau V_2)+\tau(\alpha_1+\alpha_2)F_m'(c)=0.$$

Figure 6.1 represents two populations crossing each over subject to porous media congestion with $\alpha_1=\alpha_2=1$ and $m=50$. We remark that the two populations have the same behaviour and when they cross each over, the density has to spread. In figure 6.2, we study the same behaviour but subject to the porous medium constraint on $\rho_1+2\rho_2$. We can see that the population where the constraint plays a higher role, ρ_2, has to deviate in order to let pass ρ_1 through.

In the two populations crowd motion model with linear diffusion, $\mathrm{prox}_{\tau\mathcal{E}(x,.)}(\overline{c}_1,\overline{c}_2)$ is obtained by solving

$$\inf\Big\{\sum_{i=1}^{2}\frac{1}{2}(c_i-\overline{c}_i)^2+\tau(c_i\log(c_i)+V_ic_i)\ :\ c_i\geq 0,\ \alpha_1c_1+\alpha_2c_2\leq 1\Big\}$$

whose solution is again quasi explicit. More precisely

$$\mathrm{prox}_{\tau\mathcal{E}(x,.)}(\overline{c}_1,\overline{c}_2)=(\phi_\tau(\overline{c}_1-\tau V_1),\phi_\tau(\overline{c}_2-\tau V_2))$$

if $\alpha_1\phi_\tau(\overline{c}_1-\tau V_1)+\alpha_2\phi_\tau(\overline{c}_2-\tau V_2)\leq 1$ where $\phi_\tau(\alpha)$ is the positive root of

$$c+\tau(\log c+1)=\alpha$$

and

$$\mathrm{prox}_{\tau\mathcal{E}(x,.)}(\overline{c}_1,\overline{c}_2)=\left(\psi_\tau\left(\overline{c}_1-\tau V_1-\frac{\alpha_1}{\alpha_2}(\overline{c}_2-\tau V_2)\right),\frac{1-\alpha_1\psi_\tau\left(\overline{c}_1-\tau V_1-\frac{\alpha_1}{\alpha_2}(\overline{c}_2-\tau V_2)\right)}{\alpha_2}\right)$$

where $\psi_\tau(\alpha)$ is the root in $(0,1/\alpha_1)$ of

$$\left(1+\left(\frac{\alpha_1}{\alpha_2}\right)^2\right)c+\tau\log(c)-\tau\frac{\alpha_1}{\alpha_2}\log\left(\frac{1-\alpha_1 c}{\alpha_2}\right)=\alpha+\frac{\alpha_1}{\alpha_2^2},$$

otherwise. This proximal computation therefore only involves scalar monotone equations and is therefore not more complicated than what we saw in the case of a single equation.

In figure 6.3, we see two populations which cross each other. When they start to cross each other at time $t=0.05$, we remark that the density of ρ_1 and ρ_2 decrease and the sum is saturated. Figure 6.4 represents the same case but in this example the second population is bigger than the other one, meaning that we can allow twice less people of ρ_2 than ρ_1 at every space point x. In figures 6.5 and 6.6, we add a obstacle in the middle. This can be done using a potential with very high value in this area.

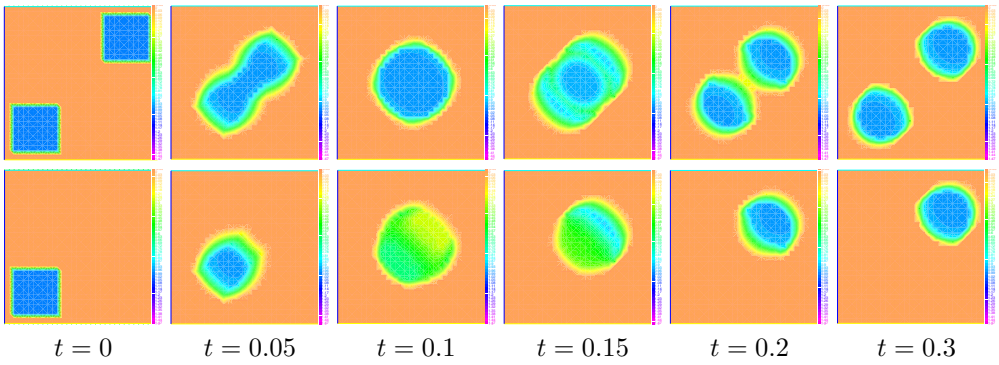


Figure 6.1: *Evolution of two species crossing each other with porous media congestion, $m=50$. Top row: display of $\rho_1+\rho_2$. Bottom row: display of ρ_1.*

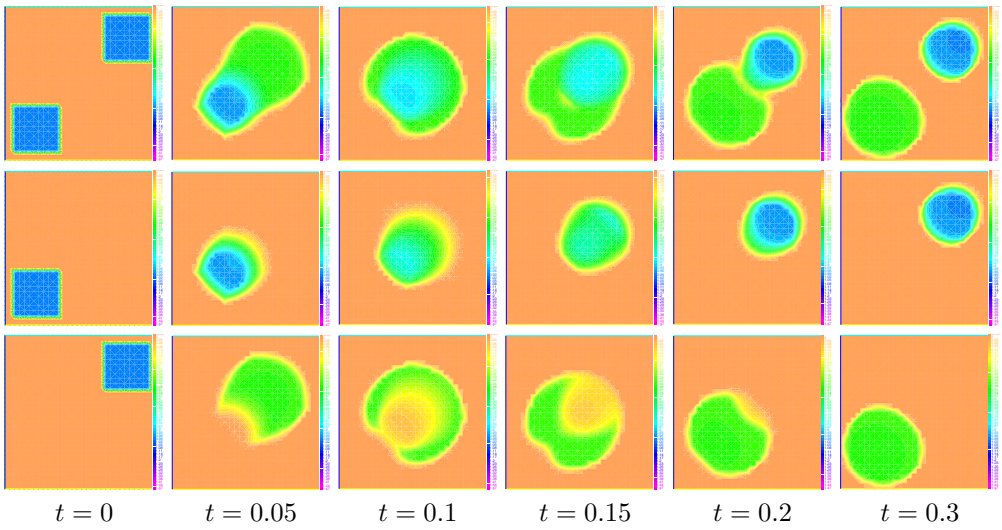


Figure 6.2: *Evolution of two species crossing each other with weighted porous media congestion,* $(\rho_1 + 2\rho_2)^m$, $m = 50$. *Top row: display of* $\rho_1 + \rho_2$. *Middle row: display of* ρ_1. *Bottom row: display of* ρ_2.

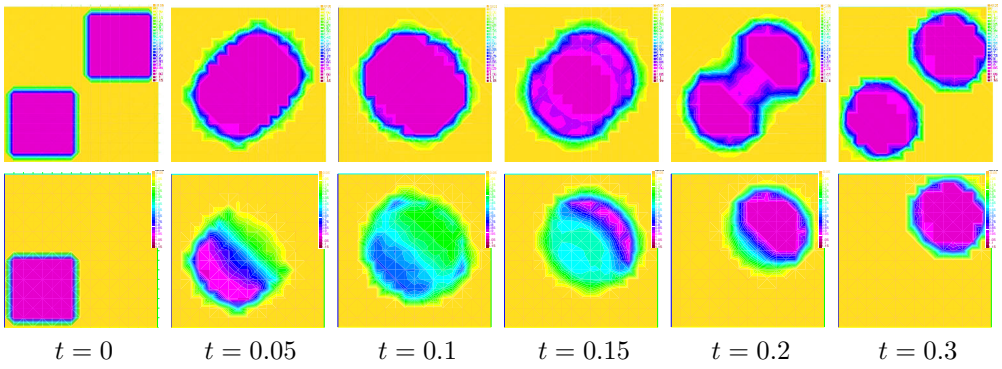


Figure 6.3: *Evolution of two species crossing each other with density constraint. Top row: display of* $\rho_1 + \rho_2$. *Bottom row: display of* ρ_1.

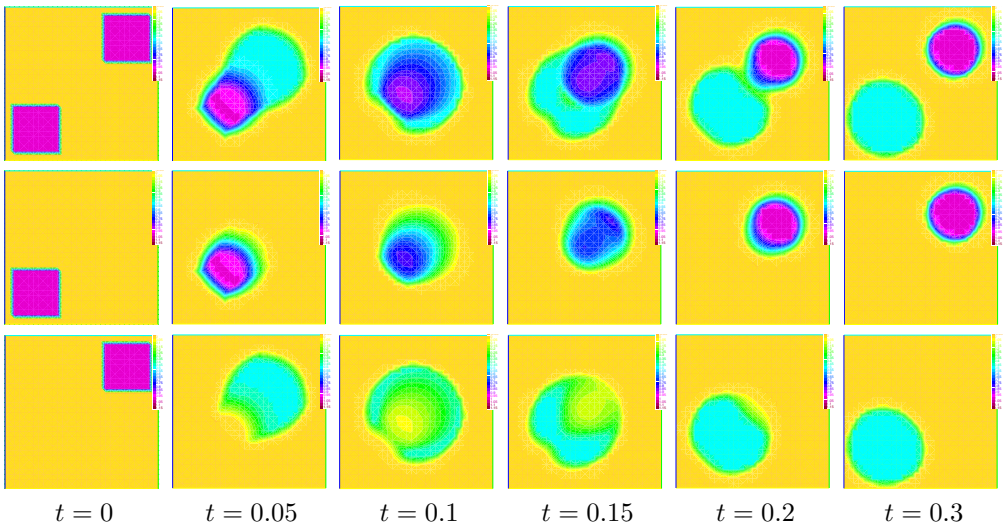


Figure 6.4: *Evolution of two species crossing each other with weighted density constraint,* $\rho_1 + 2\rho_2 \leqslant 1$. *Top row: display of* $\rho_1 + \rho_2$. *Middle row: display of* ρ_1. *Bottom row: display of* ρ_2.

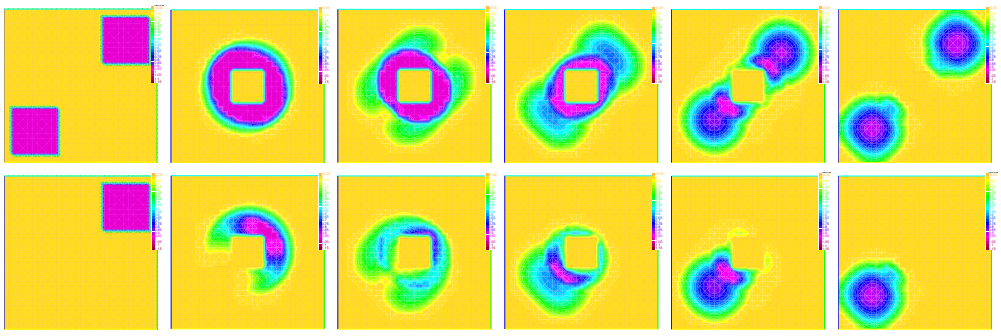


Figure 6.5: *Evolution of two species crossing each other with density constraint and an obstacle. Top row: display of* $\rho_1 + \rho_2$*. Bottom row: display of* ρ_1*.*

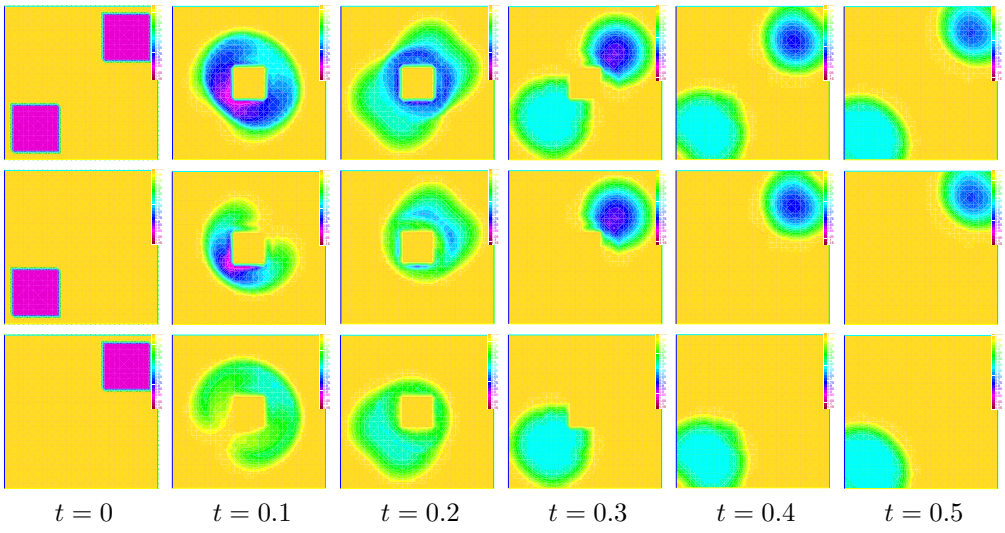


Figure 6.6: *Evolution of two species crossing each other with weighted density constraint,* $\rho_1 + 2\rho_2 \leqslant 1$*, and an obstacle. Top row: display of* $\rho_1 + \rho_2$*. Middle row: display of* ρ_1*. Bottom row: display of* ρ_2*.*

Chapter 7

Nonlinear reaction-diffusion systems

In the previous chapters, we investigated existence of solutions for equations or systems with mass conservation. Nevertheless in biology, for example for diffusive prey-predator models, this property may not be satisfied. In this chapter, we study two different methods to extend the Wasserstein gradient flow to general parabolic equations with reaction terms.

The first one was proposed by Kinderlherer and Walkington in [70]. They presented a numerical algorithm based on a splitting JKO scheme. Given a time step h and the solution at step hk, ρ_h^k, they introduced an intermediary step where the mass is updated, $\tilde{\rho}_h^{k+1} = \rho_h^k + hf(\rho_h^k)$ and then they use the usual JKO scheme

$$\rho_h^{k+1} \in \underset{\rho, |\rho| = |\tilde{\rho}_h^{k+1}|}{\operatorname{argmin}} \left\{ \frac{1}{2h} W_2^2(\rho, \tilde{\rho}_h^{k+1}) + \mathcal{F}(\rho) \right\}.$$

The motivation of this update comes from the equation we want to solve. If we integrate in time and in space

$$\partial_t \rho - \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}(\rho)\right) = f(\rho),$$

with Neumann boundary condition, $\nabla \frac{\delta \mathcal{F}}{\delta \rho}(\rho) \cdot \nu = 0$ on $\partial\Omega$, between two times $t+h, t$, then

$$\int_\Omega \rho(t+h, x)\, dx = \int_\Omega \rho(t, x)\, dx + \int_t^{t+h} \int_\Omega f(\rho(s, x))\, dx ds,$$

which implies the formula if we assume that ρ is constant in $[t, t+h)$.

In [115, 103], Petrelli and Tudorascu proved the convergence of this scheme for general diffusion problems with drift, diffusion coefficient and forcing term which may be explicitly time dependent. The proof of the convergence of this scheme is based on a maximum principle and the Fréchet-Kolmogorov theorem. The first part is an extension of the maximum principle developped by Otto in [97], to deal with potential terms. The Fréchet-Kolmogorov theorem gives strong convergence which allows to pass to the limit in the Euler-Lagrange equation.

The advantage of this scheme is that the intermediary step permits to have a good control on the L^∞-bound. In the first part of this chapter, we investigate the case of several species with interactions in reaction terms. The proof is based on the splitting scheme introduced in [70]. First, we extend the proof of the maximum principle of Petrelli and Tudorascu to systems. Moreover, we propose a new method to recover the strong convergence. We use the *bounded Lipschitz distance* to obtain time-compactness and the space compactness is recovered using the Euler-Lagrange equation. Then, the strong convergence is obtained using an extention of Aubin-Lions lemma due to Rossi and Savaré in [107]. Finally, several extensions of (7.1) are presented and we give some numerical results based on the augmented Lagrangian method introduced in [17] and presented in chapter 3.

The second method is based on a splitting method for the Wasserstein-Fisher-Rao metric. This metric was introduced in [43, 44, 71, 80, 81] by three different teams at the same time.

It is an extension of the Wasserstein metric allowing for mass variations. In [63], Galloüet and Monsaingeon proposed a splitting method to solve gradient flows for this metric. They remark that, infinitesimally, the Wasserstein-Fisher-Rao metric should be the orthogonal sum of the Wasserstein metric and the Fisher-Rao metric, which leads to a natural splitting scheme to study Wasserstein-Fisher-Rao gradient flows. For a functional $\mathcal{E}$, they split one-step of minimizing scheme for $\mathcal{E}$ between one substep of JKO scheme with $\mathcal{E}$ and one substep of minimizing scheme for $\mathcal{E}$ with respect to the Fisher-Rao metric

$$FR(\rho_0,\rho_1) := \min_{(\rho_t,r_t)\in\mathcal{A}_{FR}[\rho_0,\rho_1]} \int_0^1 \int_\Omega |r_t|^2 \, d\rho_t(x)dt = 4\int_\Omega \left|\sqrt{\frac{d\rho_0}{d\lambda}} - \sqrt{\frac{d\rho_1}{d\lambda}}\right|^2 d\lambda,$$

where the admissible set $\mathcal{A}_{FR}[\rho_0,\rho_1]$ is the set of curves $t\in[0,1]\mapsto(\rho_t,r_t)$ such that $t\mapsto\rho_t$ is weakly continuous with values in $\mathcal{M}^+(\Omega)$ with endpoints $\rho_{t=0}=\rho_0$ and $\rho_{t=1}=\rho_1$ and

$$\partial_t\rho_t = \rho_t r_t \text{ in } \mathcal{D}'((0,1)\times\Omega),$$

and λ is any reference measure such that ρ_0 and ρ_1 are both absoltely continuous with respect to λ. In [63], they proved that discrete solutions constructed using this scheme converge to the solution of

$$\partial_t\rho - \operatorname{div}(\rho\nabla(F'(\rho)+V)) = -\rho(F'(\rho)+V),$$

if

$$\mathcal{E}(\rho) := \int F(\rho) + \int V\rho.$$

However, this class of equation is very restrictive because there is a strong link between the diffusion and the reaction since we minimize the same energy at each substep. In the second section of this chapter, we propose, in collaboration with Galloüet and Monsaingeon, an extension of this method to treat a larger class of equations where there is not link between the diffusion and the reaction. Moreover, we apply this method to solve the tumor growth model introduced in [101].

7.1 Variational principle for reaction-diffusion systems

In this work, we investigate existence of solutions to general parabolic systems without cross-diffusion but with interactions in the reaction terms,

$$\begin{cases} \partial_t\rho_1 - \Delta P_1(\rho_1) - \operatorname{div}(\rho_1\nabla V_1) = f_1(\rho_1,\rho_2) \\ \partial_t\rho_2 - \Delta P_2(\rho_2) - \operatorname{div}(\rho_2\nabla V_2) = f_2(\rho_1,\rho_2), \end{cases} \tag{7.1}$$

on $\mathbb{R}^+\times\Omega$ with initial condition, $\rho_i(0,\cdot)=\rho_{i,0}$. In the sequel, Ω is an open bounded subset of $\mathbb{R}^n$ and (7.1) is endowed with no flux boundary conditions.

The section is organized as follows. In section 1, we recall some facts on Wasserstein spaces and we state our main result. Section 2 is devoted to a time-compactness result based on the maximum principle introduced in [103]. In section 3, we recover the Euler-Lagrange equations and prove the space-compactness estimate. Then we pass to the limit in section 4 using an extension of Aubin-Lions lemma. We present in section 5 several extensions of this result and give a uniqueness argument in section 6. The final section 7 is devoted to numerical simulations.

7.1.1 Preliminaries and main result

Wasserstein space and bounded Lipschitz distance In this paragraph, we recall some results from optimal transport theory and we refer the reader to the textbooks of Villani [116, 117], Ambrosio, Gigli and Savaré [4] and Santambrogio [112] for a detailed exposition.

Let Ω be an open bounded subset of $\mathbb{R}^n$, we denote $\mathcal{M}_m(\Omega)$ the set of nonnegative integrable functions with mass m. Then the Wasserstein metric of order 2 is defined for all $\rho,\mu\in\mathcal{M}_m(\Omega)$ by

$$W_{2,m}^2(\rho,\mu) := \inf_{\gamma\in\Pi_m(\rho,\mu)} \iint_{\Omega\times\Omega} |x-y|^2 \, d\gamma(x,y),$$

where the set of admissible plans is

$$\Pi_m(\rho,\mu) := \left\{\gamma \in \mathcal{M}_m(\Omega\times\Omega) \,:\, \pi_{x\#}\gamma = \rho,\, \pi_{y\#}\gamma = \mu\right\}.$$

We recall also that the 1-Wasserstein distence between $\rho,\mu \in M_m(\Omega)$ is given by

$$W_{1,m}(\rho,\mu) := \inf_{\gamma\in\Pi_m(\rho,\mu)} \iint_{\Omega\times\Omega} |x-y|\, d\gamma(x,y),$$

and the well-known Kantorovich duality states that $W_{1,m}$ can be also expressed as

$$W_{1,m}(\rho,\mu) := \sup\left\{\int_\Omega \varphi\, d(\rho-\mu) \,:\, \varphi\, 1-\text{Lipschitz on } \Omega\right\}.$$

Since we have to deal with variation of mass, we also introduce the *bounded Lipschitz distance* defined by, for all $\rho,\mu \in L^1(\Omega;\mathbb{R}^+)$,

$$d_{BL^*}(\rho,\mu) := \sup\left\{\int_\Omega \phi\, d(\rho-\mu) \,:\, \|\phi\|_{L^\infty(\Omega)} + \|\phi\|_{Lip} \leqslant 1\right\}, \tag{7.2}$$

where

$$\|\phi\|_{Lip} := \sup_{x\neq y} \frac{|\phi(x)-\phi(y)|}{|x-y|}.$$

We remark that for all $\rho,\mu \in L^1(\Omega)\cap L^\infty(\Omega)$,

$$d_{BL^*}(\rho,\mu) \leqslant \|\rho-\mu\|_{L^1(\Omega)} \leqslant |\Omega|\|\rho-\mu\|_{L^\infty(\Omega)}, \tag{7.3}$$

and, if ρ and μ have the same mass m, we have

$$d_{BL^*}(\rho,\mu) \leqslant W_1(\rho,\mu) \leqslant mW_2(\rho,\mu). \tag{7.4}$$

Assumptions and main result The nonlinear diffusion terms are given by continuous strictly convex superlinear functions $F_i : \mathbb{R}^+ \to \mathbb{R}$ of class $\mathcal{C}^2((0,+\infty))$ which satisfy $F_i(0) = 0$ and define $P_i(x) := xF_i'(x) - F_i(x)$ as the pressure associated to F_i. Since F_i' is strictly increasing, continuous on $(0,+\infty)$ and $F_i'((0,+\infty)) = (F_i'(0),+\infty)$, F_i' is inversible and $(F_i')^{-1}$ is increasing and continuous. In addition, F_i verifies one of the following technical assumptions, for all r large enough and for all $\alpha > 1, \beta > 0$, $x,y \in \mathbb{R}^+$, either

$$\alpha F_i''(\alpha r+\beta) - F_i''(r) \leqslant 0 \text{ and } (F_i)^{-1}(x+y) \leqslant A_i(F_i)^{-1}(x)(F_i)^{-1}(y),, \tag{7.5}$$

where A_i is a positive constant, or

$$F_i''(r) \geqslant 1 \text{ and } \alpha F_i''(\alpha r+\beta) - F_i''(r) \geqslant 0. \tag{7.6}$$

Moreover, we assume that P_i is continuous on $[0,+\infty)$. Define $\mathcal{F}_i : \mathcal{M}(\Omega) \to \mathbb{R}$ by

$$\mathcal{F}_i(\rho) := \begin{cases} \int_\Omega F_i(\rho(x))\, dx & \text{if } \rho \ll \mathcal{L}^n, \\ +\infty & \text{otherwise.} \end{cases}$$

Typical examples of energies we have in mind are $F_i(x) := x\log(x)$, which verifies (7.5) and gives a linear diffusion driven by the laplacian, and $F_i(x) := x^m$ $(m \geqslant 2)$, which verifies (7.6) and gives porous media diffusion.

The reaction terms are given by continuous functions $f_i : [0,+\infty)\times[0,+\infty) \to \mathbb{R}$ such that there exsits $K > 0$ such that, for all $x_1, x_2 \in [0,+\infty)$,

$$f_i(x_1,x_2) \leqslant K(1+x_1+x_2), \tag{7.7}$$

$$f_1(x_1,x_2) \geqslant -K(1+x_2)x_1, \qquad f_2(x_1,x_2) \geqslant -K(1+x_1)x_2. \tag{7.8}$$

For example, these assumptions can been verified for model of diffusive prey-predator system (cf [96]) with the choice

$$f_1(\rho_1,\rho_2) = A\varrho\left(1-\frac{\varrho}{K}\right) - \frac{B\varrho\rho_2}{1+E\varrho}, \text{ and } f_2(\rho_1,\rho_2) = \frac{C\varrho\rho_2}{1+E\varrho} - D\rho_2, \tag{7.9}$$

where $\varrho = \min(K,\rho_1)$, $A,B,C,D,E,K>0$. In this model, ρ_1 is the density of preys and ρ_2 is the density of predators. Another model is the one of pattern formation in animal coating (cf [96]) where f_1 and f_2 are given by

$$f_1(\rho_1,\rho_2) = A-\rho_1-h(\rho_1,\rho_2), \text{ and } f_2(\rho_1,\rho_2) = C(B-\rho_2)-h(\rho_1,\rho_2),$$

where $h(\rho_1,\rho_2) = \frac{D\rho_1\rho_2}{1+\rho_1+E\rho_1}$.

Remark 7.1. *Assumption* (7.7) *can be replaced by, for all* $\rho_1,\rho_2 \in L^\infty(\Omega)$,

$$f_i(\rho_1,\rho_2) \leqslant K(1+\|\rho_1\|_{L^\infty(\Omega)} + \|\rho_2\|_{L^\infty(\Omega)}).$$

This modification allows to consider nonlocal reaction terms like

$$f_i(\rho_1,\rho_2) = \int_\Omega W_{i,1}(x,y)\rho_1(y)\,dy + \int_\Omega W_{i,2}(x,y)\rho_2(y)\,dy,$$

with $W_{i,j} \geqslant 0$ *and* $W_{i,j} \in L^\infty_x(\Omega, L^p_y(\Omega))$.

Definition 7.2. *We say that* $(\rho_1,\rho_2) : [0,+\infty) \to L^1(\Omega,\mathbb{R}^+)^2$ *is a weak solution of* (7.1) *if, for all* $T<+\infty$, $\rho_i \in L^\infty((0,T),L^1(\Omega))$, $\nabla P_i(\rho_i) \in L^2((0,T)\times\Omega)$, $f_i(\rho_1,\rho_2) \in L^\infty((0,T),L^1(\Omega))$ *and*

$$\int_0^\infty \int_\Omega \partial_t\phi\rho_i - \nabla\phi\cdot\nabla P_i(\rho_i) - \nabla\phi\cdot\nabla V_i\rho_i + \phi f_i(\rho_1,\rho_2)\,dxdt = -\int_\Omega \phi(0,x)\rho_{i,0}(x)\,dx, \tag{7.10}$$

for every $\phi \in \mathcal{C}^\infty_c([0,+\infty)\times\mathbb{R}^n)$.

Now we state the main result of this section,

Theorem 7.3. *Assume that* $\rho_{i,0} \in L^1(\Omega,\mathbb{R}^+)$ *satisfies* $\|\rho_{0,i}\|_{L^\infty(\Omega)} \leqslant M_0$ *and* $V_i \in W^{1,\infty}(\Omega)$, *then system* (7.1) *admits at least one weak solution.*

The proof of this theorem is given in the next section and is based on a splitting scheme. This scheme was proposed in [70] by Kinderlehrer and Walkington to present a numerical algorithm to solve parabolic equations with forcing terms. Then in [115, 103], Petrelli and Tudorascu proved the convergence of this scheme for one equation with nonnegative forcing term.

Let h a time step, for $i\in\{1,2\}$, we define two sequences, $(\rho^k_{i,h})_k$ and $(\tilde\rho^k_{i,h})_k$ such that $\rho^0_{i,h} = \tilde\rho^0_{i,h} = \rho_{i,0}$, and for all $k \geqslant 0$,

$$\begin{cases} \tilde\rho^{k+1}_{i,h} = \rho^k_{i,h} + hf_i(\rho^k_{1,h},\rho^k_{2,h}), \\ \rho^{k+1}_{i,h} \in \operatorname{argmin}_{\rho\in\mathcal{M}_{m^{k+1}_i}} \frac{1}{2h}W^2_{2,m_{k+1}}(\rho,\tilde\rho^{k+1}_{i,h}) + \mathcal{F}_i(\rho) + \mathcal{V}_i(\rho), \end{cases} \tag{7.11}$$

where $m^{k+1}_i = \int_\Omega \tilde\rho^{k+1}_{i,h}(x)\,dx$. In the sequel, we fix a time $T>0$ and we note $t_k := hk$ for all k,h such that $hk \leqslant T$. For h fixed, we note $N := \lfloor \frac{T}{N} \rfloor$. Without loss of generality, we can assume that $T = Nh$. Then we define the piecewise interpolation $\rho_{i,h} : \mathbb{R}^+ \to L^1(\Omega;\mathbb{R}^+)$ by

$$\rho_{i,h}(t) := \rho^{k+1}_{i,h}, \qquad \text{if } t\in(hk,h(k+1)].$$

7.1.2 Properties of sequences and time compactness

First, we have to prove that the sequences $(\rho^k_{i,h})_k$ and $(\tilde\rho^k_{i,h})_k$ are well defined at least for h small enough. Let R be defined by $R := \max\left\{\|V_1\|_{L^\infty(\Omega)}, \|V_2\|_{L^\infty(\Omega)}\right\}$. In the sequel, we define M_T by

$$M_T := \max((F_1)^{-1}(F_1'((2M_0+4R+1)e^{\tilde K T})+2R), (F_2)^{-1}(F_2'((2M_0+4R+1)e^{\tilde K T})+2R)),$$

where $\tilde{K} = \max(K(1+\tilde{C}), 2K)$ and $\tilde{C} = \max(C_1, C_2)$ with

$$C_2 := \begin{cases} A_2(F_2')^{-1}(2R) & \text{if } F_2 \text{ satisfies (7.5)}, \\ 1 & \text{if } F_2 \text{ satisfies (7.6)}, \end{cases} \quad C_1 := \begin{cases} A_1(F_1')^{-1}(2R) & \text{if } F_1 \text{ satisfies (7.5)}, \\ 1 & \text{if } F_1 \text{ satisfies (7.6)}. \end{cases}$$

We start to recall a general maximum principle for the JKO scheme proved in [103].

