



**HAL**  
open science

# Mean field games : numerical methods and case of risk-averse agents

Pierre Lavigne

► **To cite this version:**

Pierre Lavigne. Mean field games : numerical methods and case of risk-averse agents. Optimization and Control [math.OC]. Institut Polytechnique de Paris, 2021. English. NNT : 2021IPPAX094 . tel-03556749

**HAL Id: tel-03556749**

**<https://theses.hal.science/tel-03556749v1>**

Submitted on 4 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT  
POLYTECHNIQUE  
DE PARIS

NNT : 2021IPPAX094

Thèse de doctorat



# Mean field games: numerical methods and case of risk-averse agents

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à l'École polytechnique

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)  
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 3 décembre 2021, par

**PIERRE LAVIGNE**

Composition du Jury :

Daniela Tonon Professeure associée, Département de mathématiques "Tullio Levi-Civita", Université de Padoue	Rapporteuse
Yves Achdou Professeur, Laboratoire Jacques-Louis Lions, Université de Paris	Rapporteur
Romuald Elie Professeur, Laboratoire d'analyse et de mathématiques appliquées, Université Paris-Est	Président du jury
Jean-David Benamou Directeur de recherche, Centre de recherche en mathématiques de la décision, Inria	Examineur
Nizar Touzi Professeur, Centre de mathématiques appliquées, École polytechnique	Examineur
Frédéric Bonnans Directeur de recherche, Laboratoire des signaux et systèmes, Inria et CentraleSupélec	Directeur de thèse
Laurent Pfeiffer Chargé de recherche, Laboratoire des signaux et systèmes, Inria et CentraleSupélec	Co-encadrant de thèse



*À mon grand-père, qui  
était si fier de moi.*

# Remerciements

Une thèse est un voyage de plusieurs années au pays des idées et des raisonnements. C'est un périple que l'on effectue entouré de guides, d'autres voyageurs et de proches qui nous accompagnent.

Académiques, amis, famille sont autant de cercles de personnes qui nous nourrissent de leurs connaissances, de leur soutien et de leur confiance. Ces compagnons de route forment les différentes facettes de notre personnalité, orientent, facilitent et rendent agréable cette expédition.

Dans cette aventure, il n'existe pas de séparation claire entre la sphère privée et professionnelle, les deux sont poreuses. La vie de jeune chercheur et le domaine personnel s'entre-déterminent en partie, parfois s'opposent, se complètent. Je crois que cette partie dédiée aux remerciements est l'occasion de prendre du recul, de se replonger dans trois années de souvenirs. Retracer le chemin qui a été emprunté, comprendre ce qui nous constitue en tant qu'individu, ce qui nous anime, et quelque part, dessine les contours de qui nous sommes. C'est pourquoi j'aimerais remercier ici les personnes qui m'ont accompagné durant mon parcours et qui ont eu, toutes à leur manière, un impact sur mon identité et mon travail.

Mes premières pensées vont évidemment à mes directeurs de thèse Frédéric et Laurent. Je remercie Frédéric de m'avoir fait confiance et d'avoir accepté d'être mon directeur de thèse. Alors étudiant en master 2 d'optimisation à Orsay, j'avais un profil atypique, puisque j'ai une formation d'économiste. Il a pris le risque d'encadrer un étudiant qui n'avait pas de formation classique en mathématiques, ce qui n'était pas une évidence. Toujours bienveillant, il a fait en sorte que ma thèse se déroule dans les meilleures conditions possibles. A partir de la deuxième année, Frédéric a proposé à Laurent de devenir co-encadrant de ma thèse. Je le remercie d'avoir assumé cette responsabilité, son arrivée dans l'équipe et son accompagnement ont marqué un tournant, autant pour l'avancement des travaux que d'un point de vue personnel. Etant son seul élève pendant un an, j'ai eu une expérience de thèse privilégiée. J'ai passé de nombreux moments à travailler avec Laurent, à l'observer et à comprendre sa manière de raisonner. Je suis conscient de la chance que j'ai eue d'avoir de telles conditions de travail, j'ai énormément appris à ses côtés. Laurent est pour moi un modèle de rigueur et de persévérance. C'est un passionné de mathématiques à la capacité de travail hors norme. Je ne compte plus le nombre de fois où nos réunions de 14h se sont terminées (précipitamment) à 19h59, heure après laquelle les portes du bâtiment de l'École se ferment. Initialement partis pour une réunion "rapide", les remarques et les idées fleurissent, de fil en aiguille le flou initial s'estompe pour laisser place à l'ordre et à l'enchaînement des arguments. Je suis particulièrement fier d'avoir été son premier étudiant de thèse et je suis conscient des liens particuliers que cela a pu créer.

Je voudrais aussi remercier Frédéric et Laurent de m'avoir encadré sur un sujet qui m'a passionné. L'approche que nous avons eue des jeux à champ moyen m'a permis

d'explorer des mathématiques très diverses, allant du calcul stochastique à l'analyse convexe, aux équations aux dérivées partielles, aux chaînes de Markov, à la théorie du contrôle, aux algorithmes d'optimisation en passant par la théorie des jeux. La pluralité des mathématiques que nous avons étudiées, travaillées, explorées au cours de ces trois années a été, pour moi, l'occasion de mesurer la multiplicité des sujets qu'ils maîtrisent avec un grand niveau de détail. Leur érudition a été une source constante d'inspiration et de challenge.

J'adresse mes remerciements à Yves Achdou et Daniela Tonon pour avoir accepté de rapporter ma thèse. Je leur suis reconnaissant du temps qu'ils ont accordé à la lecture attentive de mon manuscrit. Je remercie également Jean-David Benamou, Romuald Elie et Nizar Touzie d'avoir accepté de faire partie de mon jury de thèse.

Je souhaite remercier l'École Doctorale de Mathématiques Hadamard (EDMH) ainsi que le LabEx de Mathématiques Hadamard (LMH). Parce que mon expérience de thèse s'est ancrée dans un lieu, j'adresse mes remerciements au Centre de Mathématiques Appliquées (CMAP) et à l'École polytechnique. Je voudrais aussi remercier le Laboratoire des Signaux et Systèmes (L2S) de CentraleSupélec et l'Institut national de recherche en informatique et en automatique (Inria) pour m'avoir permis de terminer ma thèse dans les meilleures conditions possibles.

Mon parcours académique s'est construit au hasard de rencontres, d'enseignants et d'élèves qui ont pris le temps de me conseiller et de