Lemma 7.4. *Let $\mu \in L^1(\Omega, \mathbb{R}^+)$ ($\mu(\Omega) = m$) and ρ one minimizer in $\mathcal{M}_m$ of*

$$\mathcal{E}_h[\mu] \,:\, \rho \mapsto \frac{1}{2h} W_{2,m}^2(\rho, \mu) + \int_\Omega (F(\rho) + V\rho),$$

where $F : \mathbb{R}^+ \to \mathbb{R}$ is a convex continuous function of class $\mathcal{C}^2((0,+\infty))$ and $V \in W^{1,\infty}(\Omega)$. If there exists $C > 0$ such that $\mu \leqslant (F')^{-1}(C - V)$ then $\rho \leqslant (F')^{-1}(C - V)$.

Remark 7.5. *We remark that $(F')^{-1}(C - V)$ is a steady solution of the equation*

$$\partial_t \rho = \Delta P(\rho) - \operatorname{div}(\rho \nabla V) = \operatorname{div}(\rho \nabla (F'(\rho) + V).$$

Since we have a comparaison principle for this equation, at the continuous level the conclusion of lemma 7.4 is easy. Lemma 7.4 proves this result at the discrete level.

Proof. This lemma is an adaptation of the one of Otto in [97] and Agueh in [1] to handle the potential energy and the proof comes from [103]. We note $U := (F')^{-1}(C - V)$. Assume that $E := \{\rho > U\}$ has positive measure. Take $\gamma \in \Pi(\mu, \rho)$ to be the optimal transport plan, then $\gamma(E^c \times E) > 0$. Indeed, otherwise we have

$$\int_E U(x)\,dx < \int_E \rho(x)\,dx = \gamma(\Omega \times E) = \gamma(E \times E) \leqslant \gamma(E \times \Omega) = \int_E \mu(x)\,dx \leqslant \int_E U(x)\,dx,$$

which gives a contradiction. Consider $\nu := \mathbf{1}_{E^c \times E}\gamma$ and denote v_0 and v_1 its marginals. We remark that $v_0(E) = v_1(E^c) = 0$. For ε small enough, we define $\rho_\varepsilon := \rho + \varepsilon(v_0 - v_1) \in \mathcal{M}_\mu$ (because $v_0(\Omega) = \gamma(E^c \times E) = v_1(\Omega)$) and γ_ε by

$$\iint_{\Omega\times\Omega} \xi(x,y)\,d\gamma_\varepsilon(x,y) := \iint_{\Omega\times\Omega} \xi(x,y)\,d\gamma(x,y) + \varepsilon \iint_{E^c\times E} [\xi(x,x) - \xi(x,y)]\,d\gamma(x,y).$$

Since $\gamma \in \Pi(\mu, \rho)$ and ν has for marginals v_0 and v_1, we obtain that γ_ε defines a transport plan between μ and ρ_ε. We want to prove that ρ_ε verifies

$$\mathcal{E}_h[\mu](\rho_\varepsilon) < \mathcal{E}_h[\mu](\rho),$$

which gives a contradiction by definition of ρ.

We first remark that, by definition of γ_ε,

$$W_{2,m}^2(\rho_\varepsilon, \mu) - W_{2,m}^2(\rho, \mu) \leqslant -\varepsilon \iint_{E^c\times E} |x-y|^2\,d\gamma(x,y) < 0,$$

because $\gamma(E^c \times E) > 0$. So we have

$$\begin{aligned} \mathcal{E}_h[\mu](\rho_\varepsilon) &< \mathcal{E}_h[\mu](\rho) + \int_\Omega [F(\rho_\varepsilon) - F(\rho) + V\rho_\varepsilon - V\rho]\,dx \\ &< \mathcal{E}_h[\mu](\rho) + \beta(\varepsilon), \end{aligned}$$

where

$$\beta(\varepsilon) := \int_\Omega [F(\rho + \varepsilon(v_0 - v_1)) - F(\rho) + \varepsilon V(v_0 - v_1)]\,dx.$$

To conclude the proof, we have to show that β is negative for ε small enough. We have

$$\beta'(0) = -\int_\Omega (v_0 - v_1)(F'(\rho) + V)\,dx.$$

Since F' is increasing and $v_0(E) = v_1(E^c) = 0$, the decomposition

$$\beta'(0) = -\int_E v_1[F'(\rho) - F'(U)]\,dx + \int_{E^c} v_0[F'(\rho) - F'(U)]\,dx - \int_\Omega (v_0 - v_1)[V + F'(U)]\,dx$$

shows that $\beta'(0) < 0$. Indeed, the last term vanishes because by definition of U, $V + F'(U) = C$ and v_0 and v_1 have the same mass. So β is decreasing for ε small enough and since $\beta(0) = 0$, we can choose ε such that, $\beta(\varepsilon) < 0$, which gives a contradiction.

□

Proposition 7.6. *For all $h \leqslant \frac{1}{K(1+M_T)}$, $(\rho_{i,h}^k)_k$ and $(\tilde\rho_{i,h}^k)_k$ are well defined. For all k such that $hk \leqslant T$, $\rho_{i,h}^k, \tilde\rho_{i,h}^k \in L^1(\Omega, \mathbb{R}^+)$ and*

$$\|\rho_{i,h}^k\|_{L^\infty(\Omega)}, \|\tilde\rho_{i,h}^k\|_{L^\infty(\Omega)} \leqslant M_T. \tag{7.12}$$

Proof. We prove by induction that for all k these sequences are well defined and

$$\|\rho_{i,h}^k\|_{L^\infty(\Omega)}, \|\tilde\rho_{i,h}^k\|_{L^\infty(\Omega)} \leqslant (F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i), \tag{7.13}$$

with

$$\begin{cases} a_{1,k+1} = (1+Kh)a_{1,k} + Kh(1 + b_{1,2}^k), \\ a_{2,k+1} = (1+Kh)a_{2,k} + Kh(1 + b_{2,1}^k), \\ a_{i,0} = M_0 + b_i, \end{cases}$$

where b_i is defined by

$$b_i := \begin{cases} 0 & \text{if } F_i \text{ verifies (7.5)}, \\ 2R & \text{if } F_i \text{ verifies (7.6)}, \end{cases}$$

and

$$b_{1,2}^k = (F_2')^{-1}(F_2'(a_{2,k}) - b_2 + 2R) \text{ and } b_{2,1}^k = (F_1')^{-1}(F_1'(a_{1,k}) - b_1 + 2R).$$

First, if $k = 0$, then if F_i satisfies (7.5),

$$\rho_{i,h}^0 = \tilde\rho_{i,h}^0 \leqslant M_0 \leqslant (F_i')^{-1}(F_i'(M_0) + R - V_i),$$

because $(F_i')^{-1}$ is increasing and $R - V_i \geqslant 0$, so we proved the claim with $b_i = 0$. Otherwise, if F_i satisfies (7.6), then $x \mapsto F_i'(x) - x$ is increasing for x large enough and then

$$\rho_{i,h}^0 = \tilde\rho_{i,h}^0 \leqslant M_0 \leqslant (F_i')^{-1}(F_i'(M_0 + b_i) - b_i + R - V_i).$$

Now assume that for $k \geqslant 0$, $\rho_{i,h}^k$ and $\tilde\rho_{i,h}^k$ are well defined and verify (7.13). Let $i \in \{1,2\}$ and $j \in \{1,2\}$ with $j \neq i$. We first prove that $\tilde\rho_{i,h}^{k+1} \geqslant 0$. Using (7.8) and (7.13),

$$\begin{aligned} \tilde\rho_{i,h}^{k+1}(x) &\geqslant \rho_{i,h}^k(x) - hK(1 + \rho_{j,h}^k(x))\rho_{i,h}^k(x) \\ &\geqslant \rho_{i,h}^k(x)(1 - hK(1 + (F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i))) \\ &\geqslant 0, \end{aligned}$$

because $(F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i) \leqslant M_T$ (see at the end of the proof) and $h \leqslant \frac{1}{K(1+M_T)}$.

Then using the definition of $\tilde\rho_{i,h}^{k+1}$, (7.7) and (7.13),

$$\begin{aligned} \tilde\rho_{i,h}^{k+1}(x) &\leqslant (1+Kh)\|\rho_{i,h}^k\|_{L^\infty(\Omega)} + hK(1 + \|\rho_{j,h}^k\|_{L^\infty(\Omega)}) \\ &\leqslant (1+Kh)(F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i) + Kh(1 + (F_j')^{-1}(F_j'(a_{j,k}) - b_j + 2R)) \\ &\leqslant (1+Kh)(F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i) + Kh(1 + b_{i,j}^k). \end{aligned}$$

It remains to show that

$$(1+Kh)(F_i')^{-1}(F_i'(a_{i,k}) - b_i + R - V_i) + Kh(1 + b_{i,j}^k) \leqslant (F_i')^{-1}(F_i'((1+Kh)a_{i,k} + Kh(1+b_{i,j}^k)) - b_i + R - V_i),$$

that is equivalent to prove

$$F_i'((1+Kh)(F_i')^{-1}(F_i'(a_{i,k})-b_i+R-V_i)+Kh(1+b_{i,j}^k))-(R-V_i) \leqslant F_i'((1+Kh)a_{i,k}+Kh(1+b_{i,j}^k))-b_i.$$

To do so, we show that

$$q : s \mapsto F_i'((1+Kh)(F_i')^{-1}(F_i'(a_{i,k})-b_i+s)+Kh(1+b_{i,j}^k))-s$$

is increasing, if F_i satisfies (7.6), or decreasing, if F_i satisfies (7.5). Since

$$\begin{aligned} q'(s) &= (1+Kh)\left[(F_i')^{-1}(F_i'(a_{i,k})-b_i+s)\right]' F_i''((1+Kh)(F_i')^{-1}(F_i'(a_{i,k})-b_i+s)+Kh(1+b_{i,j}^k))-1 \\ &= (1+Kh)F_i''((1+Kh)r+Kh(1+b_{i,j}^k))\frac{1}{F_i''(r)}-1, \end{aligned}$$

where $r=(F_i')^{-1}(F_i'(a_{i,k})-b_i+s)$. Then we obtain the result using assumption (7.5) or (7.6). Now, if F_i satisfies (7.5), then $b_i=0 \leqslant R-V_i$ and q is decreasing so

$$q(R-V_i) \leqslant q(0),$$

and if F_i satisfies (7.6), then $b_i=2R \geqslant R-V_i$ and q is increasing so

$$q(R-V_i) \leqslant q(b_i),$$

which implies that

$$\tilde{\rho}_{i,h}^{k+1} \leqslant (F_i')^{-1}(F_i'(a_{i,k+1})-b_i+R-V_i).$$

Since $\tilde{\rho}_{i,h}^{k+1} \in L^1(\Omega;\mathbb{R}^+)$, it is well-known that $\rho_{i,h}^{k+1}$ is well defined, see for example [65] or chpter 4. Then using lemma 7.4 we obtain

$$\rho_{i,h}^{k+1} \leqslant (F_i')^{-1}(F_i'(a_{i,k+1})-b_i+R-V_i),$$

that gives the result.

To conclude the proof, we have to bound $a_{i,k}$. We remark that,

$$(F_j')^{-1}(F_j'(a_{j,k})-b_j+2R) \leqslant C_j a_{j,k},$$

where

$$C_j := \begin{cases} A_j(F_j')^{-1}(2R) & \text{if } F_j \text{ satisfies (7.5)}, \\ 1 & \text{if } F_j \text{ satisfies (7.6)}, \end{cases}$$

Then,

$$a_{i,k+1} \leqslant (1+hK)a_{i,k}+hK(1+C_j a_{j,k}).$$

If we sum, we obtain

$$\begin{aligned} a_{1,k+1}+a_{2,k+1} &\leqslant (1+hK(1+\tilde{C}))(a_{1,k}+a_{2,k})+2Kh \\ &\leqslant (1+h\tilde{K})(a_{1,k}+a_{2,k})+\tilde{K}h, \end{aligned}$$

where $\tilde{C}=\max(C_1,C_2)$ and $\tilde{K}=\max(K(1+\tilde{C}),2K)$. Then

$$a_{1,k+1}+a_{2,k+1} \leqslant e^{\tilde{K}T}(2M_0+b_1+b_2+1)-1 \leqslant e^{\tilde{K}T}(2M+4R+1),$$

which concludes the proof. □

In the next lemma, we state an usual estimate in gradient flow theory, [65], and proved in this case in [103].

Lemma 7.7. *There exists a constant $C>0$ such that for all $h \leqslant \frac{1}{K(1+M_T)}$,*

$$\sum_{k=1}^{N} W_{2,m_i^k}^2(\rho_{i,h}^k,\tilde{\rho}_{i,h}^k) \leqslant Ch. \tag{7.14}$$

Proof. The proof is the same as in [103]. Since $\rho_{i,h}^k$ is optimal in the minimization problem and $\tilde{\rho}_{i,h}^k$ is a competitor, we have

$$W_{2,m_i^k}^2(\rho_{i,h}^k, \tilde{\rho}_{i,h}^k) \leqslant 2h\left(\mathcal{F}_i(\tilde{\rho}_{i,h}^k) - \mathcal{F}_i(\rho_{i,h}^k) + \mathcal{V}_i(\tilde{\rho}_{i,h}^k) - \mathcal{V}_i(\rho_{i,h}^k)\right).$$

Moreover, F_i is convex then

$$\begin{aligned}\mathcal{F}_i(\tilde{\rho}_{i,h}^k) &\leqslant \mathcal{F}_i(\rho_{i,h}^{k-1}) + \int_\Omega F_i'(\tilde{\rho}_{i,h}^k) h f_i(\rho_{1,h}^{k-1}, \rho_{2,h}^{k-1}) \\ &\leqslant \mathcal{F}_i(\rho_{i,h}^{k-1}) + hF_i'(M_T)K(1+2M_T)|\Omega|,\end{aligned}$$

because F_i' is increasing and (7.7). Moreover,

$$\mathcal{V}_i(\tilde{\rho}_{i,h}^k) - \mathcal{V}_i(\rho_{i,h}^k) \leqslant \mathcal{V}_i(\rho_{i,h}^{k-1}) - \mathcal{V}_i(\rho_{i,h}^k) + hK(1+2M_T)|\Omega|\|V_i\|_{L^\infty(\Omega)}.$$

Then we obtain

$$W_{2,m_i^k}^2(\rho_{i,h}^k, \tilde{\rho}_{i,h}^k) \leqslant 2h\left(\mathcal{F}_i(\rho_{i,h}^{k-1}) + \mathcal{V}_i(\rho_{i,h}^{k-1}) - (\mathcal{F}_i(\rho_{i,h}^k) + \mathcal{V}_i(\rho_{i,h}^k)) + Ch\right).$$

Summing over k drom 1 to N, we have

$$\sum_{k=1}^N W_{2,m_i^k}^2(\rho_{i,h}^k, \tilde{\rho}_{i,h}^k) \leqslant 2h\left(\mathcal{F}_i(\rho_{i,0}) + \mathcal{V}_i(\rho_{i,0}) - \mathcal{F}_i(\rho_{i,h}^N) - \mathcal{V}_i(\rho_{i,h}^N) + CT\right),$$

which concludes the proof because $\mathcal{F}_i$ and $\mathcal{V}_i$ are bounded from below and $\mathcal{F}_i(\rho_{i,0}) + \mathcal{V}_i$ is finite (F_i is continuous on $[0, M_T]$ and $V_i \in L^\infty(\Omega)$).

□

The next proposition gives a time compactness result on $\rho_{i,h}$.

Proposition 7.8. *For $i \in \{1, 2\}$,*

$$\lim_{\tau \searrow 0} \sup_h \int_0^{T-\tau} d_{BL^*}(\rho_{i,h}(t+\tau), \rho_{i,h}(t))\, dt = 0.$$

Proof. First, using (7.3) and (7.4), we remark that

$$\begin{aligned}d_{BL^*}(\rho_{i,h}^k, \rho_{i,h}^{k+1}) &\leqslant d_{BL^*}(\rho_{i,h}^k, \tilde{\rho}_{i,h}^{k+1}) + d_{BL^*}(\tilde{\rho}_{i,h}^{k+1}, \rho_{i,h}^{k+1}) \\ &\leqslant C\|\rho_{i,h}^k - \tilde{\rho}_{i,h}^{k+1}\|_{L^\infty(\Omega)} + Cm_i^{k+1} W_{2,m_{k+1}}(\tilde{\rho}_{i,h}^{k+1}, \rho_{i,h}^{k+1}).\end{aligned}$$

Since $m_i^k \leqslant M_T|\Omega|$, (7.14) and $\|\rho_{i,h}^k - \tilde{\rho}_{i,h}^{k+1}\|_{L^\infty(\Omega)} = \mathcal{O}(h)$, we obtain

$$\sum_{k=0}^{N-1} d_{BL^*}(\rho_{i,h}^k, \rho_{i,h}^{k+1})^2 \leqslant Ch. \tag{7.15}$$

Let $\tau > 0$, for all $h \leqslant \frac{1}{K(1+M_T)}$, there exists $j, r \in \mathbb{N}$ such that $\tau = jh + \gamma$ with $\gamma < h$ and $(j+1)h \leqslant r\tau$, then

$$
\begin{aligned}
&\int_0^{T-\tau} d_{BL^*}(\rho_{i,h}(t+\tau),\rho_{i,h}(t))\,dt \\
&\qquad = \int_0^{T-\tau} \sum_{l=0}^{j-1} d_{BL^*}(\rho_{i,h}(t+(l+1)h),\rho_{i,h}(t+lh))\,dt \\
&\qquad\qquad + \int_0^{T-\tau} d_{BL^*}(\rho_{i,h}(t+\tau),\rho_{i,h}(t+jh))\,dt \\
&\qquad \leqslant \sum_{k=0}^{N-j-1} \int_{t_k}^{t_{k+1}} \sum_{l=0}^{j-1} d_{BL^*}(\rho_{i,h}(t+(l+1)h),\rho_{i,h}(t+lh))\,dt \\
&\qquad\qquad + \sum_{k=0}^{N-j-1} \int_{t_{k+1}-\gamma}^{t_{k+1}} d_{BL^*}(\rho_{i,h}(t+\tau),\rho_{i,h}(t+jh))\,dt \\
&\qquad \leqslant h \sum_{k=0}^{N-j-1} \sum_{l=0}^{j-1} d_{BL^*}(\rho_{i,h}^{k+l+2},\rho_{i,h}^{k+l+1}) + \gamma \sum_{k=0}^{N-j-1} d_{BL^*}(\rho_{i,h}^{k+j+2},\rho_{i,h}^{k+j+1}) \\
&\qquad \leqslant h \sum_{k=0}^{N-j-1} \sum_{l=0}^{j} d_{BL^*}(\rho_{i,h}^{k+l+2},\rho_{i,h}^{k+l+1}) \\
&\qquad \leqslant h(j+1) \sum_{k=0}^{N-1} d_{BL^*}(\rho_{i,h}^{k+1},\rho_{i,h}^{k}) \\
&\qquad \leqslant r\tau N^{1/2} \left(\sum_{k=0}^{N-1} d_{BL^*}(\rho_{i,h}^{k+1},\rho_{i,h}^{k})^2 \right)^{1/2}
\end{aligned}
$$

Then, using (7.15), we obtain

$$
\sup_h \int_0^{T-\tau} d_{BL^*}(\rho_{i,h}(t+\tau),\rho_{i,h}(t))\,dt \leqslant C\tau.
$$

□

7.1.3 Euler-Lagrange equations and space compactness

Proposition 7.9. *For every $k \geqslant 0$, $\nabla P_i(\rho_{i,h}^{k+1}) \in L^2(\Omega)$ and*

$$
(y - T_{i,h}^{k+1}(y))\rho_{i,h}^{k+1}(y) + h\nabla P_i(\rho_{i,h}^{k+1}(y)) + h\nabla V_i \rho_{i,h}^{k+1} = 0 \qquad a.e\ on\ \Omega, \tag{7.16}
$$

where $T_{i,h}^{k+1}$ is the W_2-optimal transport map between $\rho_{i,h}^{k+1}$ and $\tilde{\rho}_{i,h}^{k+1}$. Then $\rho_{i,h}$ satisfies, for all $\zeta \in \mathcal{C}_c^\infty([0,T)\times\overline{\Omega})$,

$$
\begin{aligned}
\int_0^T \int_\Omega \rho_{i,h}\partial_t\zeta \;&=\; h\sum_{k=0}^{N-1} \int_\Omega (\nabla P_i(\rho_{i,h}^{k+1}(x)) + \rho_{i,h}^{k+1}(x)\nabla V_i(x))\cdot\nabla\zeta(t_k,x)\,dx \\
&-\; h\sum_{k=0}^{N-1} \int_\Omega f_i(\rho_{1,h}^{k+1}(x),\rho_{2,h}^{k+1}(x))\zeta(t_k,x)\,dx \\
&+\; h\sum_{k=0}^{N-1} \int_{\Omega^2} \mathcal{R}[\zeta(t_k,\cdot)](x,y)\,d\gamma_{i,h}^{k+1}(x,y) \\
&-\; \int_\Omega \rho_{i,0}(x)\zeta(0,x)\,dx,
\end{aligned} \tag{7.17}
$$

where, $\gamma_{i,h}^{k+1} = (Id \times T_{i,h}^{k+1})_{\#}\rho_{i,h}^{k+1}$ is the W_2-optimal transport plan between $\rho_{i,h}^{k+1}$ and $\tilde{\rho}_{i,h}^{k+1}$ and $|\mathcal{R}[\zeta](x,y)| \leqslant \frac{1}{2}\|D^2\zeta\|_{L^\infty}|x-y|^2$.

Proof. The proof is the same as in [65, 1, 103] for example. Let ξ a smooth vector field in $\mathcal{C}_c^\infty(\Omega,\mathbb{R}^n)$ and consider Φ_τ the corresponding flow defined by

$$\partial_\tau \Phi_\tau = \xi \circ \Phi_\tau,\ \Phi_0 = Id.$$

We define a pertubation of $\rho_{i,h}^{k+1}$ by $\rho_\tau := \Phi_{\tau\#}\rho_{i,h}^{k+1}$. Then we get

$$\frac{1}{\tau}\left(\frac{1}{2h}(W^2_{2,m_i^{k+1}}(\rho_\tau,\tilde\rho_{i,h}^{k+1}) - W^2_{2,m_i^{k+1}}(\rho_{i,h}^{k+1},\tilde\rho_{i,h}^{k+1})) \right.$$
$$\left. + \mathcal{F}_i(\rho_\tau) - \mathcal{F}_i(\rho_{i,h}^{k+1}) + \mathcal{V}_i(\rho_\tau) - \mathcal{V}_i(\rho_{i,h}^{k+1})\right) \geqslant 0. \tag{7.18}$$

By standard computations, we have

$$\limsup_{\tau\searrow 0} \frac{1}{\tau}\frac{1}{2h}(W^2_{2,m_i^{k+1}}(\rho_\tau,\tilde\rho_{i,h}^{k+1}) - W^2_{2,m_i^{k+1}}(\rho_{i,h}^{k+1},\tilde\rho_{i,h}^{k+1}))$$
$$\leqslant \frac{1}{h}\int_{\Omega\times\Omega}(x-y)\cdot\xi(x)\,d\gamma_{i,h}^{k+1}(x,y), \tag{7.19}$$

with $\gamma_{i,h}^{k+1}$ is the $W_{2,m_i^{k+1}}$-optimal transport plan between $\rho_{i,h}^{k+1}$ and $\tilde\rho_{i,h}^{k+1}$ and $\gamma_{i,h}^{k+1} = (Id\times T_{i,h}^{k+1})_\#\rho_{i,h}^{k+1}$ with $T_{i,h}^{k+1} = Id - \nabla\varphi_{i,h}^{k+1}$. Moreover, using (7.12) and Lebesgue's dominated convergence theorem, we obtains

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{F}_i(\rho_\tau) - \mathcal{F}_i(\rho_{i,h}^{k+1})) \leqslant -\int_\Omega P_i(\rho_{i,h}^{k+1}(x))\,\mathrm{div}(\xi(x))\,dx. \tag{7.20}$$

Finally,

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{V}_i(\rho_\tau) - \mathcal{V}_i(\rho_{i,h}^{k+1})) \leqslant \int_\Omega \nabla V\cdot\xi\rho_{i,h}^{k+1}\,dx. \tag{7.21}$$

If we combine (7.18), (7.19), (7.20) and (7.21), and if we replace ξ by $-\xi$, we find, for all $\xi\in\mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega \nabla\varphi_{i,h}^{k+1}\cdot\xi\rho_{i,h}^{k+1} - h\int_\Omega P_i(\rho_{i,h}^{k+1})\,\mathrm{div}(\xi) + h\int_\Omega \nabla V_i\cdot\xi\rho_{i,h}^{k+1} = 0, \tag{7.22}$$

Now we claim that $P_i(\rho_h^{k+1})\in H^1(\Omega)$. Indeed, since P_i is continuous, (7.12) gives $P_i(\rho_{i,h}^{k+1})\in L^\infty(\Omega)\subset L^2(\Omega)$. Moreover, using (7.22), we obtain

$$\left|\int_\Omega P_i(\rho_{i,h}^{k+1})\,\mathrm{div}(\xi)\right| \leqslant \left[M_T^{1/2}\frac{W_{2,m_i^{k+1}}(\rho_{i,h}^{k+1},\tilde\rho_{i,h}^{k+1})}{h} + \|V_i\|_{L^\infty(\Omega)}M_T|\Omega|^{1/2}\right]\|\xi\|_{L^2(\Omega)},$$

then by duality $\nabla P_i(\rho_{i,h}^{k+1})\in L^2(\Omega)$ and $h\nabla P_i(\rho_{i,h}^{k+1}) + h\nabla V_i\rho_{i,h}^{k+1} + \nabla\varphi_{i,h}^{k+1}\rho_{i,h}^{k+1} = 0$ a.e on Ω.

For fixed $\phi\in\mathcal{C}_c^\infty(\overline\Omega)$, we integrate (7.16) against $\nabla\phi$ to obtain

$$\int_\Omega (y - T_{i,h}^{k+1}(y))\cdot\nabla\phi(y)\rho_{i,h}^{k+1}(y)\,dy + h\int_\Omega \nabla P_i(\rho_{i,h}^{k+1}(y))\cdot\nabla\phi(y)\,dy + h\int_\Omega \nabla V_i(y)\cdot\nabla\phi(y)\rho_{i,h}^{k+1}(y)\,dy = 0.$$

Using Taylor's expansion,

$$\phi(x) - \phi(y) = \nabla\phi(y)\cdot(x-y) + \mathcal{R}[\phi](x,y),$$

where $|\mathcal{R}[\phi](x,y)|\leqslant C\|D^2\phi\|_{L^\infty}|x-y|^2$. Then

$$\int_\Omega (y - T_{i,h}^{k+1}(y))\cdot\nabla\phi(y)\rho_{i,h}^{k+1}(y)\,dy = \int_\Omega(\phi(y) - \phi(T_{i,h}^{k+1}(y)))\rho_{i,h}^{k+1}(y) + \int_{\Omega^2}\mathcal{R}[\phi](x,y)\,d\gamma_{i,h}^{k+1} \tag{7.23}$$

Let $\zeta \in \mathcal{C}_c^\infty([0,T)\times\overline{\Omega})$ and extend ζ by $\zeta(0,\cdot)$ on $[-h,0)$. Then

$$\begin{aligned}
\int_0^T\int_\Omega \partial_t\zeta(t,x)\rho_{i,h}(t,x)\,dtdx &= \sum_{k=0}^{N}\int_{t_{k-1}}^{t_k}\int_\Omega \partial_t\zeta(t,x)\rho_{i,h}^k(x)\,dtdx \\
&= \sum_{k=0}^{N}\int_\Omega (\zeta(t_k,x)-\zeta(t_{k-1},x))\rho_{i,h}^k(x)\,dx \\
&= \sum_{k=0}^{N-1}\int_\Omega \zeta(t_k,x)(\rho_{i,h}^k(x)-\rho_{i,h}^{k+1}(x))\,dx - \int_\Omega \rho_{i,0}\zeta(0,x)\,dx \\
&= \sum_{k=0}^{N-1}\int_\Omega \zeta(t_k,x)(\tilde{\rho}_{i,h}^{k+1}(x)-\rho_{i,h}^{k+1}(x))\,dx - \int_\Omega \rho_{i,0}\zeta(0,x)\,dx \\
&\quad -h\sum_{k=0}^{N-1}\int_\Omega f_i(\rho_{1,h}^k(x),\rho_{2,h}^k(x))\zeta(t_k,x)\,dx,
\end{aligned}$$

and combining with (7.23) and (7.17) concludes the proof.

□

Corollary 7.10. *There exists a constant C which does not depend on h such that*

$$\|P_i(\rho_{i,h})\|_{L^2((0,T),H^1(\Omega))} \leqslant C.$$

Proof. Since P_i is continuous and $\rho_{i,h}$ satisfies (7.12), $P_i(\rho_{i,h}) \in L^2((0,T)\times\Omega)$. Moreover, using (7.16),

$$\begin{aligned}
\int_0^T\int_\Omega |\nabla P_i(\rho_{i,h})|^2 &= h\sum_{k=0}^{N-1}\int_\Omega |\nabla P_i(\rho_{i,h}^{k+1})|^2 \\
&= \frac{1}{h}\sum_{k=0}^{N-1}\int_\Omega |y-T_{i,h}^{k+1}|^2(\rho_{i,h}^{k+1})^2 + h\sum_{k=0}^{N-1}\int_\Omega |\nabla V_i\rho_{i,h}^{k+1}|^2 \\
&\leqslant M_T\frac{1}{h}\sum_{k=0}^{N-1} W_{2,m_i^{k+1}}^2(\rho_{i,h}^{k+1},\tilde{\rho}_{i,h}^{k+1}) + \|V_i\|_{L^\infty(\Omega)}^2 M_T^2|\Omega|T \\
&\leqslant C,
\end{aligned}$$

using (7.14). □

7.1.4 Convergence

In this section, we prove weak and strong convergence for $\rho_{i,h}$.

First, since $\rho_{i,h}$ is bounded in $L^\infty((0,T)\times\Omega))$, up to a subsequence, $\rho_{i,h}$ converges weakly $*$ in $L^\infty((0,T)\times\Omega))$ to ρ_i. Unfortunately, this convergence is too weak to pass to the limit in the nonlinear terms. We improve the convergence using an extension of the Aubin-Lions lemma given by Rossi and Savaré in [107] (theorem 2).

Theorem 7.11 (th. 2 in [107])**.** *On a Banach space X, let be given*

- *a normal coercive integrand $\mathcal{G} : X \to \mathbb{R}^+$, i.e, $\mathcal{G}$ is l.s.c and its sublevels are relatively compact in X,*
- *a pseudo-distance $g : X\times X \to [0,+\infty]$, i.e, g is l.s.c and*

$$[g(\rho,\mu)=0,\ \rho,\mu\in X \text{ with } \mathcal{G}(\rho),\mathcal{G}(\mu)<\infty] \Rightarrow \rho=\mu.$$

Let U be a set of measurable functions $u : (0,T) \to X$ with a fixed $T > 0$. Under the hypotheses that

$$\sup_{u\in U} \int_0^T \mathcal{G}(u(t))\,dt < +\infty \qquad \text{and} \qquad \lim_{h\searrow 0} \sup_{u\in U} \int_0^{T-h} g(u(t+h), u(t))\,dt = 0, \tag{7.24}$$

U contains a subsequence $(u_n)_{n\in\mathbb{N}}$ which converges in measure with respect to $t \in]0,T[$ to a limit $u_\star :]0,T[\to X$.

We apply this theorem in $X = L^1(\Omega)$, $U = \{\rho_{i,h} \mid h \leqslant \frac{1}{K(1+M_T)}\}$, where g is defined by

$$g(\rho,\mu) := d_{BL^*}(\rho,\mu) \qquad \text{if } \rho,\mu \in L^1(\Omega)$$

and $\mathcal{G}_i$ by

$$\mathcal{G}_i(\rho) := \begin{cases} \mathcal{F}_i(\rho) + \|P_i(\rho)\|_{H^1(\Omega)} & \text{if } \rho \in L^1(\Omega), P_i(\rho) \in H^1(\Omega) \text{ and } F_i(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

$\mathcal{G}_i$ is l.s.c and its sublevels are relatively compact in $L^1(\Omega)$ (see lemma 5.13 from chapter 5) and U satisfies (7.24) because of corollary 7.10 and (7.8). Then theorem 7.11 implies that $\rho_{i,h}$ converges in measure with respect to t in $L^1(\Omega)$ to ρ_i. Now since convergence in measure implies a.e convergence up to a subsequence, we may also assume that $\rho_{i,h}(t)$ converges strongly in $L^1(\Omega)$ to $\rho_i(t)$ t-a.e. Then Lesbegue's dominated convergence theorem imply that $\rho_{i,h}$ converges strongly in $L^1((0,T)\times\Omega)$ to ρ_i and $P_i(\rho_{i,h})$ converges strongly in $L^1((0,T)\times\Omega)$ to $P_i(\rho_i)$.

Moreover, according to corollary 7.10, $\nabla P_i(\rho_{i,h})$ is bounded in $L^2((0,T)\times\Omega)$ which implies, with the strong convergence of $P_i(\rho_{i,0})$ to $P_i(\rho_i)$, that $\nabla P_i(\rho_{i,h})$ converges weakly to $\nabla P_i(\rho_i)$ in $L^2((0,T)\times\Omega)$.

Now we can prove theorem 7.3,

Proof of theorem 7.3. The goal is to pass to the limit in (7.17). The left hand side converges to

$$\int_0^T \int_\Omega \rho_i \partial_t \zeta,$$

because $\rho_{i,h}$ converges weakly $*$ to ρ_i in $L^\infty((0,T)\times\Omega)$ and $\partial_t\zeta \in L^1((0,T)\times\Omega)$. For the same reason, the potential term converges to

$$\int_0^T \int_\Omega \nabla V_i \cdot \nabla\zeta \rho_i.$$

Using the definition of $\mathcal{R}[\zeta]$ and (7.14), the remainder term goes to 0 when h goes to 0. Since $\nabla P_i(\rho_{i,h})$ is bounded in $L^2((0,T)\times\Omega)$ and converges weakly to $\nabla P_i(\rho_i)$ in $L^2((0,T)\times\Omega)$, we have

$$h\sum_{k=0}^{N-1} \int_\Omega \nabla P_i(\rho_{i,h}^{k+1}) \cdot \nabla\zeta(t_k,x) \to \int_0^T \int_\Omega \nabla P_i(\rho_i)\cdot\nabla\zeta(t,x).$$

To conclude, since f_i is bounded (using (7.7) and (7.12)) and continuous on $[0,+\infty)\times[0,+\infty)$, by Lebesgue's dominated convergence theorem, we get

$$h\sum_{k=0}^{N-1} \int_\Omega f_i(\rho_{1,h}^k, \rho_{2,h}^k)\zeta(t_k,x) \to \int_0^T \int_\Omega f_i(\rho_1,\rho_2)\zeta.$$

□

7.1.5 More general cases

In this section, we give a non exhaustive list of possible extensions to theorem 7.1.

General diffusion problems: As in [103], we can add a drift and impose a dependence in time for the drift, the diffusion and a dependence in time and space for f_i. In other words, we solve

$$\begin{cases} \partial_t \rho_1 - a_1(t)\Delta P_1(\rho_1) - \operatorname{div}(\rho_1 \nabla V_1(t,x)) = f_1(t,x,\rho_1,\rho_2) \\ \partial_t \rho_2 - a_2(t)\Delta P_2(\rho_2) - \operatorname{div}(\rho_2 \nabla V_2(t,x)) = f_2(t,x,\rho_1,\rho_2). \end{cases} \tag{7.25}$$

Here the diffusion is given by $\tilde{F}_i(t,x) = a_i(t)F_i(x)$ and $\tilde{P}_i(t,x) = a_i(t)(xF_i'(x) - F_i(x))$ is the pressure at time t associated to $\tilde{F}_i$. Since the diffusion and the drift depend on time, Petrelli and Tudorascu consider times averages for these quantity i.e

$$\tilde{F}^k_{i,h} := \frac{1}{h}\int_{hk}^{h(k+1)} a_i(t)\,dt F_i \text{ and } V^k_{i,h} := \frac{1}{h}\int_{hk}^{h(k+1)} V_i(t,\cdot)\,dt.$$

Then the functional $\mathcal{S}_i := \mathcal{F}_i + \mathcal{V}_i$ that we want to minimize at each step becomes

$$\mathcal{S}^k_{i,h}(\rho) := \frac{1}{h}\int_{hk}^{h(k+1)} \int_\Omega (a_i(t)F_i(\rho(x)) + V_i(t,x)\rho(x))\,dxdt,$$

and the scheme (7.11) becomes

$$\begin{cases} \tilde{\rho}^{k+1}_{i,h} = \rho^k_{i,h} + f^k_{i,h}(\rho^k_{1,h}, \rho^k_{2,h}), \\ \rho^{k+1}_{i,h} \in \operatorname{argmin}_{\rho \in \mathcal{M}_{m_i^{k+1}}} \frac{1}{2h} W^2_{2,m_i^{k+1}}(\rho, \tilde{\rho}^{k+1}_{i,h}) + \mathcal{S}^k_{i,h}(\rho), \end{cases} \tag{7.26}$$

where $f^k_{i,h} := \int_{hk}^{h(k+1)} f_i(t,\cdot,\rho^k_{1,h},\rho^k_{2,h})\,dt$.

To obtain estimates that we have with scheme (7.11), we need to assume uniform estimates on V_i, F_i and f_i as in [103]. We assume that for all $T < +\infty$, there exists $C_T > 0$ such that, for all $t, s \leqslant T$,

$$\|V_i(t,\cdot)\|_{L^\infty} + \|\nabla V_i(t,\cdot)\|_{L^\infty} \leqslant C_T \text{ and } \|\nabla V_i(t,\cdot) - \nabla V_i(s,\cdot)\|_{L^\infty} \leqslant C_T|t-s|.$$

The diffusion terms $F_i : [0,+\infty)$ have to satisfy the same assumptions as before and a_i is a Lipschitz continuous function bounded from below by a positive constant i.e $a_i(t) \geqslant \alpha > 0$.

Now, we make the following assumptions on $f_i : \mathbb{R}^+ \times \Omega \times [0,+\infty) \times [0,+\infty) \to \mathbb{R}$. There exists $K > 0$ such that, for all $t, x_1, x_2 \in [0,+\infty)$ and $x \in \Omega$,

$$f_i(t,x,x_1,x_2) \leqslant K(1+x_1+x_2), \tag{7.27}$$

$$f_1(t,x,x_1,x_2) \geqslant -K(1+x_2)x_1, \qquad f_2(t,x,x_1,x_2) \geqslant -K(1+x_1)x_2. \tag{7.28}$$

Moreover, we assume that $f_i(t,x,\cdot,\cdot)$ is continuous on $[0,+\infty)\times[0,+\infty)$ and $(f_i(\cdot,x,x_1,x_2))_{(x,x_1,x_2)}$ is equicontinuous on $[0,+\infty)$ with respect to (x,x_1,x_2).

There were two main steps in our previous proof. The first one is proposition 7.12 to obtain a uniform bound on $\rho_{i,h}$. The proof can be extended to this framework and this is again a subcase of what is done in lemma 2 in [103]. We remark that using this bound, we can obtain the same type of estimate as in (7.14) and then (7.15) so we have time compactness for the sequence $(\rho_{i,h})_h$. The second step is space compactness. Since a_i is assumed to be bounded from below, we obtain the same estimate as (7.10) and then we can apply Rossi, Savaré's theorem 7.11 to retrieve strong convergence.

Nonlocal drift: Another problem we can think of is when the motion of each species depends of the others. In other words, we allow interactions in the drift,

$$\begin{cases} \partial_t \rho_1 - \Delta P_1(\rho_1) - \operatorname{div}(\rho_1 \nabla (V_1[\rho_1,\rho_2])) = f_1(t,x,\rho_1,\rho_2) \\ \partial_t \rho_2 - \Delta P_2(\rho_2) - \operatorname{div}(\rho_2 \nabla (V_2[\rho_1,\rho_2])) = f_2(t,x,\rho_1,\rho_2). \end{cases} \tag{7.29}$$

This kind of system appears naturally in a variety of domains such as biology, population dynamic, etc. Without mass creation, this problem has been study in [35, 73]. The proof of

existence is based on a semi-implicit JKO scheme introduced by Di Francesco and Fagioli in [54] to solve the same problem without diffusion. In our case we have to combine the splitting scheme (7.11) and the semi-implicit JKO scheme. If $h > 0$ is a fixed time step, we construct $(\rho_{i,h}^k)_k$ and $(\tilde{\rho}_{i,h}^k)_k$ by induction using

$$\begin{cases} \tilde{\rho}_{i,h}^{k+1} = \rho_{i,h}^k + h f_i(\rho_{1,h}^k, \rho_{2,h}^k), \\ \rho_{i,h}^{k+1} \in \operatorname{argmin}_{\rho \in \mathcal{M}_{m_i^{k+1}}} \frac{1}{2h} W_{2,m_{k+1}}^2(\rho, \tilde{\rho}_{i,h}^{k+1}) + \mathcal{F}_i(\rho) + \int_\Omega V_i[\tilde{\rho}_{1,h}^k, \tilde{\rho}_{2,h}^k]\rho. \end{cases} \tag{7.30}$$

Assume that $V_i : L^1(\Omega;\mathbb{R}^+)^2 \to \mathcal{C}^1(\Omega)$ is continuous and verifies, uniformly in $\rho_1, \rho_2 \in L^1(\Omega;\mathbb{R}^+)$,

$$\|V_i[\rho_1,\rho_2]\|_{L^\infty(\Omega)} + \|\nabla(V_i[\rho_1,\rho_2])\|_{L^\infty(\Omega)} \leqslant C(\|\rho_1\|_{L^1(\Omega)} + \|\rho_2\|_{L^1(\Omega)}), \tag{7.31}$$

and

$$\|\nabla(V_i[\rho_1,\rho_2]) - \nabla(V_i[\mu_1,\mu_2])\|_{L^\infty(\Omega)} \leqslant C(\|\rho_1-\mu_1\|_{L^1(\Omega)} + \|\rho_2-\mu_2\|_{L^1(\Omega)}). \tag{7.32}$$

For example, drifts we have in mind are of the form

$$V_i[\rho_1,\rho_2] = W_{i,1} * \rho_1 + W_{i,2} * \rho_2,$$

where the kernels $W_{i,1}$ and $W_{i,2}$ are in $W^{1,\infty}(\Omega)$.

Here, drifts are not supposed bounded so we can not apply directly proposition 7.6 but if we assume that f_i satisfy again (7.28) and (7.27) then

$$\|\rho_{i,h}^{k+1}\|_{L^1(\Omega)} = \|\tilde{\rho}_{i,h}^{k+1}\|_{L^1(\Omega)} \leqslant \|\rho_{i,h}^k\|_{L^1(\Omega)} + hK(1 + \|\rho_{1,h}^k\|_{L^1(\Omega)} + \|\rho_{2,h}^k\|_{L^1(\Omega)}).$$

Then if we sum on i, and note $A_k = \|\rho_{1,h}^k\|_{L^1(\Omega)} + \|\rho_{2,h}^k\|_{L^1(\Omega)}$, we have

$$A_{k+1} \leqslant (1+hK)A_k + hK,$$

which imply for all k, h such that $hk \leqslant T$,

$$A_k \leqslant e^{KT}A_0 + e^{KT} - 1.$$

Then $\|V_i[\tilde{\rho}_{1,h}^k, \tilde{\rho}_{2,h}^k]\|_{L^\infty(\Omega)} \leqslant R_T$ and we can apply proposition 7.6. Moreover, estimate (7.14) is rediscovered using the same argument than in [54] (proposition 4.1) or [73] (proposition 3.4), so we have time-compactnes. And the rest of the proof does not change.

Unbounded domains: Under some assumptions on the reaction term, we can extend this result to the whole space. Indeed, if we assume moreover that,

$$\|f_i(x,\rho_1,\rho_2)\|_{L^1} \leqslant C(\|\rho_1\|_{L^\infty} + \|\rho_1\|_{L^1} + \|\rho_2\|_{L^\infty} + \|\rho_2\|_{L^1}), \tag{7.33}$$

then we control the L^1-norm of ρ_i during the step where the mass is updated and then we recover

$$d_{BL^*}(\tilde{\rho}_{i,h}^{k+1}, \rho_{i,h}^k) = \mathcal{O}(h).$$

7.1.6 Uniqueness

In this section, we state an L^1-contraction result under some restriction on the diffusion, the reaction terms and assuming some regularity on solutions. The argument comes from [1, 103]. We assume that F_i satisfies: there exists $\alpha_i > 0$ and $\beta_i \in (1,3]$ such that

$$(P_i(x) - P_i(y)) \cdot (x-y) \geqslant \alpha_i |x-y|^{\beta_i}. \tag{7.34}$$

Moreover, we assume that f_i is Lipschitz i.e

$$|f_i(x_1,x_2) - f_i(y_1,y_2)| \leqslant Lip(f_i)(|x_1-y_1| + |x_2-y_2|).$$

Then, we have

Theorem 7.12. *Assume f_i satisfies (7.35) and F_i satisfies (7.34). Then under the assumption of theorem 7.3, let (ρ_1, ρ_2) and (μ_1, μ_2) two solutions of problem (7.1) with initial conditions $(\rho_{1,0}, \rho_{2,0})$ and $(\mu_{1,0}, \mu_{2,0})$. If $\partial_t\rho_1, \partial_t\rho_2, \partial_t\mu_1, \partial_t\mu_2 \in L^1((0,T)\times\Omega)$, then*

$$\|\rho_1(t) - \mu_1(t)\|_{L^1(\Omega)} + \|\rho_2(t) - \mu_2(t)\|_{L^1(\Omega)} \leqslant e^{\beta t}\left(\|\rho_{1,0} - \mu_{1,0}\|_{L^1(\Omega)} + \|\rho_{2,0} - \mu_{2,0}\|_{L^1(\Omega)}\right),$$

where $\beta := Lip(f_1) + Lip(f_2)$. In particular, if $(\rho_{1,0}, \rho_{2,0}) = (\mu_{1,0}, \mu_{2,0})$, we have uniqueness.

Proof. We note $\Omega_T := (0,T)\times\Omega$. For $\delta > 0$, define

$$\zeta_{i,\delta} := \phi_\delta(P_i(\rho_i) - P_i(\mu_i)),$$

where

$$\phi_\delta(z) := \begin{cases} 0 & \text{if } z \leqslant 0, \\ \frac{z}{\delta} & \text{if } 0 \leqslant z \leqslant \delta, \\ 1 & \text{if } z \geqslant \delta. \end{cases}$$

Using $\zeta_{i,\delta}$, or a smooth approximation of $\zeta_{i,\delta}$ in the equation satisfied by ρ_i and μ_i, we obtain

$$\begin{aligned}&\iint_{\Omega_T} \partial_t(\rho_i - \mu_i)\zeta_{i,\delta} \\ &\quad = -\iint_{\Omega_T} ((\rho_i - \mu_i)\nabla V_i \cdot \nabla\zeta_{i,\delta} + \nabla(P_i(\rho_i) - P_i(\mu_i)) \cdot \nabla\zeta_{i,\delta} - (f_i(\rho_1,\rho_2) - f_i(\mu_1,\mu_2))\zeta_{i,\delta})\, dxdt.\end{aligned}$$

We introduce now $\Omega_T^\delta := \Omega_T \cap \{0 < P_i(\rho_i) - P_i(\mu_i) < \delta\}$ and $R_T^\delta := \Omega_T \cap \{\delta \leqslant P_i(\rho_i) - P_i(\mu_i)\}$. Then by definition of $\zeta_{i,\delta}$,

$$\begin{aligned}\iint_{\Omega_T} \partial_t(\rho_i - \mu_i)\zeta_{i,\delta} &= -\frac{1}{\delta}\iint_{\Omega_T^\delta} \left((\rho_i - \mu_i)\nabla V_i \cdot \nabla(P_i(\rho_i) - P_i(\mu_i)) + |\nabla(P_i(\rho_i) - P_i(\mu_i))|^2\right) dxdt \\ &+ \frac{1}{\delta}\iint_{\Omega_T^\delta} (f_i(\rho_1,\rho_2) - f_i(\mu_1,\mu_2))(P_i(\rho_i) - P_i(\mu_i))\, dxdt \\ &+ \iint_{R_T^\delta} (f_i(\rho_1,\rho_2) - f_i(\mu_1,\mu_2))\, dxdt.\end{aligned}$$

Since solutions of (7.1) are bounded (proposition 7.12), $\rho_1, \rho_2, \mu_1, \mu_2 \leqslant M_T$, and f_i satisfies (7.7),

$$\begin{aligned}\frac{1}{\delta}\iint_{\Omega_T^\delta} (f_i(\rho_1,\rho_2) - f_i(\mu_1,\mu_2))(P_i(\rho_i) - P_i(\mu_i))\, dxdt &\leqslant 2C(1+2M_T)\iint_{\Omega_T^\delta} \frac{P_i(\rho_i) - P_i(\mu_i)}{\delta} \\ &\leqslant 2C(1+2M_T)|\Omega_T^\delta| \to 0,\end{aligned}$$

when $\delta \searrow 0$. Moreover,

$$\iint_{R_T^\delta} (f_i(\rho_1,\rho_2) - f_i(\mu_1,\mu_2))\, dxdt \leqslant Lip(f_i)(\|\rho_1 - \mu_1\|_{L^1(\Omega_T)} + \|\rho_2 - \mu_2\|_{L^1(\Omega_T)}).$$

Now, it remains to pass to the limit in the first term in the right-hand side. We first remark, using (7.34), that

$$\alpha_i|\rho_i - \mu_i|^{\beta_i - 1} \leqslant \frac{(P_i(\rho_i) - P_i(\mu_i))(\rho_i - \mu_i)}{|\rho_i - \mu_i|} \leqslant |P_i(\rho_i) - P_i(\mu_i)| \leqslant \delta$$

in Ω_T^δ. Then, using Young's inequelity,

$$\begin{aligned}\frac{1}{\delta}\iint_{\Omega_T^\delta}(\rho_i-\mu_i)\nabla V_i\cdot\nabla(P_i(\rho_i)-P_i(\mu_i))&\\ \leqslant\frac{1}{2\delta}\iint_{\Omega_T^\delta}\left(\|\nabla V_i\|_{L^\infty}(\rho_i-\mu_i)\right)^2+\frac{1}{2\delta}\|\nabla(P_i(\rho_i)-P_i(\mu_i))\|^2_{L^2(\Omega_T^\delta)}&\\ \leqslant C\|\nabla V_i\|^2_{L^\infty}\delta^{2/(\beta_i-1)-1}|\Omega_T^\delta|+\frac{1}{2\delta}\|\nabla(P_i(\rho_i)-P_i(\mu_i))\|^2_{L^2(\Omega_T^\delta)}.&\end{aligned}$$

This implies

$$\begin{aligned}-\frac{1}{\delta}\iint_{\Omega_T^\delta}\left((\rho_i-\mu_i)\nabla V_i\cdot\ \nabla(P_i(\rho_i)-P_i(\mu_i))+|\nabla(P_i(\rho_i)-P_i(\mu_i))|^2\right)dxdt&\\ \leqslant C\|\nabla V_i\|^2_{L^\infty}\delta^{2/(\beta_i-1)-1}|\Omega_T^\delta|-\frac{1}{2\delta}\|\nabla(P_i(\rho_i)-P_i(\mu_i))\|^2_{L^2(\Omega_T^\delta)}&\\ \leqslant C\|\nabla V_i\|^2_{L^\infty}\delta^{2/(\beta_i-1)-1}|\Omega_T^\delta|\to 0,&\end{aligned}$$

when δ goes to zero. Passing to the limit in δ and noticing that $\{P_i(\rho_i)-P_i(\mu_i)\geqslant 0\}=\{\rho_i-\mu_i\geqslant 0\}$, we obtain

$$\iint_{\Omega_T}\partial_t(\rho_i-\mu_i)_+\leqslant Lip(f_i)(\|\rho_1-\mu_1\|_{L^1(\Omega_T)}+\|\rho_2-\mu_2\|_{L^1(\Omega_T)}).$$

If we exchange the roles of ρ_i and μ_i and we sum inequalities obtained for $i=1$ and $i=2$, we have

$$\iint_{\Omega_T}\partial_t(|\rho_1-\mu_1|+|\rho_2-\mu_2|)\leqslant(Lip(f_1)+Lip(f_2))(\|\rho_1-\mu_1\|_{L^1(\Omega_T)}+\|\rho_2-\mu_2\|_{L^1(\Omega_T)}),$$

which concludes the proof using Gronwall's lemma.

□

Remark 7.13. *If $V_i=0$, we can dispense from assumption* (7.34) *in theorem 7.12.*

7.1.7 Simulations

In [70], Kinderlherer and Walkington present an numerical scheme to compute the solution of general parabolic equation in dimension 1. In this section, we present an alternative method to numerically solve system (7.1) based on the Benamou-Brenier formula, [11] and an augmented Lagrangian method. We solve the JKO part using the augmented lagrangian algorithm introduced in chapter 3 (see [17]) and the mass update step is easy because it is just a pointwise sum.

The first application is for the diffusive prey-predator model. We come back to the model of three species runing after each other studied in chapter 4, figure 4.3 and we add a reaction term corresponding to the prey-predator model (7.9).

Here we assume that ρ_i is the predator for ρ_{i+1} and the prey of ρ_{i-1}, where $\rho_4:=\rho_1$ and $\rho_0:=\rho_3$. Then, the reaction term for ρ_i is

$$f_i(\rho_1,\rho_2,\rho_3)=A\rho_i\left(1-\frac{\rho_i}{K}\right)-\frac{B_{i-1}\rho_i\rho_{i-1}}{1+E\rho_i}+\frac{B_i\rho_i\rho_{i+1}}{1+E\rho_{i+1}}-C\rho_2,$$

with $B_0=B_3$. Figure 7.1 represents the evolution of this model with $A=15, K=1, C=5$ and $B_i=50$ for all $i\in\{1,2,3\}$. Since the interactions are symetric, we observe that the behaviour of the species are similar. Moreover, we can see starting from time 0.5, the quantities of mass decreases since the species do not interact (see fourth row from figure 7.1). In figure 7.2, we impose that one species, ρ_3, is a better hunter than the others ($B_3=100$) and ρ_2 is a bad predator compared to the others ($B_2=40$). In addition, we impose $B_1=60$. Since the coefficient B_3 for ρ_3 is largest one, we can notice that the quantity of mass of ρ_3 increases and consequently the mas of ρ_1 decreases.

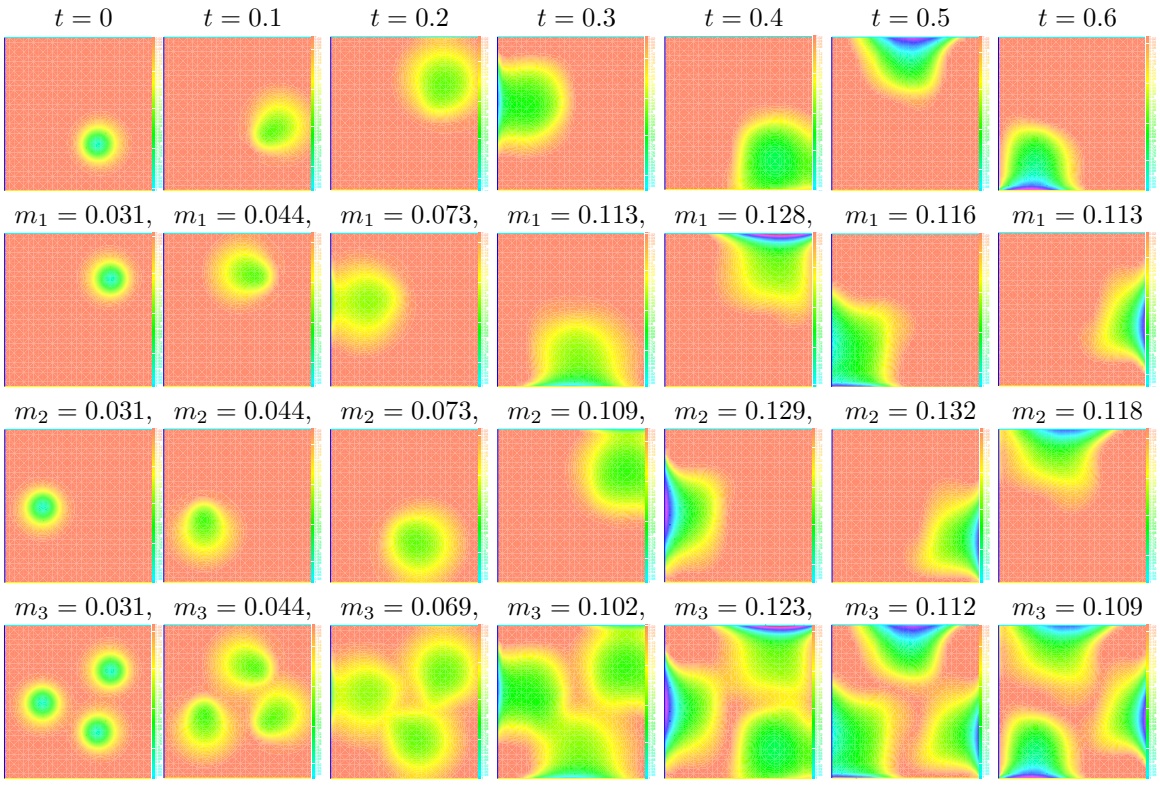


Figure 7.1: *Evolution of three species running after each other with linear diffusion and symetric prey-predator interactions. First row: display of ρ_1 with mass m_1. Second row: display of ρ_2 with mass m_2. Third row: display of ρ_3 with mass m_3. Fourth row: display of $\rho_1 + \rho_2 + \rho_3$*

Another model that we can study numerically is the tumor growth model with elliptic equation for the nutrient, studied in [102]

$$\begin{cases} \partial_t n - \Delta n^m = nG(c), \\ -\Delta c + \alpha nc = 0, \end{cases}$$

where n represents the population density of cells and c the nutrients which are diffused in the tumor. This system is endowed with no flux boundary condition for n and with Dirichlet boundary condition for nutrients, $c = 1$ on the boundary. In this context the elliptic model for the nutrients describes the fact that their diffusion is fast compared to the time scale of cell division.

This system is not covered by the previous theory (from section 7.1) but we can solve easily the elliptic equation using Freefem++ and use the scheme where at each step, we update the mass by

$$\tilde{n}_h^{k+1} = n_h^k + hn_h^k G(c_h^k),$$

where c_h^k is the solution of

$$-\Delta c + \alpha n_h^k c = 0.$$

In figure 7.3, we compute the density of cells and nutrients with

$$m = 100,\ G(c) = 200(c - 0.3),\ \alpha = 4.$$

We remark that we obtain similar results to [102]. When there are not enough nutrients to feed the cells and then cells die in the middle of the tumor.

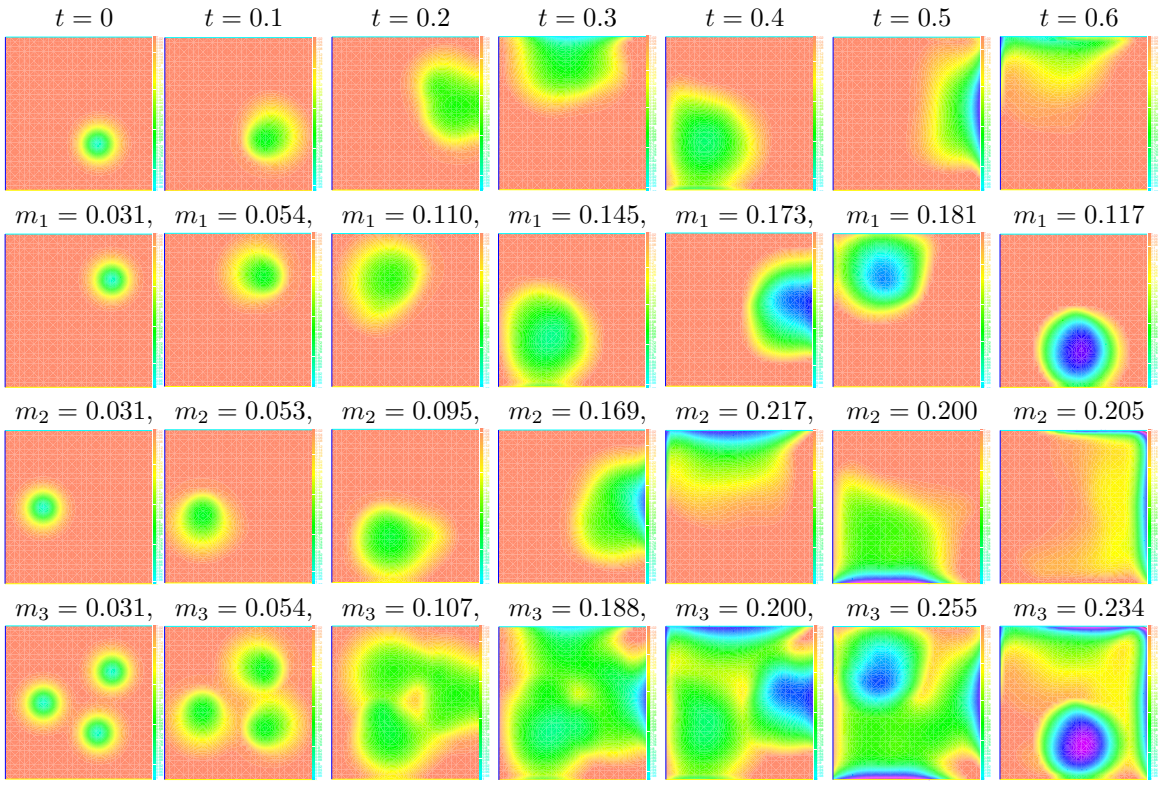


Figure 7.2: *Evolution of three species running after each other with linear diffusion and prey-predator interactions. First row: display of ρ_1 with mass m_1. Second row: display of ρ_2 with mass m_2. Third row: display of ρ_3 with mass m_3. Fourth row: display of $\rho_1 + \rho_2 + \rho_3$*

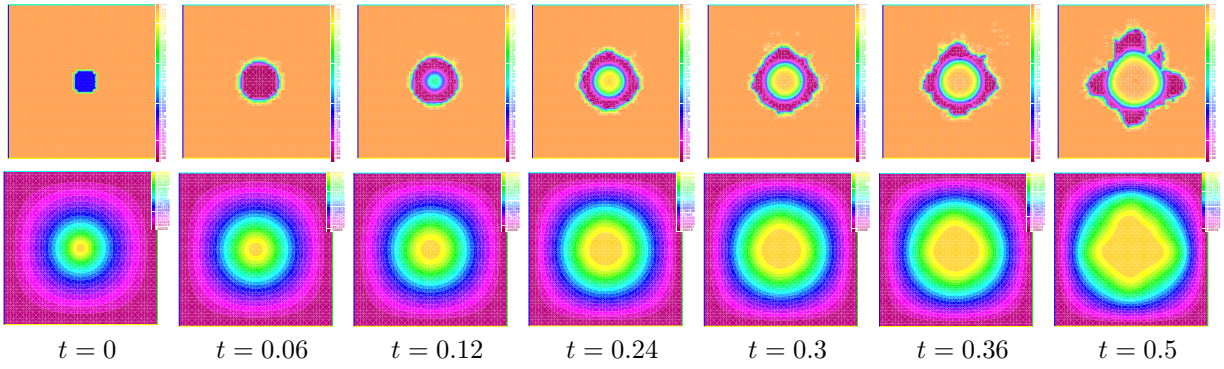


Figure 7.3: *Evolution of tumor growth with elliptic model for nutrient. Top row: display of ρ. Bottom row: display of the density of nutrient n.*

7.2 Wasserstein-Fisher-Rao splitting

In this section, we propose to solve parabolic equation of the form

$$\partial_t \rho - \operatorname{div}(\rho \nabla (F_1'(\rho) + V_1)) = -\rho (F_2'(\rho) + V_2),\ \rho_{t=0} = \rho_0, \tag{7.35}$$

in a bounded domain $\Omega \subset \mathbb{R}^n$ with Neumann boundary condition. This is a work in preparation in collaboration with Gallouët and Monsaingeon.

This section is organized as follow. In section 7.2.1, we recall some results on the Wasserstein-Fischer-Rao metric. Section 7.2.2, we extend the result of [63] to solve (7.35) then we apply our scheme to solve equations coming from tumor growth models in sections 7.2.3 and 7.2.4.

7.2.1 The Wasserstein-Fischer-Rao metric

Recently, a new optimal transport distance was introduced on the space of positive Radon measures. This distance was simultaneously introduced by three different teams, [43, 44, 71, 81, 80], and, here, we call it the Wasserstein-Fischer-Rao metric. We refer to [43, 44, 71, 81, 80] for a more detailed exposition. The advantage of this distance is that it allows mass variations, contrarily to the Wasserstein metric.

Definition 7.14. *The Wasserstein-Fisher-Rao distance between $\rho_0, \rho_1 \in \mathcal{M}^+(\Omega)$ is*

$$WFR^2(\rho_0,\rho_1) := \inf_{(\rho,v,r)\in\mathcal{A}_{WFR}[\rho_0,\rho_1]} \int_0^1 \int_\Omega (|v_t(x)|^2 + |r_t|^2)\, d\rho_t(x)dt,$$

where the admissible set $\mathcal{A}_{WFR}[\rho_0,\rho_1]$ is the set of curves $t\in[0,1]\mapsto(\rho_t,v_t,r_t)\in\mathbb{R}^+\times\mathbb{R}^n\times\mathbb{R}$ such that $t\mapsto\rho_t\in\mathcal{C}_w([0,1],\mathcal{M}^+)$ is weakly continuous, for the narrow convergence of measures, with endpoints $\rho_{|t=0}=\rho_0$, $\rho_{|t=1}=\rho_1$ and the velocity and the reaction $v,r\in L^2((0,1),L^2(\rho_t))$ solve the continuity equation with source

$$\partial_t\rho_t + \operatorname{div}(\rho_t v_t) = \rho_t r_t.$$

Comparing (7.14) with the Benamou-Brenier formula and the Fisher-Rao metric, this dynamic formulation *à la Benamou-Brenier* [11] shows that the WFR distance can be viewed as an inf-convolution of the Monge-Kantorovich and Fisher-Rao distances W_2, FR. By the results of [43, 44, 80] the infimum in the definition is always a minimum, and the corresponding minimizing curves $t\mapsto\rho_t$ are of course called geodesics. This distance is a metric in the space of positive Radon measures and metrizes the narrow convergences of measures. Interestingly, there are other possible formulations of the distance in terms of static unbalanced optimal transportation, primal-dual characterizations with relaxed marginals, lifting to probability measures on a cone over Ω, and duality with subsolutions of Hamilton-Jacobi equations. See also [80, 81] for a related version with mass penalization.

As an immediate consequence of the definition 7.14 we have a first interplay between the distances WFR, W_2, FR (see [43, 44, 71, 81, 80]):

Proposition 7.15. *Let $\rho_0,\rho_1\in\mathcal{M}_2^+$ such that $|\rho_0|=|\rho_1|$. Then*

$$WFR^2(\rho_0,\rho_1)\le W_2^2(\rho_0,\rho_1).$$

Similarly for all $\mu_0,\mu_1\in\mathcal{M}^+$ (with possibly different masses) there holds

$$WFR^2(\mu_0,\mu_1)\le FR^2(\mu_0,\mu_1).$$

Moreover, there exists the following link between the reaction and the velocity (see [43, 44, 71, 81, 80]).

Proposition 7.16 (Proposition 2.2 from [63])**.** *The definition* (7.14) *of the WFR distance can be restricted to the subclass of admissible paths (v_t,r_t) such that $v_t=\nabla r_t$.*

This connection suggests that we have no hope to treat general reaction-diffusion equation using a WFR gradient flow. However, this difficulty can be overcome using the splitting scheme introduced by Gallouët and Monsaingeon in [63] as we will see in the next section.

7.2.2 Existence result for general parabolic equations

In this section, we extend the method introduced in [63] to solve (7.35). We assume that $F_1:\mathbb{R}\to\mathbb{R}$ is given by $F_1(x)=x\log(x)$ (which gives a linear diffusion) or $F_1(x)=\frac{1}{m_1-1}x^{m_1}$, with $m_1>1$ (which gives porous media diffusion) and $F_2:\mathbb{R}\to\mathbb{R}$ is given by $F_2(x)=\frac{1}{m_2-1}x^{m_2}$, with $m_2>1$. In addition, we assume that $V_1\in W^{1,\infty}(\Omega)$ and $V_2\in L^\infty(\Omega)$. We denote $\mathcal{E}_1,\mathcal{E}_2:\mathcal{M}^+\to\mathbb{R}$ the energy given by

$$\mathcal{E}_i(\rho) := \mathcal{F}_i(\rho)+\mathcal{V}_i(\rho),$$

where

$$\mathcal{F}_i(\rho) := \begin{cases} \int_\Omega F_i(\rho) & \text{if } \rho \ll \mathcal{L}_{|\Omega} \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and } \mathcal{V}_i(\rho) := \int_\Omega V_i \rho.$$

A weak solution of (7.35) is a curve $\rho : t \in (0,+\infty) \mapsto \rho(t,\cdot) \in L^1(\Omega) \cap L^\infty(\Omega)$ such that $\nabla P_1(\rho) \in L^2([0,T]\times\Omega)$, where $P_1(x) = xF_1'(x) - F_1(x)$, for all $T < +\infty$, and

$$\int_0^{+\infty} \left(\int_\Omega (\rho \partial_t \phi - \nabla V_1 \cdot \nabla \phi \rho - \nabla P_1(\rho) \cdot \nabla \phi - \rho(F_2'(\rho) + V_2)\phi)\, dx \right) dt = -\int_\Omega \phi(0,x)\rho_0(x)\, dx,$$

for every $\phi \in \mathcal{C}_c^\infty([0,+\infty)\times\mathbb{R}^n)$.

In order to solve (7.35), we introduce the splitting scheme, which comes from [63], and which consists in splitting one step of the usual minimizing movement method between one substep of Wasserstein gradient flow and one substep of Fisher-Rao gradient flow. Let $h > 0$ be a time step, we construct two sequences $(\rho_h^k)_k$ and $(\rho_h^{k+1/2})_k$ by induction such that $\rho_h^0 := \rho_0$ and for $k \geqslant 0$

$$\begin{cases} \rho_h^{k+1/2} \in \operatorname{argmin}_{\rho\in\mathcal{M}^+, |\rho|=|\rho_h^k|} \left\{ \frac{1}{2h} W_2^2(\rho, \rho_h^k) + \mathcal{E}_1(\rho) \right\}, \\ \rho_h^{k+1} \in \operatorname{argmin}_{\rho\in\mathcal{M}^+} \left\{ \frac{1}{2h} FR_2^2(\rho, \rho_h^{k+1/2}) + \mathcal{E}_2(\rho) \right\}. \end{cases} \tag{7.36}$$

This scheme is well-posed (see [63]) and we construct two piecewise-constant interpolating curves

$$\begin{cases} \rho_h(t) = \rho_h^{k+1}, \\ \tilde{\rho}_h(t) = \rho_h^{k+1/2}, \end{cases} \quad \text{for all } t \in ((k-1)h, kh]. \tag{7.37}$$

Theorem 7.17. *Assume $\rho_0 \in L^\infty(\Omega)$ then ρ_h and $\tilde{\rho}_h$ strongly converge in $L^1((0,T)\times\Omega)$ to ρ, where ρ is one weak solution of* (7.35).

We remark that the proof of uniqueness used in theorem 7.12 can be used to prove the uniqueness of (7.35) if we assume that F_1 satisfies (7.34).

The difficulty is that we do not have direct control on the total square distance. Indeed,

$$\frac{1}{2h}\left(W_2^2(\rho_h^k, \rho_h^{k+1/2}) + FR^2(\rho_h^{k+1/2}, \rho_h^k) \right) \leqslant \mathcal{E}_1(\rho_h^k) - \mathcal{E}_1(\rho_h^{k+1/2}) + \mathcal{E}_2(\rho_h^{k+1/2}) - \mathcal{E}_2(\rho_h^{k+1}),$$

therefore when we sum on k we do not recover a telescopic sum in the right hand side. We therefore study the dissipation of $\mathcal{E}_1$ during the Fisher-Rao step driven by $\mathcal{E}_2$ and the dissipation of $\mathcal{E}_2$ during the Wasserstein step driven by $\mathcal{E}_1$.

Optimality conditions and maximum principle

We start recalling that the optimality conditions for (7.36) are (see [63])

$$\frac{-\nabla\varphi_h^{k+1/2}}{h} \rho_h^{k+1/2} = \nabla P_1(\rho_h^{k+1/2}) + \nabla V_1 \rho_h^{k+1/2} \qquad a.e, \tag{7.38}$$

where $\varphi_h^{k+1/2}$ is an optimal Kantorovich potential for $W_2(\rho_h^{k+1/2}, \rho_h^k)$, and

$$\left(\sqrt{\rho_h^{k+1}} - \sqrt{\rho_h^{k+1/2}} \right) \sqrt{\rho_h^{k+1}} = -\frac{h}{2} \rho_h^{k+1} \left(F_2'(\rho_h^{k+1}) + V_2 \right) \qquad a.e. \tag{7.39}$$

Proposition 7.18. *For all $h \leqslant h_0 < \frac{1}{\|V_2\|_\infty}$,*

$$\rho_h^{k+1/2}(x) = 0 \Rightarrow \rho_h^{k+1}(x) = 0, \tag{7.40}$$

and there exists $C = C(h_0, V_2) > 0$ independant of h, such that

$$\rho_h^{k+1}(x) \leqslant (1 + Ch)\rho_h^{k+1/2}(x) \qquad a.e. \tag{7.41}$$

Proof. Assume that $\rho_h^{k+1}(x) > 0$, otherwise the inequality is trivial. Then (7.39) implies

$$\begin{aligned}\sqrt{\rho_h^{k+1}(x)} - \sqrt{\rho_h^{k+1/2}(x)} &= -h\sqrt{\rho_h^{k+1}(x)}(F_2'(\rho_h^{k+1}(x)) + V_2(x)) \\ &\leqslant -hV_2(x)\sqrt{\rho_h^{k+1}(x)},\end{aligned}$$

since $F_2' \geqslant 0$. We deduce

$$\sqrt{\rho_h^{k+1}(x)} \leqslant \frac{1}{1-h\|V_2\|_\infty}\sqrt{\rho_h^{k+1/2}(x)},$$

and to conclude the proof of (7.41), we remark that

$$\frac{1}{(1-h\|V_2\|_\infty)^2} \leqslant 1 + h\left(\frac{2\|V_2\|_\infty + \|V_2\|_\infty^2}{1-h_0\|V_2\|_\infty}\right) = 1 + C(h_0, V_2)h.$$

Then (7.40) follows directly from (7.41). □

Remark 7.19. *We deduce directly from* (7.41) *that we have a mass control*

$$\forall t \in [0,T], \qquad \|\rho_h(t,\cdot)\|_{L^1(\Omega)}, \|\tilde\rho_h(t,\cdot)\|_{L^1(\Omega)} \leqslant e^{CT}\|\rho_0\|_{L^1(\Omega)}.$$

Now we can state a maximum principle.

Proposition 7.20. *Let $T > 0$. There exists a constant M_T such that for all $t \in [0,T]$,*

$$\|\rho_h(t)\|_\infty, \|\tilde\rho_h(t)\|_\infty \leqslant M_T.$$

Proof. The proof is the same as in [115, 103] or proposition 7.6 from section 7.1 because the step of reaction gives an at most linear estimate by proposition 7.18. □

We deduce from this proposition a bound from below for ρ_h^{k+1},

Corollary 7.21. *For h small enough,*

$$\rho_h^{k+1/2}(x) = 0 \Leftrightarrow \rho_h^{k+1}(x) = 0, \tag{7.42}$$

and

$$(1-Ch)\rho_h^{k+1/2}(x) \leqslant \rho_h^{k+1}(x) \leqslant (1+Ch)\rho_h^{k+1/2}(x) \qquad a.e. \tag{7.43}$$

Proof. We have already proved $\rho_h^{k+1/2}(x) = 0 \Leftrightarrow \rho_h^{k+1}(x) = 0$ in proposition 7.18. We refer the reader to proposition 5.1 from [63] for the inverse statement. The proof of (7.43) is the same as proposition 7.18. Indeed, we assume that $\rho_h^{k+1}(x) > 0$ otherwise the inequality is trivial. Then (7.39) gives

$$\begin{aligned}\sqrt{\rho_h^{k+1}(x)} - \sqrt{\rho_h^{k+1/2}(x))} &= -h\sqrt{\rho_h^{k+1}(x)}(F_2'(\rho_h^{k+1}(x)) + V_2(x)) \\ &\geqslant -h(F_2'(M_T) + \|V_2\|_\infty)\sqrt{\rho_h^{k+1}(x)}.\end{aligned}$$

Then the result is proved if h is chosen small enough. □

Energy dissipation

In the Wasserstein minimization problem (first line of (7.36)), taking $\rho = \rho_h^k$, we get

$$\frac{1}{2h}W_2^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant \mathcal{F}_1(\rho_h^k) - \mathcal{F}_1(\rho_h^{k+1/2}) + \int_\Omega V_1(\rho_h^k - \rho_h^{k+1/2}). \tag{7.44}$$

Since $V_1 \in W^{1,\infty}(\Omega)$ using standard methods from [54, 73], we obtain

$$\frac{1}{4h}W_2^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant \mathcal{F}_1(\rho_h^k) - \mathcal{F}_1(\rho_h^{k+1/2}) + Ch. \tag{7.45}$$

Here the difficulty is that we do not recover a telescopic sum on the internal energy. The next proposition gives a control on the dissipation of $\mathcal{F}_1$ during the Fisher-Rao step.

Proposition 7.22. *There exists a constant $C > 0$, such that for all $k \geqslant 0$,*

$$\mathcal{F}_1(\rho_h^{k+1}) \leqslant \mathcal{F}_1(\rho_h^{k+1/2}) + Ch. \tag{7.46}$$

Proof. The first case: $F_1(x) = \frac{1}{m_1-1}x^{m_1}$, with $m_1 > 1$. Since F_1 is increasing and using (7.41), we obtain

$$\begin{aligned}\mathcal{F}_1(\rho_h^{k+1}) - \mathcal{F}_1(\rho_h^{k+1/2}) &\leqslant \frac{1}{m_1-1}\int_\Omega ((1+Ch)^{m_1}-1)(\rho_h^{k+1/2})^{m_1} \\ &\leqslant Ch|\Omega|M_T^{m_1}.\end{aligned}$$

Now if $F_1(x) = x\log(x)$, we have

$$\mathcal{F}_1(\rho_h^{k+1}) = \int_{\{\rho_h^{k+1}\leqslant e^{-1}\}} \rho_h^{k+1}\log(\rho_h^{n+1}) + \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1}\log(\rho_h^{k+1}).$$

Since F_1 is increasing on $\{x \geqslant e^{-1}\}$ and using (7.41), the second term in the right hand side becomes

$$\begin{aligned}\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1}\log(\rho_h^{k+1}) &\leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} (1+Ch)\rho_h^{k+1/2}\log((1+Ch)\rho_h^{k+1/2}) \\ &\leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) \\ &+ Ch\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) \\ &+ (1+Ch)\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1/2}\log(1+Ch) \\ &\leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}} \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) + Ch,\end{aligned}$$

because $\rho_h^{k+1/2} \leqslant M_T$ and $\log(1+Ch) \leqslant Ch$. Using the same method with the bound from below (7.43), we obtain

$$\int_{\{\rho_h^{k+1}\leqslant e^{-1}\}} \rho_h^{k+1}\log(\rho_h^{k+1}) \leqslant \int_{\{\rho_h^{k+1}\leqslant e^{-1}\}} \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) + Ch.$$

Then combining these inequalities, we get

$$\int_\Omega \rho_h^{k+1}\log(\rho_h^{k+1}) \leqslant \int_\Omega \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) + Ch.$$

□

Remark 7.23. *If $F_1(x) = \frac{1}{m_1-1}x^{m_1}$, then the dissipation* (7.46) *is still true on $\mathbb{R}^n$. Indeed, since $\rho_h^{k+1/2}$ is bounded in $L^\infty(\Omega)\cap L^1(\Omega)$, the L^{m_1}-norm is bounded. Then theorem 7.17 is still true on $\mathbb{R}^n$. However, if $F_1(x) = x\log(x)$, we need to have a control on the second moment to bound the negative part.*

Now it is easy to deduce from (7.45) and proposition 7.46 that

$$\sum_{k=0}^{N-1} \frac{1}{2h} W_2^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant \mathcal{F}_1(\rho_0) - \mathcal{F}_1(\rho_h^N) + CT, \tag{7.47}$$

where $N = \lfloor \frac{T}{h} \rfloor$.

Now, taking $\rho = \rho_h^{k+1/2}$ in the second problem in (7.36), we obtain

$$\frac{1}{2h} FR_2^2(\rho_h^{k+1}, \rho_h^{k+1/2}) \leqslant \mathcal{F}_2(\rho_h^{k+1/2}) - \mathcal{F}_2(\rho_h^{k+1}) + \int_\Omega V_2(\rho_h^{k+1/2} - \rho_h^{k+1}). \tag{7.48}$$

Using (7.43) and $V_2 \in L^\infty(\Omega)$, and arguing as in proposition 7.22

$$\int_\Omega V_2(\rho_h^{k+1/2} - \rho_h^{k+1}) \leqslant \|V_2\|_\infty Ch, \tag{7.49}$$

$$\mathcal{F}_2(\rho_h^{k+1/2}) - \mathcal{F}_2(\rho_h^{k+1}) \leqslant Ch. \tag{7.50}$$

The proof of the second estimate is the same than the one of the first part of proposition 7.46 and is still true on $\mathbb{R}^n$. Combining (7.48), (7.49) and (7.50), we immediately deduce that

$$\sum_{k=0}^{N-1} \frac{1}{2h} FR_2^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant CT, \tag{7.51}$$

where $N = \lfloor \frac{T}{h} \rfloor$.

Proposition 7.24. *There exists a constant $C > 0$ such that for all h small enough and $k \geqslant 0$ such that $kh \leqslant T$,*

$$\mathcal{F}_1(\rho_h^k), \mathcal{F}_1(\rho_h^{k+1/2}) \leqslant \mathcal{F}_1(\rho_0) + CT, \tag{7.52}$$

$$\mathcal{F}_2(\rho_h^k), \mathcal{F}_2(\rho_h^{k+1/2}) \leqslant \mathcal{F}_2(\rho_0) + CT, \tag{7.53}$$

$$\sum_{k=0}^{N-1} WFR^2(\rho_h^k, \rho_h^{k+1}) \leqslant h(\mathcal{F}_1(\rho_0) + C(1+T)). \tag{7.54}$$

Proof. Proofs of (7.52) and (7.53) are trivial and to prove (7.54) we just remark that (see [63] or proposition 7.15)

$$WFR^2(\rho_h^k, \rho_h^{k+1}) \leqslant 2(FR_2^2(\rho_h^{k+1/2}, \rho_h^k) + W_2^2(\rho_h^{k+1/2}, \rho_h^k)),$$

and use (7.47) and (7.51). □

Estimates and convergences

As in [63], we recover the classical $\frac{1}{2}$-Hölder estimate, using (7.54), for all $T > t, s \geqslant 0$,

$$WFR(\rho_h(t), \rho_h(s)) \leqslant C|t - s + h|^{1/2} \text{ and } WFR(\tilde{\rho}_h(t), \tilde{\rho}_h(s)) \leqslant C|t - s + h|^{1/2}, \tag{7.55}$$

and

$$WFR(\rho_h(t), \tilde{\rho}_h(t)) \leqslant FR(\rho_h(t), \tilde{\rho}_h(t)) \leqslant C\sqrt{h}. \tag{7.56}$$

Then proposition 4.1 from [63] gives that for all $T > 0$, ρ_h and $\tilde{\rho}_h$ converge (up to a subsequence) to a WFR-continous curve $\rho \in \mathcal{C}^{1/2}([0,T], \mathcal{M}^+_{WFR})$ and

$$\sup_{t \in [0,T]} (WFR(\rho_h(t), \rho(t)) + WFR(\tilde{\rho}_h(t), \rho(t))) \to 0.$$

Proposition 7.25. *For all $T > 0$, ρ_h and $\tilde{\rho}_h$ satisfies*

$$\|P_1(\tilde{\rho}_h)\|_{L^2([0,T];H^1(\Omega))} \leqslant C(1+T), \tag{7.57}$$

Proof. The proof is a consequence of (7.38), (7.47) and proposition 7.20. □

Then using an extention of Aubin-Lions lemma due to Rossi and Savaré in [107] with (7.54) and (7.57), we obtain that $\tilde{\rho}_h$ strongly converges to ρ in $L^1([0,T]\times\Omega)$ (see [73]). Then

$$\|\rho_h-\rho\|_{L^1((0,T)\times\Omega)} \leqslant \|\rho_h-\tilde{\rho}_h\|_{L^1((0,T)\times\Omega)}+\|\tilde{\rho}_h-\rho\|_{L^1((0,T)\times\Omega)}.$$

The second term in the right hand side converges to 0 when h tends to 0 by strong convergence in L^1 of $\tilde{\rho}_h$ and using (7.43), we have

$$\|\rho_h-\tilde{\rho}_h\|_{L^1((0,T)\times\Omega)}\leqslant Ch,$$

which implies that ρ_h strongly converges to ρ in $L^1([0,T]\times\Omega)$. Moreover, since $P_1(\tilde{\rho}_h)$ is bounded in $L^2((0,T),H^1(\Omega))$, $P_1(\rho)\in L^2((0,T),H^1(\Omega))$.

Then the proof of theorem 7.17 is complete arguing as in [63].

Remark 7.26. *Theorem 7.17 can be extend to all $F_2\in\mathcal{C}^1([0,+\infty))$ such that $F_2'\geqslant C$, with $C\in\mathbb{R}$.*

7.2.3 Application to Hele-Shaw equation

In this section we are interested in the equation

$$\begin{cases}\partial_t\rho-\operatorname{div}(\rho\nabla p)=\rho(1-p),\\ p\geqslant 0, p(1-\rho)=0, \rho\leqslant 1,\\ \rho_{|t=0}=\rho_0.\end{cases}\tag{7.58}$$

on Ω a open convex subset of $\mathbb{R}^n$. This equation is motivated by tumor growth models, see for example [101, 102]. At least formally, we remark that (7.58) is the Wasserstein-Fisher-Rao gradient flow for the singular functional

$$\rho\mapsto\mathcal{F}_\infty(\rho)-\int_\Omega\rho,$$

with,

$$\mathcal{F}_\infty(\rho):=\begin{cases}0 & \text{if } \rho\leqslant 1\, a.e,\\ +\infty & \text{otherwise.}\end{cases}$$

However, the splitting approach introduced in the previous section does not work in this cas. Indeed, the scheme would be

$$\begin{cases}\rho_h^{k+1/2}\in\operatorname{argmin}_{\rho\leqslant 1,\,|\rho|=|\rho_h^k|}\frac{1}{2h}W_2^2(\rho,\rho_{h,m}^k)-\int_\Omega\rho,\\ \\ \rho_h^{k+1}\in\operatorname{argmin}_{\rho\leqslant 1}\left\{\frac{1}{2h}FR_2^2(\rho,\rho_h^{k+1/2})-\int_\Omega\rho\right\}.\end{cases}$$

The Wasserstein step conserves the mass so we can drop the mass penalization. As ρ_h^k is smaller or equal than 1, then the Wasserstein projection step (we minimise the Wasserstein distance on the convex set of densities less than 1) returns always ρ_h^k and so we do not have any diffusion effect contrarly to (7.58). Denote by $\mathcal{F}_m$, the functional defined by

$$\mathcal{F}_m(\rho):=\begin{cases}\int_\Omega\frac{\rho^m}{m-1} & \text{if } \rho^m\in L^1(\Omega),\\ +\infty & \text{otherwise.}\end{cases}$$

Since $\mathcal{F}_m$ Γ-converges to $\mathcal{F}_\infty$, see [20], we approximate $\mathcal{F}_\infty$ by $\mathcal{F}_m$.

Given a time step $h>0$, we construct by induction two sequences, $(\rho_{h,m}^{k+1/2})_n$ and $(\rho_{h,m}^k)_k$, with $\rho_{h,m}^0=\rho_0$ and for all $k\geqslant 0$,

$$\begin{cases}\rho_h^{k+1/2}\in\operatorname{argmin}_{|\rho|=|\rho_h^k|}\left\{\frac{1}{2h}W_2^2(\rho,\rho_{h,m}^k)+\mathcal{F}_m(\rho)\right\},\\ \\ \rho_h^{k+1}\in\operatorname{argmin}_{\rho\in\mathcal{M}^+}\left\{\frac{1}{2h}FR_2^2(\rho,\rho_h^{k+1/2})+\mathcal{F}_m(\rho)-\int_\Omega\rho\right\}.\end{cases}\tag{7.59}$$

We define $p_{h,m}^{k+1/2}$ and $p_{h,m}^{k+1}$ by

$$p_{h,m}^{k+1/2} := \frac{m}{m-1}(\rho_{h,m}^{k+1/2})^{m-1} \text{ and } p_{h,m}^{k+1} := \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}.$$

Then we denote $\rho_{h,m}, \tilde{\rho}_{h,m}, p_{h,m}$ and $\tilde{p}_{h,m}$ the piesewise constant interpolations of $\rho_{h,m}^{k+1}, \rho_{h,m}^{k+1/2}, p_{h,m}^{k+1/2}$ and $p_{h,m}^{k+1}$ respectively.

Theorem 7.27. *Assume $\rho_0 \in BV(\Omega)$ and $\rho_0 \leqslant 1$, then $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ strongly converge to ρ_∞ in $L^1((0,T)\times\Omega)$ when $h \searrow 0$ and $m \nearrow +\infty$. Moreover, if $mh \to 0$, then $p_{h,m}$ and $\tilde{p}_{h,m}$ converge weakly to p_∞ respectively in $L^2((0,T), H^1(\Omega))$ and $L^2((0,T)\times\Omega)$ where (ρ_∞, p_∞) is a solution of (7.58).*

Estimates and convergences

In this section, we improve some previous estimates from section 7.2.2. The next lemma gives an explicit L^∞-norm.

Lemma 7.28. *Assume that $\rho_0 \leqslant 1$, then for all $t \in \mathbb{R}^+$,*

$$\|\rho_{h,m}(t,\cdot)\|_\infty, \|\tilde{\rho}_{h,m}(t,\cdot)\|_\infty \leqslant 1.$$

Proof. Assume that $\|\rho_{h,m}^k\|_\infty \leqslant 1$, the maximum principle proved by Otto in [97] implies that $\|\rho_{h,m}^{k+1/2}\|_\infty \leqslant 1$. Now, by contradiction, we assume that $E := \{\rho_{h,m}^{k+1} > 1\}$ has positive Lebesgue measure. The optimality condition in the Fisher-Rao minimisation, (7.39), gives

$$\sqrt{\rho_{h,m}^{k+1}} - \sqrt{\rho_{h,m}^{k+1/2}} = \frac{h}{2}\sqrt{\rho_{h,m}^{k+1}}\left(1 - \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}\right).$$

On E, $1 - \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1} < 0$ then $\rho_{h,m}^{k+1} < \rho_{h,m}^{k+1/2} \leqslant 1$ which is absurd. Since $\|\rho_0\|_\infty \leqslant 1$, the lemma is proved by induction.

□

Moreover, proposition 7.18 gives, with $V_2 = -1$, that $\rho_{h,m}^{k+1} \leqslant (1+h)\rho_{h,m}^{k+1/2}$. We deduce that for all t,

$$\|\rho_{h,m}(t,\cdot)\|_{L^1(\Omega)}, \|\tilde{\rho}_{h,m}(t,\cdot)\|_{L^1(\Omega)} \leqslant e^T \|\rho_0\|_{L^1(\Omega)}.$$

Indeed, the Wasserstein part conserves the mass and the Fischer-Rao part is controlled by

$$\int_\Omega \rho_{h,m}^{k+1} \leqslant (1+h)\int_\Omega \rho_{h,m}^{k+1/2} = (1+h)\int_\Omega \rho_{h,m}^k.$$

By induction,

$$\int_\Omega \rho_{h,m}^{k+1} \leqslant (1+h)^k \int_\Omega \rho_0,$$

which gives the result.

Then testing (7.59) with $\rho = \rho_{h,m}^k$ and $\rho = \rho_{h,m}^{k+1/2}$, we obtain, as in [63],

$$\frac{1}{2h}\left(W_2^2(\rho_{h,m}^k, \rho_{h,m}^{k+1/2}) + FR^2(\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k+1})\right) \leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + \int_\Omega (\rho_{h,m}^{k+1/2} - \rho_{h,m}^{k+1}).$$

which implies, using $\rho_{h,m}^{k+1} \leqslant (1+h)\rho_{h,m}^{k+1/2}$ and the L^1-bound on $\rho_{h,m}^{k+1/2}$,

$$\frac{1}{2h}\left(W_2^2(\rho_{h,m}^k, \rho_{h,m}^{k+1/2}) + FR^2(\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k+1})\right) \leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + Ch.$$

Summing over k from 0 to $N := \lfloor T/h \rfloor$ and using proposition 7.15, we get

$$\sum_{k=0}^{N-1} WFR^2(\rho_{h,m}^k, \rho_{h,m}^{k+1}) \leqslant h(\mathcal{F}_m(\rho_0) - \mathcal{F}_m(\rho_{h,m}^N) + CT),$$

with C a constant independant of m and h. In addition, $\mathcal{F}_m$ is nonegative and

$$\mathcal{F}_m(\rho_0) \leqslant \frac{1}{m-1}\int_\Omega \rho_0 \leqslant C,$$

with C independant of m because $\rho_0 \leqslant 1$ and in $L^1(\Omega)$.

Consequently, we obtain the classical $\frac{1}{2}$-Hölder estimate, for all $t, s \leqslant T$,

$$\left\{ \begin{array}{l} WFR(\rho_{h,m}(t,\cdot), \rho_{h,m}(s,\cdot)) \leqslant C|t-s+h|^{1/2}, \\ WFR(\tilde{\rho}_{h,m}(t,\cdot), \tilde{\rho}_{h,m}(s,\cdot)) \leqslant C|t-s+h|^{1/2}. \end{array} \right. \tag{7.60}$$

Lemma 7.29. *If $\rho_0 \in BV(\Omega)$,*

$$\sup_{t\in[0,T]} \left\{ \|\rho_{h,m}(t,\cdot)\|_{BV(\Omega)}, \|\tilde{\rho}_{h,m}(t,\cdot)\|_{BV(\Omega)} \right\} \leqslant e^T \|\rho_0\|_{BV(\Omega)}.$$

Proof. The argument follows the lines of proposition 5.1 from [63]. First we remark that using theorem 1.1 from [52], the BV-norm is nonincreasing during the Wasserstein step,

$$\|\rho_{h,m}^{k+1/2}\|_{BV(\Omega)} \leqslant \|\rho_{h,m}^{k}\|_{BV(\Omega)}.$$

Now we prove that $\rho_{h,m}^{k+1} = R(\rho_{h,m}^{k+1/2})$ with R a $(1+h)$-Lispchitz function, using the implicit function theorem. This will prove that

$$\|\rho_{h,m}^{k+1}\|_{BV(\Omega)} \leqslant (1+h)\|\rho_{h,m}^{k+1/2}\|_{BV(\Omega)},$$

see [3], and conclude the proof. We already know from (7.43) that $\rho_{h,m}^{k+1/2}$ and $\rho_{h,m}^{k+1}$ have the same support. In this support and from (7.39) $\rho_{h,m}^{k+1}(x)$ is solution of

$$f(\rho, \rho_{h,m}^{k+1/2}(x)) = 0 \text{ with } f(\rho,\mu) = \sqrt{\rho}\left(1 - \frac{h}{2}\left(1 - \frac{m}{m-1}\rho^{m-1}\right)\right) - \sqrt{\mu}.$$

For $\mu > 0$, the implicit function theorem gives the existence of a $\mathcal{C}^1$ map, R, such that the solution, ρ, of $f(\rho,\mu) = 0$ is given by $\rho = R(\mu)$. Moreover,

$$\begin{array}{rcl} \dfrac{dR(\mu)}{d\mu} = -\dfrac{\partial_\mu f}{\partial_\rho f}_{|\rho=R(\mu)} & = & \dfrac{\frac{1}{2\sqrt{\mu}}}{\frac{1}{2\sqrt{\rho}}\left(1 - \frac{h}{2}\left(1 - \frac{m}{m-1}\rho^{m-1}\right)\right) + \frac{h}{2}m\sqrt{\rho}\rho^{m-2}} \\ & \leqslant & \dfrac{\rho}{\mu} \leqslant 1+h, \end{array}$$

Extending R by continuity $R(0) = 0$, we proved as required that $\rho_{h,m}^{k+1} = R(\rho_{h,m}^{k+1/2})$ with R a Lipschitz function with Lipschitz constant $1+h$.

□

Then Rossi-Savaré theorem [107], using the $\frac{1}{2}$-Hölder estimates (7.60), gives that $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ converges strongly in $L^1((0,T)\times\Omega)$ to ρ_∞. Moreover, since $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ are bounded in $L^1 \cap L^\infty((0,T)\times\Omega)$ and by definition of $p_{h,m}$ and $\tilde{p}_{h,m}$, we obtain that $p_{h,m}$ and $\tilde{p}_{h,m}$ converge weakly in $L^q((0,T)\times\Omega)$, for every $q \in [1,+\infty]$, to p_∞ and $\tilde{p}_\infty$. Now we have to show that $p_\infty = \tilde{p}_\infty$. We start by recalling that $\rho_{h,m}^{k+1} \leqslant (1+h)\rho_{h,m}^{k+1/2}$ and $\rho_{h,m}^{k+1}, \tilde{\rho}_{h,m}^{k+1} \in L^1(\Omega;[0,1])$. Then, for all t, and $k = \lfloor t/h \rfloor$,

$$\begin{array}{rcl} \displaystyle\int_\Omega |p_{m,h}(t,\cdot) - \tilde{p}_{m,h}(t,\cdot)| & = & \displaystyle\int_\Omega \frac{m}{m-1}|\rho_{h,m}^{m-1}(t,\cdot) - \tilde{\rho}_{h,m}^{m-1}(t,\cdot)| \\ & = & \displaystyle\int_\Omega \frac{m}{m-1}|(\rho_{h,m}^{k+1})^{m-1} - (\rho_{h,m}^{k+1/2})^{m-1}| \\ & \leqslant & \displaystyle m\int_\Omega |\rho_{h,m}^{k+1} - \rho_h^{k+1/2}| \\ & \leqslant & \displaystyle m\left(\int_\Omega \rho_{h,m}^{k+1/2}\right) h = Chm \longrightarrow 0, \end{array}$$

since $hm \to 0$ by assumption. In the next lemma, we improve the convergence of $\tilde{p}_{h,m}$ in order to pass to the limit in the diffusion term.

Lemma 7.30. *There exists a constant C, independantof h and m, such that*

$$\|\tilde{p}_{h,m}\|_{L^2((0,T),H^1(\Omega))} \leqslant C.$$

Consequently, up to a subsequence, $\tilde{p}_{h,m}$ converges weakly in $L^2((0,T),H^1(\Omega))$ to p_∞.

Proof. The proof is based on the flow interchange argument developped by Matthes, McCann and Savaré in [86] and recalled in chapter 4. Let η be the solution of

$$\begin{cases} \partial_t \eta = \Delta \eta^{m-1} + \varepsilon \Delta \eta, \\ \eta(0) = \rho_h^{n+1/2}. \end{cases}$$

It is well known, [4], that η is the Wasserstein gradient flow of

$$\mathcal{G}(\rho) := \int_\Omega \frac{\rho^{m-1}}{m-2} + \varepsilon \int_\Omega \rho \log(\rho).$$

Since $\mathcal{G}$ is geodesically convex, η satisfies the evolution variational inequality (EVI)

$$\frac{1}{2}\frac{d^+}{dt}_{|t=s} W_2^2(\eta(s),\rho) \leqslant \mathcal{G}(\rho) - \mathcal{G}(\eta(s)),$$

for all $s > 0$ and for all $\rho \in \mathcal{P}^{\mathrm{ac}}(\Omega)$, where

$$\frac{d^+}{dt} f(t) := \limsup_{s \to 0^+} \frac{f(t+s) - f(t)}{s}.$$

Using the minimization scheme (7.59), we obtain that

$$\frac{1}{2}\frac{d^+}{dt}_{|t=s} W_2^2(\eta(s),\rho_{h,m}^k) \geqslant -h \frac{d^+}{dt}_{|t=s} \mathcal{F}_m(\eta(s)).$$

Since η is smooth, the following calculations are allowed for all $s > 0$,

$$\begin{aligned} \partial_s \mathcal{F}_m(\eta(s)) &= \int_\Omega \frac{m}{m-1} \eta(s)^{m-1} (\Delta \eta(s)^{m-1} + \varepsilon \Delta \eta(s)) \\ &= -\int_\Omega \frac{m}{m-1} |\nabla \eta(s)^{m-1}|^2 - \varepsilon \int_\Omega m \eta(s)^{m-2} |\nabla \eta(s)|^2 \\ &\leqslant -\int_\Omega \frac{m}{m-1} |\nabla \eta(s)^{m-1}|^2. \end{aligned}$$

By an easy lower semi-continuity argument, we obtain that

$$\int_\Omega \frac{m-1}{m} |\nabla p_{h,m}^{k+1/2}|^2 = \int_\Omega \frac{m}{m-1} |\nabla (\rho_{h,m}^{k+1/2})^{m-1}|^2 \leqslant \liminf_{s \searrow 0} \frac{d^+}{dt}_{|t=s} \mathcal{F}_m(\eta(s)).$$

Then we have

$$\begin{aligned} h \int_\Omega \frac{m-1}{m} |\nabla p_{h,m}^{k+1/2}|^2 \leqslant{}& \mathcal{F}_{m-1}(\rho_{h,m}^k) - \mathcal{F}_{m-1}(\rho_{h,m}^{k+1/2}) \\ &+ \varepsilon \left(\int_\Omega \rho_{h,m}^k \log(\rho_{h,m}^k) - \int_\Omega \rho_{h,m}^{k+1/2} \log(\rho_{h,m}^{k+1/2}) \right). \end{aligned}$$

Arguing as in proposition 7.46, we have

$$\mathcal{F}_{m-1}(\rho_{h,m}^{k+1}) \leqslant \mathcal{F}_{m-1}(\rho_{h,m}^{k+1/2}) + Ch,$$

with C independant of h and m. Moreover, one step of JKO scheme controls the second moment, the entropy is bounded and then we can pass to the limit in $\varepsilon \searrow 0$:

$$h \int_\Omega \frac{m-1}{m} |\nabla p_{h,m}^{k+1/2}|^2 \leqslant \mathcal{F}_{m-1}(\rho_{h,m}^k) - \mathcal{F}_{m-1}(\rho_{h,m}^{k+1}) + Ch.$$

Summing over k gives

$$\int_0^T \int_\Omega |\nabla \tilde{p}_{h,m}(t,x)|^2 \, dxdt \leqslant 2\mathcal{F}_{m-1}(\rho_0) + CT,$$

for all $T < +\infty$. And since $\mathcal{F}_{m-1}(\rho_0) \leqslant \|\rho_0\|_{L^1(\Omega)}$, the result is proved.

□

Properties of p_∞ and conclusion

In the next lemma, we show that p_∞ satisfies conditions in (7.58).

Lemma 7.31. *The weak limit of $\tilde{p}_{h,m}$, p_∞, satisfies*

$$p_\infty \geqslant 0, p_\infty \leqslant 1 \text{ and } p_\infty(1-\rho_\infty) = 0 \text{ a.e.},$$

where ρ_∞ is the limit of $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$.

Proof. It is obvious that $p_\infty \geqslant 0$. Lemma 7.28 with $p_{h,m} = \frac{m}{m-1}\rho_{h,m}^{m-1} \leqslant \frac{m}{m-1}$ gives that $\rho_\infty \leqslant 1$ and $p_\infty \leqslant 1$ a.e. Now we notice that for $(t,x) \in (0,T)\times\Omega$, if $\rho_\infty(t,x) < 1$ then $\rho_{h,m}^{m-1}(t,x) \to 0$, otherwise $\rho_{h,m}(t,x) \to 1$. In both cases,

$$p_{h,m}(t,x)(1-\rho_{h,m}(t,x)) = \frac{m}{m-1}\rho_{h,m}^{m-1}(t,x)(1-\rho_{h,m}(t,x)) \to 0.$$

Since $p_{h,m}(1-\rho_{h,m})$ is bounded in $L^\infty((0,T)\times\Omega)$ and converges a.e to 0, using Lebesgue's dominated convergence Theorem, we obtain for all $\varphi \in \mathcal{C}_c^\infty((0,T)\times\Omega)$,

$$\int_0^T \int_\Omega p_{h,m}(1-\rho_{h,m})\varphi \to 0.$$

In addition, since $\rho_{h,m}$ strongly converges to ρ_∞ in $L^1((0,T)\times\Omega)$ and $p_{h,m}$ is bounded in $L^q((0,T)\times\Omega)$, for all $q \in [1,+\infty]$, and then converges weakly to p_∞ in $L^q((0,T)\times\Omega)$ we deduce that

$$\int_\Omega p_{h,m}(1-\rho_{h,m})\varphi \to \int_\Omega p_\infty(1-\rho_\infty)\varphi,$$

which proves the lemma. □

The end of the proof of theorem 7.27 is as in [63]. We write the Euler-Lagrange equations of (7.59) and since $\rho_{h,m}, \tilde{\rho}_{h,m}$ converge strongly in $L^q((0,T)\times\Omega)$, for all $q \in [1,+\infty)$, and $p_{h,m}, \nabla\tilde{p}_{h,m}$ converge weakly in $L^2((0,T)\times\Omega)$ the result is proved.

Numerical simulation

Since the scheme (7.59) is constructive, we can do numerical simulations. We use the algorithm presented in 3 (or [17]) to solve the Wasserstein step and the Fisher-Rao step is a convex pointwise minimization problem: for all $x \in \Omega$,

$$\rho_{h,m}^{k+1}(x) = \operatorname*{argmin}_{\rho\in\mathbb{R}^+} 4\left|\sqrt{\rho} - \sqrt{\rho_{h,m}^{k+1/2}(x)}\right|^2 + 2h\left(\frac{\rho^m}{m-1} - 1\right).$$

In figure 7.4, we see the evolution of the approximative solution $\rho_{h,m}$ for $m = 100$ and with a time step $h = 0.005$. We remark that the tumor starts to saturate the constraint in its support and then grows.

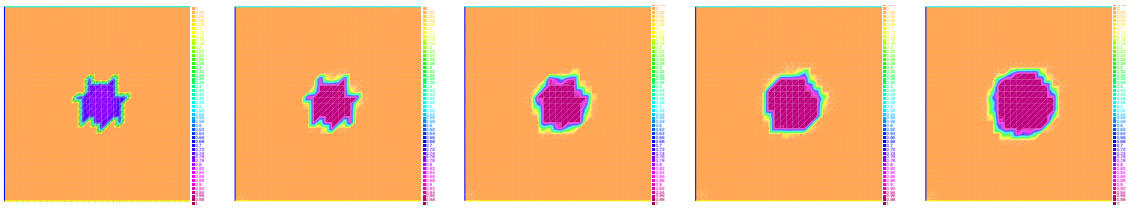


Figure 7.4: *Shape of approximative solution* $\rho_{h,m}$ *of equation* (7.58) *with* $m = 100$, $h = 0.005$.

7.2.4 Model of tumor growth with nutrient

Now, we want to apply this method to study models of tumor growth with nutriments, see [101],

$$\begin{cases} \partial_t \rho - \operatorname{div}(\rho \nabla p) = \rho\left((1-p)(c+c_1) - c_2\right), \\ \partial_t c - \Delta c = -\rho c, \\ p \geqslant 0, p(1-\rho) = 0, \rho \leqslant 1, \\ \rho_{|t=0} = \rho_0,\ c_{|t=0} = c_0, \end{cases} \tag{7.61}$$

on Ω a convex bounded subset of $\mathbb{R}^n$ and where c_1 and c_2 are two positive constants. This system is endowed with Neumann boundary conditions.

We introduce a semi-implicit splitting scheme to solve this problem. We construct four sequences $\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k}, c_{h,m}^{k+1/2}$ and $c_{h,m}^{k}$ such that for all $k \geqslant 0$,

$$\begin{cases} \rho_{h,m}^{k+1/2} \in \operatorname{argmin}_{\rho \in \mathcal{M}^+, |\rho| = |\rho_{h,m}^k|} \left\{ \frac{1}{2h} W_2^2(\rho, \rho_{h,m}^k) + \mathcal{F}_m(\rho) \right\}, \\ c_{h,m}^{k+1/2} \in \operatorname{argmin}_{c \in \mathcal{M}^+, |c| = |c_{h,m}^k|} \left\{ \frac{1}{2h} W_2^2(c, c_{h,m}^k) + \mathcal{E}(\rho) \right\}, \end{cases} \tag{7.62}$$

and

$$\begin{cases} \rho_{h,m}^{k+1} \in \operatorname{argmin}_{\rho \in \mathcal{M}^+} \left\{ \frac{1}{2h} FR^2(\rho, \rho_h^{k+1/2}) + \mathcal{E}_1(\rho | c_{h,m}^{k+1/2}) \right\}, \\ c_{h,m}^{k+1} \in \operatorname{argmin}_{c \in \mathcal{M}^+} \left\{ \frac{1}{2h} FR^2(c, c_{h,m}^{k+1/2}) + \mathcal{E}_2(c | \rho_{h,m}^{k+1/2}) \right\}, \end{cases} \tag{7.63}$$

where $\mathcal{E}, \mathcal{E}_1$ and $\mathcal{E}_2$ are defined as follow,

$$\mathcal{E}(\rho) = \int_\Omega \rho \log(\rho),$$

$$\mathcal{E}_1(\rho|c) := \int_\Omega \left(c_{h,m}^{k+1/2} + c_1\right) \frac{\rho^m}{m-1} + \int_\Omega (c_2 - c_{h,m}^{k+1/2} - c_1)\rho,$$

and

$$\mathcal{E}_2(c|\rho) := \int_\Omega \rho c.$$

These sequences are well-defined and we define $p_{h,m}^{k+1/2}$ and $p_{h,m}^{k+1}$ by $p_{h,m}^{k+1/2} := \frac{m}{m-1}(\rho_{h,m}^{k+1/2})^{m-1}$ and $p_{h,m}^{k+1} := \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}$. Then we denote $\rho_{h,m}, \tilde{\rho}_{h,m}, p_{h,m}, \tilde{p}_{h,m}, c_{h,m}$ and $\tilde{c}_{h,m}$ and the piecewise constant interpolations of $\rho_{h,m}^{k+1}, \rho_{h,m}^{k+1/2}, p_{h,m}^{k+1}, p_{h,m}^{k+1/2}, c_{h,m}^{k+1}$ and $c_{h,m}^{k+1/2}$ respectively.

Theorem 7.32. *Assume* $\rho_0 \in BV(\Omega)$ *with* $\rho_0 \leqslant 1$ *and* $c_0 \in L^\infty(\Omega)$. *Then* $\rho_{h,m}$ *and* $\tilde{\rho}_{h,m}$ *strongly converge to* ρ_∞ *in* $L^1((0,T)\times\Omega)$ *and* $c_{h,m}$ *and* $\tilde{c}_{h,m}$ *strongly converge to* c_∞ *in* $L^1((0,T)\times\Omega)$ *when* $h \searrow 0$ *and* $m \nearrow +\infty$. *Moreover, if* $mh \to 0$, *then* $p_{h,m}$ *and* $\tilde{p}_{h,m}$ *converge weakly to* p_∞ *respectively in* $L^2((0,T), H^1(\Omega))$ *and* $L^2((0,T)\times\Omega)$ *where* $(\rho_\infty, p_\infty, c_\infty)$ *is the solution of* (7.61).

Estimates and convergences

We start by recalling the optimality conditions for the scheme (7.62)-(7.63). The Euler-Lagrange equations satisfy by $\rho_{h,m}^{k+1}$ and $\rho_{h,m}^{k+1/2}$ are

$$\begin{cases} \rho_{h,m}^{k+1/2}\nabla p_{h,m}^{k+1/2} = \frac{\nabla\varphi}{h}\rho_{h,m}^{k+1/2}, \\ \sqrt{\rho_{h,m}^{k+1}} - \sqrt{\rho_{h,m}^{k+1/2}} = \frac{h}{2}\sqrt{\rho_{h,m}^{k+1}}\left((1-p_{h,m}^{k+1})(c_{h,m}^{k+1/2}+c_1)-c_2\right), \end{cases} \tag{7.64}$$

where φ is a Kantorovich potential for $W_2(\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k})$ and

$$\begin{cases} \nabla c_{h,m}^{k+1/2} = \frac{\nabla\psi}{h}c_{h,m}^{k+1/2}, \\ \sqrt{c_{h,m}^{k+1}} - \sqrt{c_{h,m}^{k+1/2}} = -\frac{h}{2}\sqrt{c_{h,m}^{k+1}}\rho_{h,m}^{k+1/2}, \end{cases} \tag{7.65}$$

with ψ a Kantorovich potential for $W_2(c_{h,m}^{k+1/2}, c_{h,m}^{k})$.

Using the optimality conditions for Fischer-Rao steps, we obtain directly the following estimate on the L^∞ bound.

Lemma 7.33. *For all $k \geqslant 0$,*

$$\|c_{h,m}^{k+1}\|_{L^\infty(\Omega)} \leqslant \|c_{h,m}^{k+1/2}\|_{L^\infty(\Omega)} \leqslant \|c_{h,m}^{k}\|_{L^\infty(\Omega)}.$$

In particular, for all $t \in \mathbb{R}^+$,

$$\|c_{h,m}(t,\cdot)\|_{L^\infty(\Omega)}, \|\tilde{c}_{h,m}(t,\cdot)\|_{L^\infty(\Omega)} \leqslant \|c_0\|_{L^\infty(\Omega)}.$$

Moreover,

$$\|\rho_{h,m}(t,\cdot)\|_\infty, \|\tilde{\rho}_{h,m}(t,\cdot)\|_\infty \leqslant 1.$$

Consequently,

$$\rho_{h,m}^{k+1} \leqslant (1+Ch)\rho_{h,m}^{k+1/2}.$$

Proof. The proof of estimates on $c_{h,m}$ and $\tilde{c}_{h,m}$ is obvious because one step of Wasserstein gradient flow with the Entropy decreases the L^∞-norm (see [98, 1]) and since the product $\sqrt{c_{h,m}^{k+1}}\rho_{h,m}^{k+1/2}$ is nonnegative, the L^∞-norm also decreases during the Fischer-Rao step. Moreover the proof for $\rho_{h,m}(t,\cdot)$ and $\tilde{\rho}_{h,m}(t,\cdot)$ is the same as in lemma 7.28. Now, using the fact that $\Phi(p_{h,m}^{k+1}, c_{h,m}^{k+1/2}) := (1-p_{h,m}^{k+1})(c_{h,m}^{k+1/2}+c_1)-c_2$ is bounded uniformly in k and arguing as in proposition 7.18, we conclude the proof. □

With these bounds it is easy to prove that

$$\begin{aligned} \mathcal{F}_m(\rho_{h,m}^{k+1}) &\leqslant \mathcal{F}_m(\rho_{h,m}^{k+1/2}) + Ch, \\ \mathcal{E}_1(\rho_{h,m}^{k+1/2}|c_{h,m}^{k+1/2}) - \mathcal{E}_1(\rho_{h,m}^{k+1}|c_{h,m}^{k+1/2}) &\leqslant Ch, \\ \mathcal{E}(c_{h,m}^{k+1}) &\leqslant \mathcal{E}(c_{h,m}^{k+1/2}) + Ch, \\ \mathcal{E}_2(c_{h,m}^{k+1/2}|\rho_{h,m}^{k+1/2}) - \mathcal{E}_2(c_{h,m}^{k+1}|\rho_{h,m}^{k+1/2}) &\leqslant Ch, \end{aligned}$$

see proposition 7.46. One readily checks that C does not depend on m. Then we obtain the usual Hölder estimates which imply that $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ converge to ρ_∞ in $L^\infty([0,T], L^1(\Omega))$ and $c_{h,m}$ and $\tilde{c}_{h,m}$ converge to c_∞ in $L^\infty([0,T], L^1(\Omega))$. Now we need to improve the convergence of these sequences. For $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$, this follows from

$$\sup_{t\in[0,T]} \left\{\|\rho_{h,m}(t,\cdot)\|_{BV(\Omega)}, \|\tilde{\rho}_{h,m}(t,\cdot)\|_{BV(\Omega)}\right\} \leqslant e^{CT}\|\rho_0\|_{BV(\Omega)}.$$

The proof is the same as in lemma 7.29. Then the Rossi-Savaré theorem, [107], gives that $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ converges strongly in $L^1((0,T)\times\Omega)$ to ρ_∞. Now using the Euler-Lagrange equation satisfied by $c_{h,m}^{k+1/2}$, we obtain that $\tilde{c}_{h,m}$ is bounded in $L^2((0,T),H^1(\Omega))$. Using again the Rossi-Savaré theorem, [107], we find that $\tilde{c}_{h,m}$ converges strongly in $L^1((0,T)\times\Omega)$ to c_∞. Now, since $\rho_{h,m}^{1/2}$ is bounded in $L^\infty(\Omega)$, for $c_{h,m}^{k+1}\geqslant(1-h)c_{h,m}^{k+1/2}$ (see the general case (7.43)) which implies that $c_{h,m}$ converges strongly in $L^1((0,T)\times\Omega)$ to c_∞.

To conclude the proof we have to charaterize the limits of $p_{h,m}^{k+1}$ and $\tilde{p}_{h,m}^{k+1}$. This can be done arguing as in the previous section. Moreover, all the nonlinear terms such as $\tilde{\rho}_{h,m}c_{h,m}$ and $\rho_{h,m}\Phi(p_{h,m},\tilde{c}_{h,m})$, with $\Phi(p,c)=(1-p)(c+c_1)-c_2$ converge weakly to $\rho_\infty c_\infty$ and $\rho_\infty\Phi(p_\infty,c_\infty)$ because $\rho_{h,m},\tilde{\rho}_{h,m},c_{h,m}$ and $\tilde{c}_{h,m}$ converge strongly and $p_{h,m}$ converge weakly.

Chapter 8

Systems of PDEs coupled with multi-marginal problems

Recently, Kinderlehrer, Monsaingeon and Xu proposed in [69] a gradient flow approach to solve the Poisson-Nernst-Planck system

$$\begin{cases} \partial_t \rho - \alpha \Delta \rho^m - \operatorname{div}(\rho \nabla (U + \varphi)) = 0, \\ \partial_t \mu - \beta \Delta \mu^m - \operatorname{div}(\mu \nabla (V - \varphi)) = 0, \\ -\Delta \varphi = \rho - \mu. \end{cases}$$

This system is used to model ionic transport of sereval interacting species, for example. Inspired by this work we are interested by a "nonlinear" version where species ρ and μ are coupled with the Monge-Ampère equation instead of the Poisson equation,

$$\begin{cases} \partial_t \rho - \alpha \Delta \rho^m - \operatorname{div}(\rho \nabla U) - \operatorname{div}(\rho \nabla \varphi) = 0, \\ \partial_t \mu - \beta \Delta \mu^m - \operatorname{div}(\mu \nabla V) - \operatorname{div}(\mu \nabla \varphi^c) = 0, \\ \det(I - D^2 \varphi) \mu (Id - \nabla \varphi) = \rho, \end{cases} \tag{8.1}$$

where φ^c is the c-transform of φ, $\varphi^c(x) = \sup_y |x-y|^2 - \varphi(y)$ and $|x|^2 - \varphi$ is convex. This kind of systems can arise in urban planning. In a series of work [25, 26, 36, 31, 32, 109, 108, 111] (list non-exhaustive), a static model of urban planning was proposed. A simplified model consists in considering an urban area region Ω where residents and services given by two probability densities on Ω, ρ_1 and ρ_2, want to minimize a quantity, $\mathcal{E}(\rho_1, \rho_2)$, to reach an ideal organization in the city. The total cost has to take into account a transportation cost between residential areas and service areas, a congestion effect for residential areas because the population does not want to live in very crowded area and on the contrary services want to be more concentrated in order to increase efficiency and decrease management costs. Particularly, the cost functional $\mathcal{E}$ can be taken as

$$\mathcal{E}(\rho_1, \rho_2) = W_c(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2), \tag{8.2}$$

where W_c is the value of an optimal transport problem with the cost c. The cost c can have several interpretations. For example, people working in services area want to be close to it in order to decrease car travel. $\mathcal{F}$ is an internal energy given by a convex superlinear function F,

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho(x))\, dx & \text{if } \mathcal{F}(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Since F is superlinear and convex, $\mathcal{F}$ can be rewritten as

$$\mathcal{F}(\rho) = \int_\Omega \frac{F(\rho)}{\rho} \rho,$$

with $\rho \mapsto \frac{F(\rho)}{\rho}$ is a increasing function which can be seen as the unhapiness of a citizen when he lives in a place where the population density is ρ.

And finally, $\mathcal{G}$ is on the form

$$\mathcal{G}(\rho) = \iint_{\Omega\times\Omega} h(|x-y|)\, d\rho(x) d\rho(y)$$

with h an increasing function modeling interactions between different services.

Since a city is always in evolution, it seems very natural to look for the gradient flow of $\mathcal{E}$ in the product of Wasserstein spaces. In the case where c is the quadratic cost, formaly we find to a system of the form (8.2) and φ is a Kantorovich potential of $W_2(\rho_1, \rho_2)$ which implies that is satisfies the Monge-Ampère equation

$$\det(I - D^2\varphi)\mu(Id - \nabla\varphi) = \rho.$$

In this chapter, we propose to investigate a generalization of problem (8.1). We extend to several species and it is very natural to assume that the cost is different for different populations. For example, workers have to pay for gas to go to work but firms do not. We will use the semi-implicit JKO scheme introduced in [54] to deal with these different costs. We remark that in (8.2), the interaction between densities is only in the transport term and we will see at the end possible extensions in the case of two sectors in the city.

The organization of the chapter is the following. In section 8.1.2, we explain our problem and state our main result. Sections 8.2, 8.3 and 8.4 are devoted to the demonstration of the existence of solutions for our evolution problem. The proof is based on a semi-implicit JKO scheme and on Rossi-Savaré theorem to obtain strong regularity. Using displacement convexity argument, we prove a uniqueness result for some class of functionals in section 8.5. In the final section 8.6.2, we propose some extensions and describe some open problems.

8.1 Preliminaries and main result

8.1.1 Multi-marginal transportation

In this part, we recall some results from the multi-marginal transport theory that we will used in the sequel. We refer to [100, 56] for a complete survey on this subject. The usual transport optimal can be extend to several marginals $\rho_1, \dots, \rho_l \in \mathcal{P}(\Omega)$, where Ω is a open bounded domain of $\mathbb{R}^n$. Let c be a cost function from Ω^l to $\mathbb{R}$, we define the multi-marginal transport problem $\mathcal{W}_c$ by

$$\mathcal{W}_c(\rho_1, \dots, \rho_l) := \inf\left\{ \int_{\Omega^l} c(x_1, \dots, x_l)\, d\gamma(x_1, \dots, x_l) \,:\, \gamma \in \Pi(\rho_1, \dots, \rho_l) \right\}.$$

The existence of an optimal transport plan is the same as in the usual optimal transport case. Then, as in the classical 2-marginal case, if we assume that c is continuous on $\overline{\Omega}^l$, the dual formulation holds

$$\mathcal{W}_c(\rho_1, \dots, \rho_l) = \sup\left\{ \sum_{i=1}^{l} \int_{\Omega} u_i(x_i)\, d\rho_i(x_i) \,:\, \sum_{i=1}^{l} u_i(x_i) \leqslant c(x_1, \dots, x_l) \right\}.$$

We called Kantorovich potential any $(u_1, \dots, u_l)$ optimal for the dual formulation. Any Kantorovich potentials $u_1, \dots, u_l$ are c-conjugate function i.e

$$u_i(x_i) = \inf\left\{ c(x_1, \cdot, x_l) - \sum_{j=1, j\neq i}^{l} u_j(x_j),\, x_j \in \Omega \right\}, \text{ for all } i = 1, \dots, l.$$

For any γ optimal transport plan and $(u_1, \dots, u_l)$ Kantorovich potential, we get

$$\sum_{i=1}^{l} u_i(x_i) = c(x_1, \dots, x_l) \qquad \gamma - a.e.$$

Moreover, if we assume that ρ_i is absolutely continuous with respect to the Lebesgue measure and c is differentiable in the i-th variable, then u_i are Lipschitz functions and

$$\nabla u_i(x_i) = \nabla_{x_i} c(x_1, \dots, x_l) \qquad \gamma - a.e.$$

8.1.2 Assumptions and main result

Let Ω be an open bounded subset of $\mathbb{R}^n$ and $l \in \mathbb{N}^*$. For all $i \in [\![1, l]\!]$, we define a cost function $c_i : \mathbb{R}^{nl} \to \mathbb{R}$ continuous on $\overline{\Omega}^l$ and differentiable with respect to x_i such that $\nabla_{x_i} c_i$ is continuous on $\overline{\Omega}^l$ and bounded on $\overline{\Omega}^l$.

The multi-marginal interaction energy $\mathcal{W}_i : \mathcal{P}(\Omega)^l \to \mathbb{R}$ is defined by

$$\mathcal{W}_i(\rho_1, \dots, \rho_l) := \inf \left\{ \int_{\Omega^l} c_i(x_1, \dots, x_l)\, d\gamma(x_1, \dots, x_l) \,:\, \gamma \in \Pi(\rho_1, \dots, \rho_l) \right\},$$

where $\Pi(\rho_1, \dots, \rho_l) := \left\{ \gamma \in \mathcal{P}^{\mathrm{ac}}(\Omega^l) \,:\, {\pi^i}_{\#}\gamma = \rho_i \right\}$ and π^i denotes the canonical projection from Ω^l to Ω.

For all i, let $F_i : \mathbb{R}^+ \to \mathbb{R}$ be a strictly convex superlinear function which satisfies

$$F_i(0) = 0, \text{ and } P_i(x) \leqslant C(1 + F_i(x)). \tag{8.3}$$

where $P_i(x) := xF_i'(x) - F_i(x)$ is the pressure associated to F_i.

The goal of this paper is to solve the following nonlinear diffusion system with nonlocal interactions:

$$\partial_t \rho_i = \Delta P_i(\rho_i) + \operatorname{div}(\rho_i \nabla u_i),\ \rho_{i|t=0} = \rho_{i,0} \tag{8.4}$$

on $(0, +\infty) \times \Omega$, for all $i \in [\![1, l]\!]$ where u_i is an optimal Kantorovich potential of $\mathcal{W}_i(\rho_1, \dots, \rho_l)$, so that

$$\nabla u_i(x_i) = \nabla_{x_i} c_i(x_1, \dots, x_l) \qquad \lambda_i - a.e.$$

with λ_i solving $\mathcal{W}_i(\rho_1, \dots, \rho_l)$. Since Ω is a bounded subset of $\mathbb{R}^n$, (8.4) is equipped with natural Neumann boundary condition on $\partial\Omega$ i.e

$$(\nabla P_i(\rho_i) + \nabla u_i \rho_i) \cdot \nu = 0 \qquad \text{on } \partial\Omega,$$

where ν is the outward normal to $\partial\Omega$.

A weak solution of (8.4) is a curve $t \in (0, +\infty) \mapsto (\rho_1(t), \dots, \rho_l(t)) \in \mathcal{P}^{\mathrm{ac}}(\Omega)^l$ such that $\nabla P_i(\rho_i) \in \mathcal{M}^n((0, T) \times \Omega)$, for all $T < +\infty$, and

$$\int_0^{+\infty} \left(\int_\Omega \partial_t \Phi \rho_i \, dx - \int_\Omega \nabla\Phi \cdot d\nabla P_i(\rho_i) - \int_{\Omega^l} \nabla_{x_i} c_i(x_1, \dots, x_l) \cdot \nabla\Phi(x_i)\, d\lambda_i(x_1, \dots, x_l) \right) dt$$
$$= - \int_\Omega \Phi(0, x) \rho_{i,0}(x)\, dx,$$

for every $\Phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$ and where λ_i is an optimal transport plan of $\mathcal{W}_i(\rho_1, \dots, \rho_l)$.

Since any Kantorovich potential u_i of $\mathcal{W}_i(\rho_1, \dots, \rho_l)$ satisfies

$$\nabla u_i(x_i) = \nabla_{x_i} c_i(x_1, \dots, x_l) \qquad \lambda_i - a.e,$$

we remark that

$$\int_{\Omega^l} \nabla_{x_i} c_i(x_1, \dots, x_l) \cdot \nabla\Phi(x_i)\, d\lambda_i(x_1, \dots, x_l) = \int_\Omega \nabla u_i(x_i) \cdot \cdot \nabla\Phi(x_i)\, d\rho_i(x_i).$$

Since the left-hand side is linear contrarily to the right-hand side, it is easier to pass to the limit in λ than in $\nabla u_i \rho_i$. That is why we choose this definition for the weak solution of (8.4).

Theorem 8.1. *If $\rho_{i,0} \in \mathcal{P}^{\mathrm{ac}}(\Omega)$ satisfy*

$$\mathcal{F}_i(\rho_{i,0}) < +\infty, \qquad \text{with} \qquad \mathcal{F}_i(\rho) := \begin{cases} \int_\Omega F_i(\rho(x))\, dx & \text{if } \rho \ll \mathcal{L}^n, \\ +\infty & \text{otherwise.} \end{cases} \tag{8.5}$$

Then (8.4) admits at least one weak solution.

8.2 Semi-implicit JKO scheme

In this section we introduce our semi-implicit JKO scheme and we find the usual estimates.
Let $h > 0$ be a time step, we construct l sequences with the following iterative discrete scheme: for all $i \in [\![1, l]\!]$, $\rho^0_{i,h} = \rho_{i,0}$ and for all $k \geqslant 0$, $\rho^{k+1}_{i,h}$ minimizes

$$\mathcal{E}_{i,h}(\rho|\boldsymbol{\rho}^k_h) := W_2^2(\rho, \rho^k_{i,h}) + 2h\left(\mathcal{F}_i(\rho) + \mathcal{W}^k_i(\rho)\right),$$

on $\rho \in \mathcal{P}^{ac}(\Omega)$, with $\mathcal{W}^k_i(\rho) := \mathcal{W}_i(\rho^k_{1,h}, \dots, \rho^k_{i-1,h}, \rho, \rho^k_{i+1,h}, \dots, \rho^k_{l,h})$ and $\boldsymbol{\rho}^k_h := (\rho^k_{1,h}, \dots, \rho^k_{l,h})$.

These sequences are well defined (the proof is the same as in the usual JKO scheme because of the regularity of c_i). Now we define the piecewise constant interpolations by, $\rho_{i,h}(0) = \rho_{i,0}$ and for all $t > 0$,

$$\rho_{i,h}(t) := \rho^{k+1}_{i,h} \text{ if } t \in (hk, h(k+1)]. \tag{8.6}$$

In the next proposition, we state usual estimates from JKO scheme.

Proposition 8.2. *For all $T > 0$. There exists $C > 0$ such that, for all h, k, with $hk < T$, $N = \lceil \frac{T}{h} \rceil$, for $i \in [\![1, l]\!]$, we have*

$$\mathcal{F}_i(\rho^k_{i,h}) \leqslant C, \tag{8.7}$$

$$\textstyle\sum_{k=0}^{N-1} W_2^2(\rho^k_{i,h}, \rho^{k+1}_{i,h}) \leqslant Ch. \tag{8.8}$$

Proof. We start to prove (8.8). Since $\rho^{k+1}_{i,h}$ is optimal in the minimization problem and $\rho^k_{i,h}$ is a competitor, we have

$$W_2^2(\rho^k_{i,h}, \rho^{k+1}_{i,h}) \leqslant 2h\left(\mathcal{F}_i(\rho^k_{i,h}) - \mathcal{F}_i(\rho^{k+1}_{i,h}) + \mathcal{W}^k_i(\rho^k_{i,h}) - \mathcal{W}^k_i(\rho^{k+1}_{i,h})\right). \tag{8.9}$$

Let γ the W_2-optimal transport plan between $\rho^{k+1}_{i,h}$ and $\rho^k_{i,h}$ and T the W_2-optimal transport map associated to γ i.e $\gamma = (I \times T)_\# \rho^{k+1}_{i,h}$. Let $\lambda \in \Pi\left(\rho^k_{1,h}, \dots, \rho^k_{i-1,h}, \rho^{k+1}_{i,h}, \rho^k_{i+1,h}, \dots, \rho^k_{l,h}\right)$ optimal for $\mathcal{W}_i$. We introduce λ_T defined by

$$\int_{\Omega^l} \varphi(x_1, \dots, x_l)\, d\lambda_T(x_1, \dots, x_l) := \int_{\Omega^l} \varphi(x_1, \dots, T(x_i), \dots, x_l)\, d\lambda(x_1, \dots, x_l).$$

By definition, $\lambda_T \in \Pi\left(\rho^k_{1,h}, \dots, \rho^k_{i-1,h}, \rho^k_{i,h}, \rho^k_{i+1,h}, \dots, \rho^k_{l,h}\right)$. Then,

$$\begin{aligned}
\mathcal{W}^k_i(\rho^k_{i,h}) - \mathcal{W}^k_i(\rho^{k+1}_{i,h}) &\leqslant \int_{\Omega^l} \left[c_i(x_1, \dots, T(x_i), \dots, x_l) - c_i(x_1, \dots, x_l)\right] d\lambda(x_1, \dots, x_l) \\
&\leqslant \|\nabla_{x_i} c_i\|_{L^\infty} \int_{\Omega^l} |T(x_i) - x_i|\, d\lambda(x_1, \dots, x_l) \\
&\leqslant CW_2(\rho^{k+1}_{i,h}, \rho^k_{i,h}),
\end{aligned}$$

because of the assumption on $\nabla_{x_i} c_i$ and Cauchy-Schwarz inequality. Then using this inequality in (8.9) and Young's inequality, we find

$$\frac{1}{2} W_2^2(\rho^k_{i,h}, \rho^{k+1}_{i,h}) \leqslant 2h\left(\mathcal{F}_i(\rho^k_{i,h}) - \mathcal{F}_i(\rho^{k+1}_{i,h}) + Ch\right).$$

Summing over k, we find

$$\begin{aligned}
\sum_{k=0}^{N-1} W_2^2(\rho^k_{i,h}, \rho^{k+1}_{i,h}) &\leqslant 4h\left(\sum_{k=0}^{N-1} (\mathcal{F}_i(\rho^k_{i,h}) - \mathcal{F}_i(\rho^{k+1}_{i,h})) + C^2T\right) \\
&\leqslant 4h\left(\mathcal{F}_i(\rho_{i,0}) - \mathcal{F}_i(\rho^N_{i,h}) + C^2T\right).
\end{aligned} \tag{8.10}$$

Since Ω is bounded, $\mathcal{F}_i$ is bounded from below and using the assumption (8.5), we conclude (8.8). The estimate (8.7) is obvious using (8.10).

□

8.3 Discrete system and stronger estimates

The goal of this section is to find the discrete system solved by the piecewise constant interpolations (8.6) and then we will deduce estimates on the BV-norm of the pressure.

Let $\lambda_{i,h}^{k+1}$ an optimal transport map for $\mathcal{W}_i\left(\rho_{1,h}^k,\dots,\rho_{i-1,h}^k,\rho_{i,h}^{k+1},\rho_{i+1,h}^k,\dots,\rho_{l,h}^k\right)$ and $u_{i,h}^{k+1}$ is a Kantorovich potentiel so we have

$$\nabla u_{i,h}^{k+1}(x_i)=\nabla_{x_i}c_i(x_1,\dots,x_l)\qquad \lambda_{i,h}^{k+1}-a.e. \tag{8.11}$$

Proposition 8.3. *For all $i\in[\![1,l]\!]$ and for all $k\geqslant 0$, we have*

$$h\left(\nabla u_{i,h}^{k+1}\rho_{i,h}^{k+1}+\nabla P_i(\rho_{i,h}^{k+1})\right)=-\nabla\varphi_{i,h}^{k+1}\rho_{i,h}^{k+1}\qquad a.e, \tag{8.12}$$

where $\varphi_{i,h}^{k+1}$ is the (unique) Kantorovich potential from $\rho_{i,h}^{k+1}$ to $\rho_{i,h}^k$ for W_2.

Proof. The proof is similar to chapter 4, proposition 4.18 or section 7 from [73]. We start by taking the first vartiation in the semi-implicit JKO scheme. Let $\xi\in\mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$ be given and Φ_τ the corresponding flow defined by

$$\partial_\tau\Phi_\tau=\xi\circ\Phi_\tau,\ \Phi_0=Id.$$

We define the pertubation ρ_τ of $\rho_{i,h}^{k+1}$ by $\rho_\tau:=\Phi_{\tau\#}\rho_{i,h}^{k+1}$. Then we get

$$\frac{1}{\tau}\left(\mathcal{E}_{i,h}(\rho_\tau|\boldsymbol{\rho}_h^k)-\mathcal{E}_{i,h}(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right)\geqslant 0. \tag{8.13}$$

By standard computations, we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(W_2^2(\rho_\tau,\rho_{i,h}^k)-W_2^2(\rho_{i,h}^{k+1},\rho_{i,h}^k))\leqslant\int_{\Omega\times\Omega}(x-y)\cdot\xi(x)\,d\gamma_{i,h}^{k+1}(x,y), \tag{8.14}$$

with $\gamma_{i,h}^{k+1}$ is an W_2-optimal transport plan in $\Gamma(\rho_{i,h}^{k+1},\rho_{i,h}^k)$ and $\gamma_{i,h}^{k+1}=(Id\times T_{i,h}^{k+1})_\#\rho_{i,h}^{k+1}$ with $T_{i,h}^{k+1}=Id-\nabla\varphi_{i,h}^{k+1}$. Moreover, using (8.3), (8.7) and Lebesgue's dominated convergence theorem, we obtains

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{F}_i(\rho_\tau)-\mathcal{F}_i(\rho_{i,h}^{k+1}))\leqslant-\int_\Omega P_i(\rho_{i,h}^{k+1}(x))\,\mathrm{div}(\xi(x))\,dx. \tag{8.15}$$

Finally,

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(W_i^k(\rho_\tau)-W_i^k(\rho_{i,h}^{k+1}))\leqslant\int_{\Omega^l}\nabla_{x_i}c_i(x_1,\dots,x_l)\cdot\xi(x_i)\,d\lambda_{i,h}^{k+1}(x_1,\dots,x_l), \tag{8.16}$$

If we combine (8.11), (8.13), (8.14), (8.15) and (8.16), and replacing ξ by $-\xi$, we find, for all $\xi\in\mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega\nabla\varphi_{i,h}^{k+1}\cdot\xi\rho_{i,h}^{k+1}-h\int_\Omega P_i(\rho_{i,h}^{k+1})\,\mathrm{div}(\xi)+h\int_\Omega\nabla u_{i,h}^{k+1}\cdot\xi\rho_{i,h}^{k+1}=0, \tag{8.17}$$

Now we claim that $P_i(\rho_{i,h}^{k+1})\in W^{1,1}(\Omega)$. Indeed, since P_i is controled by F_i, (8.7) gives $P_i(\rho_{i,h}^{k+1})\in L^1(\Omega)$. Moreover, using (8.17), we obtain

$$\left|\int_\Omega P_i(\rho_{i,h}^{k+1})\,\mathrm{div}(\xi)\right|\leqslant\left[\int_\Omega\frac{|\nabla\varphi_{i,h}^k(y)|}{h}\rho_{i,h}^{k+1}+\|\nabla_{x_i}c_i\|_{L^\infty}\right]\|\xi\|_{L^\infty(\Omega)}\leqslant\left[\frac{W_2(\rho_{i,h}^k,\rho_{i,h}^{k+1})}{h}+C\right]\|\xi\|_{L^\infty(\Omega)}.$$

By duality, this implies $P_i(\rho_{i,h}^{k+1})\in BV(\Omega)$ and $\nabla P_i(\rho_{i,h}^{k+1})=\left(-\nabla u_{i,h}^{k+1}\rho_{i,h}^{k+1}-\frac{\nabla\varphi_{i,h}^{k+1}}{h}\rho_{i,h}^{k+1}\right)$ in $\mathcal{M}^n(\Omega)$. In fact, $P_i(\rho_{i,h}^{k+1})$ is in $W^{1,1}(\Omega)$ because $\nabla u_{i,h}^{k+1}\rho_{i,h}^{k+1}+\frac{\nabla\varphi_{i,h}^k}{h}\rho_{i,h}^{k+1}\in L^1(\Omega)$ and then we proved (8.12).

□

The following proposition shows that $(\rho_{1,h},\dots,\rho_{l,h})$ is solution of a discrete approximation of system (8.4).

Proposition 8.4. *Let $h>0$, for all $T>0$, let N such that $Nh=T$ and for all $\phi\in\mathcal{C}_c^\infty([0,T)\times\Omega)$, then*

$$\begin{aligned}\int_0^T\int_\Omega \rho_{i,h}(t,x)\partial_t\phi(t,x)\,dxdt &= h\sum_{k=0}^{N-1}\int_\Omega \nabla P_i(\rho_{i,h}^{k+1}(x))\cdot\nabla\phi(t_k,x)\,dx\\ &+ h\sum_{k=0}^{N-1}\int_{\Omega^l}\nabla_{x_i}c_i(x_1,\dots,x_l)\cdot\nabla\phi(t_k,x_i)\,d\lambda_{i,h}^{k+1}(x_1,\dots,x_l)\\ &+ \sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\phi(t_k,\cdot)](x,y)d\gamma_{i,h}^{k+1}(x,y)\\ &- \int_\Omega\rho_{i,0}(x)\phi(0,x)\,dx,\end{aligned}$$

with, for all $\phi\in\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)|\leqslant\frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\Omega)}|x-y|^2,$$

$\gamma_{i,h}^{k+1}$ is an optimal transport plan in $\Gamma(\rho_{i,h}^k,\rho_{i,h}^{k+1})$.

Proof. This is a consequence of (8.12) (see chapter 4 proposition 4.18 or [73] section 7). □

Now we will obtain a stronger estimate to pass to the limit in the nonlinear diffusion.

Proposition 8.5. *For all $T>0$, we have*

$$\|P_i(\rho_{i,h})\|_{L^1((0,T);W^{1,1}(\Omega))}\leqslant CT. \tag{8.18}$$

Proof. If we integrate (8.12), we obtain

$$h\int_\Omega|\nabla P_i(\rho_{i,h}^{k+1})|\leqslant W_2(\rho_{i,h}^k,\rho_{i,h}^{k+1})+Ch,$$

Then we sum from $k=0$ to $N-1$ and thanks to (8.8), we have

$$\int_0^T\int_\Omega|\nabla P_i(\rho_{i,h})|\leqslant CT.$$

We conclude thanks to (8.3) and (8.7). □

8.4 Convergences and proof of main theorem

8.4.1 Weak and strong convergences

As usual, thanks to (8.8) and a rafined version of Arzelà-Ascoli theorem (see [4]), $\rho_{i,h}$ converges to ρ_i in $L^\infty([0,T];\mathcal{P}^{\mathrm{ac}}(\Omega))$, for all $T>0$.

Unfortunately, we need a strong convergence in order to pass to the limit in the nonlinear diffusion term. To do that we will apply an extension of Aubin-Lions lemma proved by Rossi and Savaré (theorem 2 from [107]). We start to introduce $\mathcal{G}_i:L^1(\Omega)\to(-\infty,+\infty]$ defined by

$$\mathcal{G}_i(\rho):=\begin{cases}\mathcal{F}_i(\rho)+\|P_i(\rho)\|_{BV(\Omega)} & \text{if } \rho\in\mathcal{P}^{\mathrm{ac}}(\Omega), F_i(\rho)\in L^1(\Omega) \text{ and } P_i(\rho)\in BV(\Omega),\\ +\infty & \text{otherwise.}\end{cases}$$

$\mathcal{G}_i$ is l.s.c and its sublevels are relatively compact in $L^1(\Omega)$. The l.s.c comes from the l.s.c of the BV-norm. Now we have to show that its sublevels are relativly compact in $L^1(\Omega)$. Let

$A_c := \left\{\rho \in L^1(\Omega) : \mathcal{G}_i(\rho) \leqslant c\right\}$. Let $(\rho_k)_k$ be a sequence in A_c then $P_i(\rho_k)$ is bounded in $BV(\Omega)$ thus $P_i(\rho_k)$ converges to f in $L^1(\Omega)$, because the injection of $BV(\Omega)$ into $L^1(\Omega)$ is compact. Since P_i is continuous, one to one and its inverse is continuous, ρ_k converges to $\rho := P_i^{-1}(f)$ a.e. Moreover, since $\mathcal{G}_i(\rho_k) \leqslant c$ and F_i is superlinear, using Vitali's convergence theorem, we obtain that ρ_k converges to ρ strongly in $L^1(\Omega)$.

Assumptions in the Rossi-Savaré theorem are satisfied because of (8.8), (8.7) and (8.18) then we can apply this theorem in $L^1(\Omega)$ and we find that, using a diagonal argument, up to a subsequence, for all i, $\rho_{i,h}$ converges in measure with respect to t in $L^1(\Omega)$ to ρ_i. Now since convergence in measure implies a.e convergence up to a subsequence, we may also assume that $\rho_{i,h}(t)$ converges strongly in $L^1(\Omega)$ to $\rho_i(t)$ t-a.e. Then Lesbegue's dominated convergence theorem implies that $\rho_{i,h}$ converges strongly in $L^1((0,T)\times\Omega)$ to ρ_i.

Now we want to prove that $P_i(\rho_{i,h})$ converges strongly to $P_i(\rho_i)$ in $L^1((0,T)\times\Omega)$. First, we know that $P_i(\rho_{i,h})$ is uniformly bounded in $L^\infty((0,T),L^1(\Omega))$, using (8.3), and thanks to (8.18), we have that $P_i(\rho_{i,h})$ is uniformly bounded in $L^1((0,T),W^{1,1}(\Omega))$. Then the Sobolev embedding gives that $P_i(\rho_{i,h})$ is uniformly bounded in $L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^{n/n-1}(\Omega))$. We recall now the lemma giving the uniform integrability of $P_i(\rho_{i,h})$ established in chapter 5 (lemma 5.14):

Lemma 8.6. *Let $p > 1$, $q := \frac{2p-1}{p}$ and $f \in L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^p(\Omega))$ then $f \in L^q((0,T)\times\Omega)$ and we have*

$$\|f\|^q_{L^q((0,T)\times\Omega)} \leqslant \|f\|^{q-1}_{L^\infty_t(L^1_x)} \|f\|_{L^1_t(L^p_x)}.$$

Then, using lemma 8.6, $P_i(\rho_{i,h})$ is uniformly bounded in $L^{(n+1)/n}((0,T)\times\Omega)$. This implies that $P_i(\rho_{i,h})$ is uniformly integrable and Vitali's convergence theorem gives that $P_i(\rho_{i,h})$ converges strongly to $P_i(\rho_i)$ in $L^1((0,T)\times\Omega)$.

Then we can conclude that $\nabla P_i(\rho_{i,h})$ converges to $\nabla P_i(\rho_i)$ in $\mathcal{M}^n((0,T)\times\Omega)$.

8.4.2 Convergence of $\mathcal{W}_i$-optimal transport plans

Before treating the convergence to the interaction term, we introduce several notations. We note $\lambda^{k+1}_{i,h}$ an optimal transport plan for $\mathcal{W}_i\left(\rho^k_{1,h},\dots,\rho^k_{i-1,h},\rho^{k+1}_{i,h},\rho^k_{i+1,h},\dots,\rho^k_{l,h}\right)$. Then we define the interpolation $\lambda_{i,h}$ by

$$\lambda_{i,h}(t) := \lambda^k_{i,h} \text{ if } t \in (h(k-1),hk].$$

We introduce the notation $\check{\boldsymbol{\rho}}^k_{i,h} := \rho^k_{1,h}\otimes\cdots\otimes\rho^k_{i-1,h}\otimes\rho^k_{i+1,h}\otimes\cdots\otimes\rho^k_{l,h}$ and we define the shifted interpolation

$$\check{\boldsymbol{\rho}}_{i,h}(t) := \check{\boldsymbol{\rho}}^k_{i,h} \text{ if } t \in (hk,h(k+1)] \text{ and } \check{\boldsymbol{\rho}}_{i,h}(0) := \check{\boldsymbol{\rho}}_{i,0} \text{ if } t = 0.$$

It is clear that $\check{\boldsymbol{\rho}}_{i,h}$ narrowly converges to $\check{\boldsymbol{\rho}}_i := \rho_1\otimes\cdots\otimes\rho_{i-1},\rho_{i+1}\otimes\cdots\otimes\rho_l$ in $L^\infty([0,T],\mathcal{P}^{\mathrm{ac}}(\Omega^{l-1}))$.

To simplify notations, we assume that $\lambda_{i,h}(t) \in \Pi\left(\rho_{i,h}(t),\check{\boldsymbol{\rho}}_{i,h}(t)\right)$ and if we identify $\lambda_{i,h}$ and $T^{-1}\lambda_{i,h}$, we can assume that $\lambda_{i,h} \in \mathcal{P}([0,T]\times\Omega\times\Omega^{l-1})$. $\lambda_{i,h}$ is tight in $\mathcal{P}([0,T]\times\Omega\times\Omega^{l-1})$ then narrowly converges to λ_i in $\mathcal{P}([0,T]\times\Omega\times\Omega^{l-1})$.

Moreover, if we note $\pi^{1,2},\pi^{1,3}$ the projections from $[0,T]\times\Omega\times\Omega^{l-1}$ to $[0,T]\times\Omega$ and $[0,T]\times\Omega^{l-1}$, with $\pi^{1,2}(t,x,y) = (t,x)$ and $\pi^{1,3}(t,x,y) = (t,y)$, and $\pi^1(t,x,y) = t$, then we have $\pi^{1,2}_\#\lambda_{i,h} = \rho_{i,h}(t)dt$, $\pi^{1,3}_\#\lambda_{i,h} = \check{\boldsymbol{\rho}}_{i,h}(t)dt$ and $\pi^1_\#\lambda_{i,h} = T^{-1}\mathcal{L}_{|[0,T]}$. If we pass to the limit, since $\rho_{i,h}(t)dt$ and $\check{\boldsymbol{\rho}}_{i,h}(t)dt$ narrowly converge to $\rho_i(t)dt$ and $\check{\boldsymbol{\rho}}_i(t)dt$, then we have $\pi^{1,2}_\#\lambda_i = \rho_i(t)dt$, $\pi^{1,3}_\#\lambda_i = \check{\boldsymbol{\rho}}_i(t)dt$ and $\pi^1_\#\lambda_i = T^{-1}\mathcal{L}_{|[0,T]}$.

Now, using lemma 5.3.2 of [4], we obtain that the disintegration of λ_i with respect to $T^{-1}\mathcal{L}_{|[0,T]}$, $\lambda_i(t)$, belongs to $\Pi(\rho_i(t),\check{\boldsymbol{\rho}}_i(t))$, t-a.e in $[0,T]$.

It remains to prove that $\lambda_i(t)$ is optimal for $\mathcal{W}_i\left(\rho_1(t),\dots,\rho_l(t)\right)$. We start to establish an approximation result for the optimal transport plan between $\rho_i(t)$ and $\check{\boldsymbol{\rho}}_i(t)$.

Lemma 8.7. *Let $\overline{\lambda}_i(t)$ an optimal transport plan for $\mathcal{W}_i(\rho_1(t),\dots,\rho_l(t))$. There exists a sequence of transport plans $\overline{\lambda}_{i,h}(t) \in \Pi(\rho_{i,h}(t), \check{\boldsymbol{\rho}}_{i,h}(t))$ such that*

$$\sup_{t\in[0,T]} W_1(\overline{\lambda}_i(t), \overline{\lambda}_{i,h}(t)) \to 0.$$

Proof. The proof follows the lines of the one of lemma 6.2 in [6]. Let $\gamma_i(t) \in \Pi(\rho_i(t), \rho_{i,h}(t))$ the optimal transport plan for W_2 and let $\check{\boldsymbol{\gamma}}_i(t) \in \Pi(\check{\boldsymbol{\rho}}_i(t), \check{\boldsymbol{\rho}}_{i,h}(t))$ the optimal transport plan for W_2. Let us disintegrate $\gamma_i(t)$ and $\check{\boldsymbol{\gamma}}_i(t)$ as $\gamma_i(t) = \rho_i(t) \otimes \gamma_i^x(t)$ and $\check{\boldsymbol{\gamma}}_i(t) = \check{\boldsymbol{\rho}}_i(t) \otimes \check{\boldsymbol{\gamma}}_i^y(t)$. Now we define $\overline{\lambda}_{i,h}(t)$ by

$$\overline{\lambda}_{i,h}(t) = \int_{\Omega\times\Omega} \gamma_i^x(t) \otimes \check{\boldsymbol{\gamma}}_i^y(t)\, \overline{\lambda}_i(t, dx, dy).$$

By construction, $\overline{\lambda}_{i,h}(t) \in \Pi(\rho_{i,h}(t), \check{\boldsymbol{\rho}}_{i,h}(t))$. Then we introduce π a transport plan between $\overline{\lambda}_i(t)$ and $\overline{\lambda}_{i,h}(t)$ defined by for all $\varphi \in \mathcal{C}(\Omega^4)$,

$$\int_{\Omega^4} \varphi(x,y,x',y')\, d\pi(x,y,x',y') = \int_{\Omega^2} \left(\int_{\Omega^2} \varphi(x,y,x',y')\, \gamma_i^x(t,dx') \check{\boldsymbol{\gamma}}_i^y(t,dy') \right) \overline{\lambda}_i(t,dx,dy).$$

Since $\pi \in \Pi(\overline{\lambda}_i(t), \overline{\lambda}_{i,h}(t))$ we have

$$\begin{aligned}
W_1(\overline{\lambda}_i(t), \overline{\lambda}_{i,h}(t)) &\leqslant \int_{\Omega^4} (|x-x'| + |y-y'|)\, d\pi(x,y,x',y') \\
&\leqslant \int_{\Omega^2} |x-x'| \gamma_i^x(t,dx') \rho_i(t,dx) + \int_{\Omega^2} |y-y'| \check{\boldsymbol{\gamma}}_i^y(t,dy') \check{\boldsymbol{\rho}}_i(t,dy) \\
&\leqslant \int_{\Omega^2} |x-x'| \gamma_i(t,dx,dx') + \int_{\Omega^2} |y-y'| \check{\boldsymbol{\gamma}}_i(t,dy,dy') \\
&\leqslant W_2(\rho_{i,h}(t), \rho_i(t)) + W_2(\check{\boldsymbol{\rho}}_{i,h}(t), \check{\boldsymbol{\rho}}_i(t)).
\end{aligned}$$

Now, using the fact that $\rho_{i,h}$ and $\check{\boldsymbol{\rho}}_{i,h}$ converge in $L^\infty([0,T], \mathcal{P}^{\mathrm{ac}}(\Omega))$ and in $L^\infty([0,T], \mathcal{P}^{\mathrm{ac}}(\Omega^{l-1}))$ to ρ_i and $\check{\boldsymbol{\rho}}_i$ respectively, we have the result. □

In the next proposition, we use the previous lemma to show that $\lambda_i(t)$ is optimal for $\mathcal{W}_i(\rho_1(t),\dots,\rho_l(t))$ t-a.e in $[0,T]$.

Proposition 8.8. *For almost every $t \in [0,T]$, $\lambda_i(t)$ is optimal for $\mathcal{W}_i(\rho_1(t),\dots,\rho_l(t))$.*

Proof. Let $\overline{\lambda}_i(t)$ an optimal transport plan for $\mathcal{W}_i(\rho_1(t),\dots,\rho_l(t))$. First of all, we define $\overline{\lambda}_{i,h}(t)$ as in lemma 8.7. Since $\lambda_{i,h}$ is optimal for $\mathcal{W}_i(\rho_{i,h}, \check{\boldsymbol{\rho}}_{i,h})$ and $\overline{\lambda}_{i,h}(t) \in \Pi(\rho_{i,h}, \check{\boldsymbol{\rho}}_{i,h})$, we have

$$\int_{\Omega\times\Omega^{l-1}} c_i(x,y) \overline{\lambda}_{i,h}(t,dx,dy) \geqslant \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \lambda_{i,h}(t,dx,dy).$$

So for all nonnegative function $\varphi \in \mathcal{C}^\infty([0,T])$ we get

$$\int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \overline{\lambda}_{i,h}(t,dx,dy)\varphi(t)\, dt \geqslant \int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \lambda_{i,h}(t,dx,dy)\varphi(t)\, dt.$$

Since Ω is bounded and according to lemma 8.7,

$$\int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \overline{\lambda}_{i,h}(t,dx,dy)\varphi(t)\, dt \to \int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \overline{\lambda}_i(t,dx,dy)\varphi(t)\, dt.$$

In addition, since $\lambda_{i,h}$ narrowly converges to λ_i in $\mathcal{P}([0,T]\times\Omega\times\Omega^{l-1})$, we have

$$\int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \lambda_{i,h}(t,dx,dy)\varphi(t)\, dt \to \int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y) \lambda_i(t,dx,dy)\varphi(t)\, dt.$$

And then

$$\int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y)\overline{\lambda}_i(t,dx,dy)\varphi(t)\,dt \geqslant \int_0^T \int_{\Omega\times\Omega^{l-1}} c_i(x,y)\lambda_i(t,dx,dy)\varphi(t)\,dt.$$

The inequality holds for all nonnegative function $\varphi \in \mathcal{C}^\infty([0,T])$, then we obtain for almost every $t \in [0,T]$

$$\mathcal{W}_i\left(\rho_1(t),\dots,\rho_l(t)\right) \geqslant \int_{\Omega\times\Omega^{l-1}} c_i(x,y)\lambda_i(t,dx,dy),$$

and the proof is concluded. □

8.4.3 Proof of theorem 8.1

In this section, we have to take the limit in the system of proposition 8.4. The linear term (with time derivative) and the diffusion term converge to the desired result because $\rho_{i,h}$ converges to ρ_i strongly in $L^1(]0,T[\times\Omega)$ and $\nabla P_i(\rho_{i,h})$ converges to $\nabla P_i(\rho_i)$ in $\mathcal{M}^n([0,T]\times\Omega)$. The remainder term goes to 0 when h goes to 0 because of (8.8). So we just have to check the convergence of multi-marginal interaction term. But, since $\lambda_{i,h}$ converges to λ_i in $\mathcal{P}([0,T]\times\Omega^l)$, we have

$$h\sum_{k=0}^{N-1}\int_\Omega \nabla_{x_i}c_i(x_1,\dots,x_l)\cdot\nabla\phi(t_k,x_i)\,d\lambda_{i,h}^{k+1}(x,y) \to \int_0^T\int_\Omega \nabla_{x_i}c_i(x_1,\dots,x_l)\cdot\nabla\phi(t,x_i)\,d\lambda_i(t,x,y)dt.$$

Moreover, $\lambda_i(t)$ is an optimal transport plan for $\mathcal{W}_i(\rho_1,\dots,\rho_l)$ then for any Kantorovich potential u_i,

$$\nabla u_i(x_i) = \nabla_{x_i}c_i(x_1,\dots,x_l) \qquad \lambda_i(t)-a.e,$$

which concludes the proof.

8.5 Uniqueness

We will give an uniqueness result based on a displacement convexity argument and some examples of problem where this condition is satisfied. In the following, Ω is assumed convex.

8.5.1 Displacement convexity in product Wasserstein space

In this section, we will establish several cases where the multi-marginal energy is geodesically convex in the product Wasserstein space.

Definition 8.9. *Let $\lambda \in \mathbb{R}$. A functional $\mathcal{W} : \mathcal{P}(\Omega)^l \to (-\infty,+\infty]$ is said λ-geodesically convex in $\mathcal{P}(\Omega)^l$ if for every $i \in [\![1,l]\!]$ and for every couple $(\mu_i^0,\mu_i^1) \in \mathcal{P}(\Omega)^2$*

$$\mathcal{W}(\mu_1^t,\dots,\mu_l^t) \leqslant (1-t)\mathcal{W}(\mu_1^0,\dots,\mu_l^0) + t\mathcal{W}(\mu_1^1,\dots,\mu_l^1) - \frac{\lambda}{2}t(1-t)W_2^2((\mu_1^0,\dots,\mu_l^0),(\mu_1^1,\dots,\mu_l^1)),$$

where μ_i^t is a constant speed geodesic between μ_i^0 and μ_i^1 and W_2 is the product distance on $\mathcal{P}(\Omega)^l$.

We say that $F : [0,+\infty) \to \mathbb{R}$ satisfy McCann's condition if

$$x \in (0,+\infty) \mapsto x^n F(x^{-n}) \text{ is convex nonincreasing.} \tag{8.19}$$

McCann showed in [93] that if F satisfy (8.19), then $\mathcal{F}$, defined by

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho) & \text{if } F(\rho)\in L^1(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$

is geodesically convex ($\lambda = 0$).

Now we will find some examples of multi-marginal energy which are λ-geodesically convex.

A general classe of functions in 1D: In this paragraph, Ω is a convex bounded subset of $\mathbb{R}$.

Proposition 8.10. *Let $h : \Omega \to \mathbb{R}$ be a convex function. The functional $\mathcal{W} : \mathcal{P}^{\mathrm{ac}}(\Omega) \times \mathcal{P}^{\mathrm{ac}}(\Omega) \to \mathbb{R}$ defined by*

$$\mathcal{W}(\mu,\nu) := \inf_{\gamma\in\Pi(\mu,\nu)} \int_{\Omega\times\Omega} h(x-y)\, d\gamma(x,y),$$

is geodesically convex in $\mathcal{P}^{\mathrm{ac}}(\Omega) \times \mathcal{P}^{\mathrm{ac}}(\Omega)$.

Proof. Let $\mu_0, \mu_1, \nu_0, \nu_1 \in \mathcal{P}^{\mathrm{ac}}(\Omega)$ and let μ_t and ν_t be constant speed geodesics between μ_0 to μ_1 and ν_0 to ν_1, i.e $\mu_t = {T_t^\mu}_{\#}\mu_0$ and $\nu_t = {T_t^\nu}_{\#}\nu_0$, where $T_t^\mu = (1-t)Id + tT^\mu$, $T_t^\nu = (1-t)Id + tT^\nu$ and T^μ and T^ν are respectively the optimal maps between μ_0 to μ_1 and ν_0 to ν_1.

Let γ_0 the optimal transport plan between μ_0 and ν_0. Since μ_0 is absolutely continuous, $\gamma_0 = (Id, S)_{\#}\mu_0$ where S is a non decreasing map. We define $\gamma_t = (T_t^\mu, T_t^\nu)_{\#}\gamma_0$ and we remark that $\gamma_1 = (T^\mu, T^\nu \circ S)_{\#}\mu_0$.

If we take $(T^\mu(x), T^\nu \circ S(x))$ and $(T^\mu(x), T^\nu \circ S(x'))$ in the support of γ_1 where $x, x' \in \operatorname{supp}\mu_0$ we have that $T^\mu(x) < T^\mu(x')$ imply that $x \leqslant x'$ because T^μ is non decreasing and then since $T^\nu \circ S$ is the composition of two non decreasing maps, we obtain $T^\nu \circ S(x) \leqslant T^\nu \circ S(x')$. Lemma 2.8 of [112] gives that γ_1 is the optimal transport plan between μ_1 and ν_1.

Then we have

$$\begin{aligned}
\mathcal{W}(\mu_t,\nu_t) &\leqslant \int_{\Omega\times\Omega} h(x-y)\, d\gamma_t(x,y)\\
&\leqslant \int_{\Omega\times\Omega} h(T_t^\mu(x) - T_t^\nu(y))\, d\gamma_0(x,y)\\
&\leqslant (1-t)\int_{\Omega\times\Omega} h(x-y)\, d\gamma_0(x,y) + t\int_{\Omega\times\Omega} h(T^\mu(x) - T^\nu(y))\, d\gamma_0(x,y)\\
&\leqslant (1-t)\mathcal{W}(\mu_0,\nu_0) + t\mathcal{W}(\mu_1,\nu_1).
\end{aligned}$$

□

Remark 8.11. *This result can not be generalized in higher dimension. Indeed, in dimension $n > 1$, it is well known that $W_2(\cdot,\sigma)$ is not λ-convex along geodesic on $\mathcal{P}(\Omega)$ (see example 9.1.5 from [4]).*

$-W_2^2$ in dimension n: In this paragraph, we will show that $-W_2^2$ is (-2)-geodesically convex on $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$. We denote $\Pi_0(\rho,\mu)$ the set of optimal transport plans for problem $W_2(\rho,\mu)$.

Proposition 8.12. *Let $\mu_1, \mu_2, \mu_3, \mu_4 \in \mathcal{P}(\Omega)$ and let $\mu^{1,2} \in \Pi_0(\mu_2,\mu_1)$, $\mu^{3,4} \in \Pi_0(\mu_3,\mu_4)$ and $\mu^{t,t} \in \Pi_0(\mu_t^{1\to2}, \mu_t^{3\to4})$, with $\mu_t^{1\to2} = ((1-t)x_2 + tx_1)_{\#}\mu^{1,2}$ and $\mu_t^{3\to4} = ((1-t)x_3 + tx_4)_{\#}\mu^{3,4}$. Then there exists a plan*

$$\boldsymbol{\mu} \in \Pi(\mu^{1,2},\mu^{3,4}) \subset \mathcal{P}(\Omega^4) \text{ such that } {\pi_t^{1\to2,3\to4}}_{\#}\boldsymbol{\mu} = \mu^{t,t},$$

where $\pi_t^{1\to2,3\to4}(x_1,x_2,x_3,x_4) := ((1-t)x_2 + tx_1, (1-t)x_3 + tx_4)$.

Proof. The proof follows the lines of the proof of lemma 7.3.1 in [4]. Let $\Sigma_t, \tilde{\Sigma}_t : \Omega^2 \to \Omega^2$ and $\Lambda_t : \Omega^4 \to \Omega^4$ be the homeomorphisms defined by

$$\Sigma_t(x_1,x_2) := (x_1, (1-t)x_2 + tx_1), \qquad \tilde{\Sigma}_t(x_3,x_4) := ((1-t)x_3 + tx_4, x_4),$$

and

$$\Lambda_t(x_1,x_2,x_3,x_4) := (x_1, (1-t)x_2 + tx_1, (1-t)x_3 + tx_4, x_4).$$

We notice that $\boldsymbol{\mu}$ has the desired properties if and only if $\boldsymbol{\nu} := \Lambda_{t\#}\boldsymbol{\mu}$ verifies

$${\pi^{1,2}}_{\#}\boldsymbol{\nu} = \Sigma_{t\#}\mu^{1,2},\ {\pi^{2,3}}_{\#}\boldsymbol{\nu} = \mu^{t,t} \text{ and } {\pi^{3,4}}_{\#}\boldsymbol{\nu} = \tilde{\Sigma}_{t\#}\mu^{3,4}. \tag{8.20}$$

Since $\Sigma_{t\#}\mu^{1,2} \in \Pi(\mu_2, \mu_t^{1\to2})$, $\mu^{t,t} \in \Pi(\mu_t^{1\to2}, \mu_t^{3\to4})$ and $\tilde{\Sigma}_{t\#}\mu^{3,4} \in \Pi(\mu_t^{3\to4}, \mu_4)$, gluing lemma proves that there exists a plan $\boldsymbol{\nu} \in \mathcal{P}(\Omega^4)$ fulfilling (8.20). In addition, existence of $\boldsymbol{\mu}$ is proved because Λ_t is inversible. □

Let $\mu_0, \nu_0, \mu_1, \nu_1 \in \mathcal{P}(\Omega)$, $\gamma_\mu \in \Pi_0(\mu_1, \mu_0)$ and $\gamma_\nu \in \Pi_0(\nu_0, \nu_1)$. We note μ_t and ν_t the constant speed geodesics connecting respectively μ_0 to μ_1 and ν_0 to ν_1. Let γ_t the optimal transport plan between μ_t and ν_t. Proposition 8.20 says that there exists a plan $\boldsymbol{\gamma}$ such that

$$\boldsymbol{\gamma} \in \Pi(\gamma_\mu, \gamma_\nu) \text{ and } \pi^{1\to 2,3\to 4}{}_{\#}\boldsymbol{\gamma} = \gamma_t.$$

By definition of $\boldsymbol{\gamma}$, we have

$$\begin{aligned} W_2^2(\mu_t, \nu_t) &= \int_{\Omega^4} |(1-t)x_2 + tx_1 - (1-t)x_3 - tx_4|^2 \, d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) \\ &= \int_{\Omega^4} |(1-t)(x_2 - x_3) + t(x_1 - x_4)|^2 \, d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4). \end{aligned}$$

Moreover, the Hilbertian identity

$$|(1-t)a + tb|^2 = (1-t)|a|^2 + t|b|^2 - t(1-t)|b-a|^2$$

gives

$$\begin{aligned} W_2^2(\mu_t, \nu_t) &= (1-t)\int_{\Omega^4} |x_2 - x_3|^2 d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) + t\int_{\Omega^4} |x_1 - x_4|^2 \, d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) \\ &\quad -t(1-t)\int_{\Omega^4} |x_2 - x_3 - x_1 + x_4|^2 d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4). \end{aligned}$$

But, since $\pi^{2,3}{}_{\#}\boldsymbol{\gamma} \in \Pi(\mu_0, \nu_0)$ and $\pi^{1,4}{}_{\#}\boldsymbol{\gamma} \in \Pi(\mu_1, \nu_1)$, we have

$$\begin{aligned} \int_{\Omega^4} |x_2 - x_3|^2 d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) &\geqslant W_2^2(\mu_0, \nu_0), \\ \int_{\Omega^4} |x_1 - x_4|^2 \, d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) &\geqslant W_2^2(\mu_1, \nu_1). \end{aligned}$$

In addition,

$$\begin{aligned} \int_{\Omega^4} |x_2 - x_3 - x_1 + x_4|^2 d\boldsymbol{\gamma}(x_1, x_2, x_3, x_4) &\leqslant 2\left[\int_{\Omega^4} |x_2 - x_1|^2 \, d\boldsymbol{\gamma} + \int_{\Omega^4} |x_4 - x_3|^2 \, d\boldsymbol{\gamma}\right] \\ &\leqslant 2\left[\int_{\Omega^2} |x_2 - x_1|^2 \, d\gamma_\mu + \int_{\Omega^2} |x_4 - x_3|^2 \, d\gamma_\nu\right] \\ &\leqslant 2\left[W_2^2(\mu_0, \mu_1) + W_2^2(\nu_0, \nu_1)\right] \\ &\leqslant 2W_2^2((\mu_0, \nu_0), (\mu_1, \nu_1)). \end{aligned}$$

Then, if we combine all these inequalities, we obtain

$$W_2^2(\mu_t, \nu_t) \geqslant (1-t)W_2^2(\mu_0, \nu_0) + tW_2^2(\mu_1, \nu_1) - 2t(1-t)W_2^2((\mu_0, \nu_0), (\mu_1, \nu_1)). \tag{8.21}$$

As far we know, the identification of others examples is still an open question.

8.5.2 Uniqueness of solution

For the purpose of this chapter, it is enough to work in $\mathcal{P}^{\text{ac}}(\Omega)$ because diffusion implies that solutions are absolutely continuous. First, we define the Fréchet subdifferential for $\mathcal{W} : \mathcal{P}^{\text{ac}}(\Omega)^l \to (-\infty, +\infty]$ by extending the definition in [4].

Definition 8.13. *Let $\mathcal{W} : \mathcal{P}^{\mathrm{ac}}(\Omega)^l \to (-\infty, +\infty]$ be a functional and let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_l) \in L^2((\mu_1, \dots, \mu_l), \Omega)$, i.e*

$$\sum_{i=1}^{l} \int_\Omega |\xi_i|^2 \mu_i < +\infty.$$

We say that $\boldsymbol{\xi}$ is in the Fréchet subdifferential $\partial \mathcal{W}(\mu_1, \dots, \mu_l)$ if

$$\liminf_{\boldsymbol{\nu} \to \boldsymbol{\mu}} \frac{\mathcal{W}(\boldsymbol{\nu}) - \mathcal{W}(\boldsymbol{\mu}) - \sum_{i=1}^{l} \int_\Omega \langle \xi_i(x), T_{\mu_i}^{\nu_i}(x) - x \rangle \mu_i(x)\, dx}{W_2(\boldsymbol{\nu}, \boldsymbol{\mu})} \geqslant 0, \tag{8.22}$$

where $\boldsymbol{\mu} := (\mu_1, \dots, \mu_l)$ and $T_{\mu_i}^{\nu_i}$ is the optimal transport map between μ_i and ν_i.

In the next proposition, we want to characterize the subdifferential of λ-geodesically convex functionals.

Proposition 8.14. *Let $\mathcal{W} : \mathcal{P}^{\mathrm{ac}}(\Omega)^l \to (-\infty, +\infty]$ be a λ-geodesically convex functional. Then a vector $\boldsymbol{\xi} \in L^2(\boldsymbol{\mu}, \Omega)$ belongs to the Fréchet subdifferential of $\mathcal{W}$ at $\boldsymbol{\mu}$ iff*

$$\mathcal{W}(\boldsymbol{\nu}) - \mathcal{W}(\boldsymbol{\mu}) \geqslant \sum_{i=1}^{l} \int_\Omega \langle \xi_i(x), T_{\mu_i}^{\nu_i}(x) - x \rangle \mu_i(x)\, dx + \frac{\lambda}{2} W_2^2(\boldsymbol{\nu}, \boldsymbol{\mu}), \tag{8.23}$$

for all $\boldsymbol{\nu} \in \mathcal{P}^{\mathrm{ac}}(\Omega)^l$. Moreover, if $\boldsymbol{\xi} \in \partial \mathcal{W}(\boldsymbol{\mu})$ and $\boldsymbol{\kappa} \in \partial \mathcal{W}(\boldsymbol{\nu})$ then

$$\sum_{i=1}^{l} \int_\Omega \langle \xi_i(x) - \kappa_i(T_{\mu_i}^{\nu_i}(x)), T_{\mu_i}^{\nu_i}(x) - x \rangle \mu_i(x)\, dx \leqslant -\lambda W_2^2(\boldsymbol{\mu}, \boldsymbol{\nu}). \tag{8.24}$$

Proof. The proof is the same as in the characterization by Variational inequalities and monotonicity in [4] p. 231.

□

Now we can find the following uniqueness result

Theorem 8.15. *Assume F_i satisfy (8.19) and $\mathcal{W}_i$ is a λ_i-geodesically convex functional. Let $\boldsymbol{\rho}^1 := (\rho_1^1, \dots, \rho_l^1)$ and $\boldsymbol{\rho}^2 := (\rho_1^2, \dots, \rho_l^2)$ two weak solutions of (8.4) with initial conditions $\rho_i^1(0, \cdot) = \rho_{i,0}^1$ and $\rho_i^2(0, \cdot) = \rho_{i,0}^2$. If for all $T < +\infty$,*

$$\int_0^T \sum_{i=1}^{l} \|v_{i,t}^1\|_{L^2(\rho_{i,t}^1)}\, dt + \int_0^T \sum_{i=1}^{l} \|v_{i,t}^2\|_{L^2(\rho_{i,t}^2)}\, dt < +\infty, \tag{8.25}$$

with, for $j \in \{1, 2\}$,

$$v_{i,t}^j := -\nabla F_i'(\rho_{i,t}^j) - \nabla u_i^j,$$

then for every $t \in [0, T]$,

$$W_2^2(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2) \leqslant e^{-\left(2 \sum_{i=1}^{l} \lambda_i\right) t} W_2^2(\boldsymbol{\rho}_0^1, \boldsymbol{\rho}_0^2),$$

Proof. Using theorem 5.24, corollary 5.25 from [112] and assumption (8.25), we obtain

$$\frac{d}{dt}\left(\frac{1}{2} W_2^2(\rho_{i,t}^1, \rho_{i,t}^2)\right) = \int_\Omega \langle x - T_{i,t}(x), v_{i,t}^1(x) - v_{i,t}^2(T_{i,t}(x)) \rangle \rho_i^1(x)\, dx,$$

where $T_{i,t}$ is the optimal transport map between $\rho_{i,t}^1$ and $\rho_{i,t}^2$. Since F_i satisfies McCann's condition, we have

$$\int_\Omega \langle x - T_{i,t}(x), \nabla F_i'(\rho_{i,t}^2)(T_{i,t}(x)) - \nabla F_i'(\rho_{i,t}^1(x)) \rangle \rho_{i,t}^1(x)\, dx \leqslant 0,$$

In addition, $\mathcal{W}_i$ is λ_i-geodesically convex then (8.24) gives

$$\sum_{i=1}^{l}\int_\Omega \langle \nabla u^1_{i,t}(x) - \nabla u^2_{i,t}(T^{\nu_i}_{\mu_i}(x)), T^{\nu_i}_{\mu_i}(x) - x\rangle \rho^1_{i,t}(x)\,dx \leqslant -\sum_{i=1}^{l}\lambda_i W_2^2(\boldsymbol{\rho}^1_t, \boldsymbol{\rho}^2_t).$$

Then summing over i and combining these inequalities, we obtain

$$\frac{d}{dt}\left(W_2^2(\boldsymbol{\rho}^1_t, \boldsymbol{\rho}^2_t)\right) \leqslant -\left(2\sum_{i=1}^{l}\lambda_i\right) W_2^2(\boldsymbol{\rho}^1_t, \boldsymbol{\rho}^2_t).$$

Gronwall's lemma concludes the proof. □

Remark 8.16. *Assumption* (8.25) *in theorem 8.15 is made to ensure the absolute continuity of* $W_2(\rho^1_{i,t}, \rho^2_{i,t})$ *and it can been checked using* (8.12) *(see for example proposition 4.24 from chapter 4 or [73]).*

8.6 Extensions and open problems

This section is devoted to extensions of this problem and open problems for future work on this subject.

8.6.1 Urban planning

We will propose some extensions of the previous result with applications to urban planning. In the introduction of this chapter, we briefly discuss about one model of urban planning where residents, ρ_1, and services, ρ_2, want to minimize a cost functional

$$\mathcal{E}(\rho_1, \rho_2) = W_c(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2),$$

with $c : \mathbb{R}^n \to \mathbb{R}$ is a λ-convex function,

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega \rho(x)^2\,dx & \text{if } \mathcal{F}(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$\mathcal{G}(\rho) = \alpha \iint_{\Omega\times\Omega} |x-y|^2\,d\rho(x)d\rho(y).$$

In [25, 26, 109, 108], the authors proved that the functional $\mathcal{E}$ admits minimizers (ρ_1^*, ρ_2^*) which are charatherized by

- ρ_1^* is concentrated on a ball $B(x_0, r_\alpha) \cap \Omega$ and is given by

$$\rho_1^*(x) = \frac{\alpha}{2\alpha+1}(r_\alpha^2 - |x-x_0|^2),$$

  where x_0 is the barycentre of both ρ_1^* and ρ_2^*,

- ρ_2^* is concentrated on the ball $B(x_0, r_\alpha/(2\alpha+1)) \cap \Omega$ and it is the image of ρ_1^* under the homothety of ratio $1/(2\alpha+1)$ and centre x_0.

The first question we have in mind is: Do the solutions given by theorem 8.1 for the energy $\mathcal{E}$ converge to minimizers of $\mathcal{E}$ when t goes to $+\infty$?

We recall that the gradient flow of energy $\mathcal{E}$ is a weak solution, (ρ_1, ρ_2), of system

$$\begin{cases} \partial_t\rho_1 - \Delta\rho_1^2 - \operatorname{div}(\rho_1\nabla\varphi) = 0, \\ \partial_t\rho_2 - 2\alpha\operatorname{div}(\rho_2(|\cdot| * \rho_2)) - \operatorname{div}(\rho_2\nabla\psi) = 0, \end{cases}$$

with (φ, ψ) is a pair of optimal Kantorovich potentials for $W_c(\rho_1, \rho_2)$. In dimension $n = 1$, we proved that W_c is geodesically convex if c is convex then $\mathcal{E}$ is displacement convex in $\mathcal{P}^{\mathrm{ac}}(\Omega) \times \mathcal{P}^{\mathrm{ac}}(\Omega)$. Moreover, theorem 8.15 gives a contraction principle in the Wasserstein product space for displacement convex energies which implies that

$$W_2((\rho_1(t), \rho_2(t)), (\rho_1^*, \rho_2^*)) \leqslant e^{-\lambda t} W_2((\rho_1(0), \rho_2(0)), (\rho_1^*, \rho_2^*)).$$

Then if λ is positive, $(\rho_1(t), \rho_2(t))$ converges to minimizers of $\mathcal{E}$ when t goes to $+\infty$. In higher dimension, we can not apply this argument because even if c is λ-convex, a priori W_c is not geodesically convex. However, in the very simple case for two species where $c = \frac{|\cdot|^2}{2}$, $V_1 = V_2 = V$ and $\mathcal{F}$ is the Entropy, then denoting (ρ_1, ρ_2) solution of

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \operatorname{div}(\rho_1 \nabla V) - \operatorname{div}(\rho_1 \nabla \varphi_1) = 0, \\ \partial_t \rho_2 - \Delta \rho_2 - \operatorname{div}(\rho_2 \nabla V) - \operatorname{div}(\rho_2 \nabla \varphi_2) = 0, \end{cases}$$

where φ_1, φ_2 are the optimal Kantorovich potential (up to a constant) for $W_2(\rho_1, \rho_2)$, we can use the same argument that in theorem 8.15 to find the asymptotic behaviour of ρ_1 and ρ_2. Indeed,

$$\begin{aligned} \frac{d}{dt}\left(\frac{1}{2} W_2^2(\rho_1(t), \rho_2(t))\right) &\leqslant \int (x - T(x)) \cdot (\nabla \log(\rho_1(x)) - \nabla \log(\rho_2(T(x)))) \rho_1(x)\, dx \\ &+ \int (x - T(x)) \cdot (\nabla V(x) - \nabla V(T(x))) \rho_1(x)\, dx \\ &+ \int (x - T(x)) \cdot (\nabla \varphi_1(x) - \nabla \varphi_2(T(x))) \rho_1(x)\, dx, \end{aligned}$$

where T is the optimal transport map for $W_2(\rho_1, \rho_2)$. We remark that since ρ_1 and ρ_2 are positive densities,

$$\nabla \varphi_2(T(x)) = T(x) - S(T(x)) = T(x) - x = \nabla \varphi_1(x) \qquad a.e,$$

with S the optimal transport map for $W_2(\rho_2, \rho_1)$. If we assume that V is λ-convex, for some $\lambda > 0$, and using the displacement convexity of the Entropy and the potential energy then we obtain

$$\frac{d}{dt}\left(\frac{1}{2} W_2^2(\rho_1(t), \rho_2(t))\right) \leqslant -\lambda W_2^2(\rho_1(t), \rho_2(t)),$$

which implies that $\rho_1(t)$ and $\rho_2(t)$ converge to the same limit, when $t \nearrow +\infty$, which is the minimizer of

$$\rho \mapsto \int \rho \log(\rho) + \int V \rho.$$

In the previous model, we only focus on one sector, but we can assume that residential areas and service areas are split in two sectors denoted 1 and 2. This model was proposed in 2008 and studied in the static case, by Carlier, Ekeland and Rochet [34]. Here we are more interested in the evolution case. We denote f_1 and f_2 the densities of firms and w_1 and w_2 the densities of residents in these areas. Now the congestion costs is on the all population which implies that the congestion fonctional is

$$\mathcal{F}(w, f) := \frac{\alpha_w}{2} \int (w_1 + w_2)^2 + \frac{\alpha_f}{2} \int (f_1 + f_2)^2,$$

where $w := (w_1, w_2)$ and $f := (f_1, f_2)$ and α_w and α_f represent positive parameters. Moreover all residents in sector i want to work or use services from the same sector consequently, w and f have to minimize

$$\mathcal{W}_c(w, f) := \beta W_c(w_1, f_1) + \beta W_c(w_2, f_2).$$

We have to take into account externalities created by residents and firms, so they have to minimize

$$\begin{aligned} \mathcal{G}(w, f) := \sum_{i=1}^{2} \frac{\kappa_i}{2} \iint |x - y|^2 f_i(x) f_i(y)\, dxdy + \frac{\sigma_i}{2} \iint |x - y|^2 f_i(x)(w_1(y) + w_2(y))\, dxdy \\ + \frac{\theta}{4} \iint |x - y|^2 (w_1(x) + w_2(x))(w_1(y) + w_2(y))\, dxdy, \end{aligned}$$

where $\kappa_1, \kappa_2, \theta$ are positive parameters and σ_1 and σ_2 can be positive (services) or negative (polluting industries). Under some assumptions in parameters, this minimizing problem admits at least one solution. In future, we want to combine arguments developped at chapter 6, to study the existence of a gradient flow in a Wasserstein product space for the functional $\mathcal{F} + \mathcal{W}_c + \mathcal{G}$ and the asymptotic convergence.

8.6.2 Entropic regularization

In this section, we present future works on this subject with numerical applications. Indeed, the computation of the gradient flow of (8.4) can be very challenging because at each time step we have two solve two transport problems. To deal with this issue, we want to introduce a smoothing regularization. For simplicity, we restraint ourselves to the case of one density which interacts with a prescribed measure σ. In other words, (8.4) becomes

$$\rho_h^{k+1} \in \operatorname{argmin} \frac{1}{2h} W_2^2(\rho, \rho_h^k) + \mathcal{F}(\rho) + \mathcal{W}_\varepsilon(\rho, \sigma),$$

where $\mathcal{W}_\varepsilon$ is given by

$$\mathcal{W}_\varepsilon(\rho, \sigma) := \inf_{\gamma \in \Pi(\rho,\sigma)} \iint_{\Omega\times\Omega} \frac{|x-y|^p}{p}\, d\gamma(x,y) + \varepsilon \iint_{\Omega\times\Omega} \gamma \log(\gamma).$$

We remark that $\mathcal{W}_\varepsilon$ can be rewritten has the infimum over transport plan of the relative entropy with respect to a Gibbs measure $G(x,y) = \exp(-|x-y|^p/p\varepsilon)$,

$$\mathcal{W}_\varepsilon(\rho, \sigma) = \inf_{\gamma \in \Pi(\rho,\sigma)} \mathcal{H}(\gamma|G),$$

with

$$\mathcal{H}(\gamma|G) := \int_{\Omega\times\Omega} \frac{\gamma}{G} \log\left(\frac{\gamma}{G}\right) dG(x,y).$$

This problem was introduced by Schrödinger in [113] (see for example [78] for a survey on this problem). In [77] and more recently in [30] it was proved the following Γ-convergence result

Theorem 8.17. *Let $\rho, \mu \in \mathcal{P}_p(\mathbb{R}^n)$ with finite entropy, then*

$$\mathcal{F}_\varepsilon : \gamma \mapsto \begin{cases} \iint_{\Omega\times\Omega} \frac{|x-y|^p}{p}\, d\gamma(x,y) + \varepsilon \iint_{\Omega\times\Omega} \gamma\log(\gamma) & \text{if } \gamma \in \Pi(\rho,\mu), \\ +\infty & \text{otherwise,} \end{cases}$$

Γ-converges to

$$\mathcal{F} : \gamma \mapsto \begin{cases} \iint_{\Omega\times\Omega} \frac{|x-y|^p}{p}\, d\gamma(x,y) & \text{if } \gamma \in \Pi(\rho,\mu), \\ +\infty & \text{otherwise,} \end{cases}$$

with respect to the weak topology on $\mathcal{P}_p(\mathbb{R}^n)$ i.e

- *for any sequence $(\gamma_\varepsilon) \subset \mathcal{P}(\Omega\times\Omega)$ converging weakly to $\gamma \in \mathcal{P}(\Omega\times\Omega)$, we have*

$$\liminf_{n\to+\infty} \mathcal{F}_\varepsilon(\gamma_\varepsilon) \geqslant \mathcal{F}(\gamma),$$

- *for any $\gamma \in \mathcal{P}(\Omega\times\Omega)$, there exists $(\gamma_\varepsilon) \subset \mathcal{P}(\Omega\times\Omega)$ converging to γ in $\mathcal{P}(\Omega\times\Omega)$ such that*

$$\limsup_{n\to+\infty} \mathcal{F}_\varepsilon(\gamma_\varepsilon) \leqslant \mathcal{F}(\gamma).$$

Consequently, a sequence of minimizer of $\mathcal{W}_\varepsilon(\rho,\mu)$ converges to an optimal transport plan for $W_p(\rho,\mu)$.

In [104], Peyré used the entropic regularization to solve gradient flows in Wasserstein space. He regularized the W_2 metric which permit to have a very simple algorithm to solve each JKO step. In addition, he applied this algorithm to obtain numerical simulations of sytem with two species interacting with the W_2 distance between each others as system (8.1). In a future work, we want to extend this method to several species and combine this with the Entropy of the sum to solve numerically the evolution for the two-sectors case as it was introduced in previously.

In our setting, replaced the $W_p^p(\rho,\sigma)$ by $\mathcal{W}_\varepsilon(\rho,\sigma)$ means that the population ρ wants to reach σ but is subject to a noise. In the following, we will discuss about future works on gradient flow of $\mathcal{W}_\varepsilon$ with ε fixed and when ε goes to 0. Our first aim is to prove the convergence of the JKO scheme for $\mathcal{W}_\varepsilon$ with a fixed ε. One difficulty in the previous proof is the identification of the limit of optimal transport plans λ_h for $\mathcal{W}_\varepsilon(\rho_h,\sigma)$ as an optimal transport plan for $\mathcal{W}_\varepsilon(\rho,\sigma)$. To solve this problem, we would like to prove the following Γ-convergence result:

Conjecture 8.18. *If ρ_n converges to ρ in $L^1(\Omega)$. We define $\mathcal{G}_n$ by*

$$\mathcal{G}_n(\gamma) := \begin{cases} \mathcal{H}(\gamma|G) & \text{if } \gamma \in \Pi(\rho_n,\sigma), \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$\mathcal{G}(\gamma) := \begin{cases} \mathcal{H}(\gamma|G) & \text{if } \gamma \in \Pi(\rho,\sigma), \\ +\infty & \text{otherwise,} \end{cases}$$

where $G(x,y) = \exp(-|x-y|^p/p\varepsilon)$. Then $\mathcal{G}_n$ Γ-converges to $\mathcal{G}$ i.e

- *for any sequence $(\gamma_n) \subset \mathcal{P}(\Omega\times\Omega)$ converging weakly to $\gamma \in \mathcal{P}(\Omega\times\Omega)$, we have*

$$\liminf_{n\to+\infty} \mathcal{G}_n(\gamma_n|G) \geqslant \mathcal{G}(\gamma|G),$$

- *for any $\gamma \in \mathcal{P}(\Omega\times\Omega)$, there exists $(\gamma_n) \subset \mathcal{P}(\Omega\times\Omega)$ converging to γ in $\mathcal{P}(\Omega\times\Omega)$ such that*

$$\limsup_{n\to+\infty} \mathcal{G}_n(\gamma_n|G) \leqslant \mathcal{G}(\gamma|G).$$

Indeed, if we prove this conjecture, arguing as in proposition 8.8, we obtain the desired result. The liminf condition is not the issue here and is easy to prove by lower semi-continuity of the relative entropy. For the limsup condition, we will give a formal proof when ρ_n and ρ are bounded from below.

Formal proof. If $\gamma \notin \Pi(\rho,\sigma)$ then for n large enough, $\gamma_n \notin \Pi(\rho_n,\sigma)$. Then we assume that $\gamma_n \in \Pi(\rho_n,\sigma)$ and $\gamma \in \Pi(\rho,\sigma)$. Then the liminf inequalities comes directly from the duality formula of the relative entropy, see lemma 9.4.4 of [4],

$$\mathcal{H}(\gamma|G) = \sup\left\{ \int_{\Omega\times\Omega} S\,d\gamma - \int_{\Omega\times\Omega} (e^S - 1)\,dG \;:\; S \in \mathcal{C}_b(\Omega\times\Omega) \right\}.$$

Now we give a formal proof for the limsup condition. Assume that ρ_n and ρ are smooth, $\mathcal{C}^{0,\alpha}(\Omega)$, and bounded from below. Then the transport map T_n from ρ to ρ_n is given by the gradient of a convex function u_n such that $u_n \in \mathcal{C}^{2,\alpha}(\Omega)$ and solves the Monge-Ampère equation

$$\rho_n(\nabla u_n(x)) \det(D^2 u_n(x)) = \rho(x) \qquad \text{for all } x,$$

according to Caffarelli's regularity theorem 2.6 or [27, 28]. Since ρ_n converges to ρ a.e, T_n converges to I a.e. We define γ_n by $\gamma_n := (T_n, I)_\#\gamma$. Then we have

$$\gamma_n(T_n(x), y) \det(\nabla T_n(x)) = \gamma(x,y).$$

We obtain

$$\begin{aligned}\mathcal{H}(\gamma_n|G) &= \iint_{\Omega\times\Omega} \frac{\gamma(T_n^{-1}(x),y)}{\det(\nabla T_n(T_n^{-1}(x)))} \log\left(\frac{\gamma(T_n^{-1}(x),y)}{G(x,y)\det(\nabla T_n(T_n^{-1}(x)))}\right) dxdy \\ &= \iint_{\Omega\times\Omega} \gamma(x,y)\log\left(\frac{\gamma(x,y)}{G(T_n(x),y)\det(\nabla T_n(x))}\right) dxdy \\ &= \iint_{\Omega\times\Omega} \gamma(x,y)\log\left(\frac{\gamma(x,y)}{G(T_n(x),y)}\right) dxdy - \iint_{\Omega\times\Omega} \gamma(x,y)\log(\det(\nabla T_n(x)))\,dxdy.\end{aligned}$$

Since T_n converges a.e to I, we have

$$\limsup_{n\to+\infty} \iint_{\Omega\times\Omega} \gamma(x,y)\log\left(\frac{\gamma(x,y)}{G(T_n(x),y)}\right) dxdy = \mathcal{H}(\gamma|G).$$

Moreover, at least formally, $\det(\nabla T_n(x)) \to 1$, which implies that

$$\limsup_{n\to+\infty} \left(-\iint_{\Omega\times\Omega} \gamma(x,y)\log(\det(\nabla T_n(x)))\,dxdy\right) = 0.$$

Then we obtain

$$\limsup_{n\to+\infty} \mathcal{H}(\gamma_n|G) \leqslant \mathcal{H}(\gamma|G).$$

□

To conclude, we hope to obtain rigourously this result and there are several others issues as the regularity of Kantorovich potentials of the regularized problem and their stability with respect to marginals that we need to understand in order to have an existence result for the entropic smoothing problem.

Bibliography

[1] Martial Agueh. Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Adv. Differential Equations*, 10(3):309–360, 2005.

[2] Damon Alexander, Inwon Kim, and Yao Yao. Quasi-static evolution and congested crowd transport. *Nonlinearity*, 27(4):823–858, 2014.

[3] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[5] Jean-Pierre Aubin. Un théorème de compacité. *C. R. Acad. Sci. Paris*, 256:5042–5044, 1963.

[6] Jean-Bernard Baillon and Guillaume Carlier. From discrete to continuous Wardrop equilibria. *Netw. Heterog. Media*, 7(2):219–241, 2012.

[7] A. Bakhta and V. Ehrlacher. Global existence of bounded weak solutions to degenerate cross-diffusion equations in moving domain. *ArXiv e-prints*, August 2015.

[8] J. D. Benamou and Y. Brenier. Mixed L^2-Wasserstein optimal mapping between prescribed density functions. *J. Optim. Theory Appl.*, 111(2):255–271, 2001.

[9] J.-D. Benamou, G. Carlier, Q. Mérigot, and É. Oudet. Discretization of functionals involving the monge-ampère operator, 2014.

[10] Jean-David Benamou. Numerical resolution of an "unbalanced" mass transport problem. *M2AN Math. Model. Numer. Anal.*, 37(5):851–868, 2003.

[11] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[12] Jean-David Benamou, Yann Brenier, and Kevin Guittet. Numerical analysis of a multi-phasic mass transport problem. *Contemporary Mathematics*, 353:1–18, 2004.

[13] Jean-David Benamou and Guillaume Carlier. Augmented Lagrangian Methods for Transport Optimization, Mean Field Games and Degenerate Elliptic Equations. *J. Optim. Theory Appl.*, 167(1):1–26, 2015.

[14] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.

[15] Jean-David Benamou, Francis Collino, and Jean-Marie Mirebeau. Monotone and consistent discretization of the monge-ampere operator. *arXiv preprint arXiv:1409.6694*, 2014.

[16] Jean-David Benamou, Brittany D. Froese, and Adam M. Oberman. Numerical solution of the optimal transportation problem using the Monge-Ampère equation. *J. Comput. Phys.*, 260:107–126, 2014.

[17] Benamou, Jean-David, Carlier, Guillaume, and Laborde, Maxime. An augmented lagrangian approach to wasserstein gradient flows and applications. *ESAIM: ProcS*, 54:1–17, 2016.

[18] Adrien Blanchet, Vincent Calvez, and José A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.

[19] Guy Bouchitté and Giuseppe Buttazzo. New lower semicontinuity results for nonconvex functionals defined on measures. *Nonlinear Anal.*, 15(7):679–692, 1990.

[20] Andrea Braides. Γ*-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[21] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[22] Martin Burger, José A. Carrillo, and Marie-Therese Wolfram. A mixed finite element method for nonlinear diffusion equations. *Kinet. Relat. Models*, 3(1):59–83, 2010.

[23] Martin Burger, Marzena Franek, and Carola-Bibiane Schönlieb. Regularized regression and density estimation based on optimal transport. *Appl. Math. Res. Express. AMRX*, (2):209–253, 2012.

[24] G. Buttazzo, C. Jimenez, and E. Oudet. An optimization problem for mass transportation with congested dynamics. *SIAM J. Control Optim.*, 48(3):1961–1976, 2009.

[25] Giuseppe Buttazzo and Filippo Santambrogio. A model for the optimal planning of an urban area. *SIAM J. Math. Anal.*, 37(2):514–530 (electronic), 2005.

[26] Giuseppe Buttazzo and Filippo Santambrogio. A mass transportation model for the optimal planning of an urban region. *SIAM Rev.*, 51(3):593–610, 2009.

[27] Luis A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.

[28] Luis A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[29] Clément Cancès and Cindy Guichard. Entropy-diminishing CVFE scheme for solving anisotropic degenerate diffusion equations. In *Finite volumes for complex applications. VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 187–196. Springer, Cham, 2014.

[30] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer. Convergence of Entropic Schemes for Optimal Transport and Gradient Flows. *ArXiv e-prints*, December 2015.

[31] G. Carlier and I. Ekeland. The structure of cities. *J. Global Optim.*, 29(4):371–376, 2004.

[32] G. Carlier and I. Ekeland. Equilibrium structure of a bidimensional asymmetric city. *Nonlinear Anal. Real World Appl.*, 8(3):725–748, 2007.

[33] G. Carlier and M. Laborde. A splitting method for nonlinear diffusions with nonlocal, non-potential drifts. *Nonlinear Analysis: Theory, Methods & Applications*, 150:1 – 18, 2017.

[34] Guillaume Carlier, Ivar Ekeland, and Jean-Charles Rochet. private communication. 2008.

[35] Guillaume Carlier and Maxime Laborde. Remarks on continuity equations with nonlinear diffusion and nonlocal drifts. *Journal of Mathematical Analysis and Applications*, pages –, 2016.

[36] Guillaume Carlier and Filippo Santambrogio. A variational model for urban planning with traffic congestion. *ESAIM Control Optim. Calc. Var.*, 11(4):595–613 (electronic), 2005.

[37] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[38] J. A. Carrillo, Y. Huang, F. S. Patacchini, and G. Wolansky. Numerical study of a particle method for gradient flows. *Preprint*, 2015.

[39] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4):1273–1309, 2010.

[40] J. A. Carrillo and J. S. Moll. Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms. *SIAM J. Sci. Comput.*, 31(6):4305–4329, 2009/10.

[41] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[42] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[43] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fischer-Rao. *arXiv preprint arXiv:1506.06430*, 2015.

[44] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: geometry and Kantorovich formulation. *arXiv preprint arXiv:1508.05216*, 2015.

[45] Rinaldo M. Colombo, Mauro Garavello, and Magali Lécureux-Mercier. A class of nonlocal models for pedestrian traffic. *Math. Models Methods Appl. Sci.*, 22(4):1150023, 34, 2012.

[46] Rinaldo M. Colombo and Magali Lécureux-Mercier. Nonlocal crowd dynamics models for several populations. *Acta Math. Sci. Ser. B Engl. Ed.*, 32(1):177–196, 2012.

[47] Gianluca Crippa and Magali Lécureux-Mercier. Existence and uniqueness of measure solutions for a system of continuity equations with non-local flow. *NoDEA Nonlinear Differential Equations Appl.*, 20(3):523–537, 2013.

[48] Julien Dambrine, Nicolas Meunier, Bertrand Maury, and Aude Roudneff-Chupin. A congestion model for cell migration. *Commun. Pure Appl. Anal.*, 11(1):243–260, 2012.

[49] Sara Daneri and Giuseppe Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.

[50] Djairo Guedes De Figueiredo. *Lectures on the Ekeland variational principle with applications and detours*. Springer Berlin, 1989.

[51] Ennio De Giorgi. New problems on minimizing movementstt. *Boundary Value Problems for PDE and Applications*, pages 81–98, 1993.

[52] Guido De Philippis, Alpár Richárd Mészáros, Filippo Santambrogio, and Bozhidar Velichkov. BV estimates in optimal transportation and applications. *Arch. Ration. Mech. Anal.*, 219(2):829–860, 2016.

[53] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. On the entropic structure of reaction-cross diffusion systems. *Comm. Partial Differential Equations*, 40(9):1705–1747, 2015.

[54] Marco Di Francesco and Simone Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity*, 26(10):2777–2808, 2013.

[55] Marco Di Francesco and Daniel Matthes. Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. Partial Differential Equations*, 50(1-2):199–230, 2014.

[56] Simone Di Marino, Augusto Gerolin, and Luca Nenna. Optimal transportation theory with repulsive costs. *arXiv preprint arXiv:1506.04565*, 2015.

[57] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[58] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

[59] Bertram Düring, Daniel Matthes, and Josipa Pina Milišić. A gradient flow scheme for nonlinear fourth order equations. *Discrete Contin. Dyn. Syst. Ser. B*, 14(3):935–959, 2010.

[60] Jonathan Eckstein and Dimitri P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 55(3, Ser. A):293–318, 1992.

[61] Michel Fortin and Roland Glowinski. *Augmented Lagrangian methods*, volume 15 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.

[62] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element methods. *Computers and Mathematics with applications*, 2:17–40, 1976.

[63] Thomas Gallouët and Leonard Monsaingeon. A jko splitting scheme for kantorovich-fisher-rao gradient flows. working paper or preprint, February 2016.

[64] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[65] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[66] Ansgar Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.

[67] Ansgar Jüngel and Nicola Zamponi. A cross-diffusion system derived from a fokker-planck equation with partial averaging. *arXiv preprint arXiv:1601.05039*, 2016.

[68] Leonid V Kantorovich. On the transfer of masses. In *Dokl. Akad. Nauk. SSSR*, volume 37, pages 227–229, 1942.

[69] David Kinderlehrer, Léonard Monsaingeon, and Xiang Xu. A wasserstein gradient flow approach to poisson-nernst-planck equations. *arXiv preprint arXiv:1501.04437*, 2015.

[70] David Kinderlehrer and Noel J. Walkington. Approximation of parabolic equations using the Wasserstein metric. 33(4):837–852, 1999.

[71] Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A new optimal transport distance on the space of finite radon measures. *arXiv preprint arXiv:1505.07746*, 2015.

[72] Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A fitness-driven cross-diffusion system from population dynamics as a gradient flow. *Journal of Differential Equations*, 261(5):2784–2808, 2016.

[73] M. Laborde. On some non linear evolution systems which are perturbations of Wasserstein gradient flows. *to appear in Radon Ser. Comput. Appl. Math.*, 2015.

[74] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[75] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

[76] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[77] Christian Léonard. From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.*, 262(4):1879–1920, 2012.

[78] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[79] Thomas Lepoutre, Michel Pierre, and Guillaume Rolland. Global well-posedness of a conservative relaxed cross diffusion system. *SIAM J. Math. Anal.*, 44(3):1674–1693, 2012.

[80] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal Entropy-Transport problems and a new Hellinger-Kantorovich distance between positive measures. *arXiv preprint arXiv:1508.07941*, 2015.

[81] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *arXiv preprint arXiv:1509.00068*, 2015.

[82] J.-L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications. Vol. 1.* Travaux et Recherches Mathématiques, No. 17. Dunod, Paris, 1968.

[83] Stefano Lisini and Antonio Marigonda. On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscripta Math.*, 133(1-2):197–224, 2010.

[84] Stefano Lisini, Daniel Matthes, and Giuseppe Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.

[85] Simone Di Marino and Alpar Richard Meszaros. Uniqueness issues for evolution equations with density constraints, 2016.

[86] Daniel Matthes, Robert J. McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.

[87] Daniel Matthes and Horst Osberger. Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.*, 48(3):697–726, 2014.

[88] Daniel Matthes and Horst Osberger. Convergence of a fully discrete variational scheme for a thin-film equation, 2015.

[89] Bertrand Maury, Aude Roudneff-Chupin, and Filippo Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.

[90] Bertrand Maury, Aude Roudneff-Chupin, and Filippo Santambrogio. Congestion-driven dendritic growth. *Discrete Contin. Dyn. Syst.*, 34(4):1575–1604, 2014.

[91] Bertrand Maury, Aude Roudneff-Chupin, Filippo Santambrogio, and Juliette Venel. Handling congestion in crowd motion modeling. *Netw. Heterog. Media*, 6(3):485–519, 2011.

[92] Bertrand Maury and Juliette Venel. *Handling of Contacts in Crowd Motion Simulations*, pages 171–180. Springer Berlin Heidelberg, 2009.

[93] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[94] Alpar Richard Meszaros and Filippo Santambrogio. Advection-diffusion equations with density constraints, 2016.

[95] Gaspard Monge. Mémoire sur la théory des déblais et des remblais. *Histoire de l'Académie Royale des sciences de Paris*, pages 666–704, 1781.

[96] J. D. Murray. *Mathematical biology. II*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2003. Spatial models and biomedical applications.

[97] Felix Otto. Double degenerate diffusion equations as steepest descent, 1996.

[98] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[99] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. Optimal transport with proximal splitting. *SIAM J. Imaging Sci.*, 7(1):212–238, 2014.

[100] Brendan Pass. Multi-marginal optimal transport: theory and applications. *ESAIM Math. Model. Numer. Anal.*, 49(6):1771–1790, 2015.

[101] Benoît Perthame, Fernando Quirós, and Juan Luis Vázquez. The Hele-Shaw asymptotics for mechanical models of tumor growth. *Arch. Ration. Mech. Anal.*, 212(1):93–127, 2014.

[102] Benoît Perthame, Min Tang, and Nicolas Vauchelet. Traveling wave solution of the Hele-Shaw model of tumor growth with nutrient. *Math. Models Methods Appl. Sci.*, 24(13):2601–2626, 2014.

[103] Luca Petrelli and Adrian Tudorascu. Variational principle for general diffusion problems. *Appl. Math. Optim.*, 50(3):229–257, 2004.

[104] Gabriel Peyré. Entropic wasserstein gradient flows, 2015.

[105] Michel Pierre. Global existence in reaction-diffusion systems with control of mass: a survey. *Milan J. Math.*, 78(2):417–455, 2010.

[106] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[107] Riccarda Rossi and Giuseppe Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2, 2003.

[108] Filippo Santambrogio. Variational problems in transport theory with mass concentration, 2006.

[109] Filippo Santambrogio. Transport and concentration problems with interaction effects. *J. Glob. Opt.*, 38(1):129–141, 2007.

[110] Filippo Santambrogio. Gradient flows in Wasserstein spaces and applications to crowd movement. In *Seminaire: Equations aux Dérivées Partielles. 2009–2010*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXVII, 16. École Polytech., Palaiseau, 2012.

[111] Filippo Santambrogio. Models and applications of optimal transport in economics, traffic, and urban planning. In *Optimal transportation*, volume 413 of *London Math. Soc. Lecture Note Ser.*, pages 22–40. Cambridge Univ. Press, Cambridge, 2014.

[112] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians*. Progress in Nonlinear Differential Equations and Their Applications 87. Birkasauser Verlag, Basel, 2015.

[113] E. Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Ann. Inst. H. Poincaré*, 2(4):269–310, 1932.

[114] Jacques Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[115] A. Tudorascu. One-phase Stefan problems; a mass transfer approach. *Adv. Math. Sci. Appl.*, 14(1):151–185, 2004.

[116] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[117] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[118] Jonathan Zinsl. Geodesically convex energies and confinement of solutions for a multicomponent system of nonlocal interaction equations. Technical report, 2014.

[119] Jonathan Zinsl and Daniel Matthes. Transport distances and geodesic convexity for systems of degenerate diffusion equations. *Calc. Var. Partial Differential Equations*, 54(4):3397–3438, 2015.

Résumé

Depuis l'article fondateur de Jordan, Kinderlehrer et Otto en 1998, il est bien connu qu'une large classe d'équations paraboliques peuvent être vues comme des flots de gradient dans l'espace de Wasserstein. Le but de cette thèse est d'étendre cette théorie à certaines équations et systèmes qui n'ont pas exactement une structure de flot de gradient. Les interactions étudiées sont de différentes natures. Le premier chapitre traite des systèmes avec des interactions non locales dans la dérive. Nous étudions ensuite des systèmes de diffusions croisées s'appliquant aux modèles de congestion pour plusieurs populations. Un autre modèle étudié est celui où le couplage se trouve dans le terme de réaction comme les systèmes proie-prédateur avec diffusion ou encore les modèles de croissance tumorale. Nous étudierons enfin des systèmes de type nouveau où l'interaction est donnée par un problème de transport multi-marges. Une grande partie de ces problèmes est illustrée de simulations numériques.

Abstract

Since 1998 and the seminal work of Jordan, Kinderlehrer and Otto, it is well known that a large class of parabolic equations can be seen as gradient flows in the Wasserstein space. This thesis is devoted to extensions of this theory to equations and systems which do not have exactly a gradient flow structure. We study different kind of couplings. First, we treat the case of nonlocal interactions in the drift. Then, we study cross diffusion systems which model congestion for several species. We are also interested in reaction-diffusion systems as diffusive prey-predator systems or tumor growth models. Finally, we introduce a new class of systems where the interaction is given by a multi-marginal transport problem. In many cases, we give numerical simulations to illustrate our theorical results.

Mots Clés

distance de Wasserstein, flots de gradient, schéma JKO, splitting, dérive non locale, diffusions non linéaires, diffusions croisées, systèmes de réaction-diffusion, équations d'Hele-Shaw, transport optimal, transport multi-marges, formule de Benamou-Brenier, lagrangien augmenté, mouvement de foules, espèces en interaction

Keywords

Wasserstein distance, gradient flows, JKO scheme, splitting, nonlocal drift, nonlinear diffusions, cross-diffusion, reaction-diffusion systems, Hele-Shaw equation, optimal transport, multi-marginal transport, Benamou-Brenier formula, augmented lagrangian, crowd motions, interacting species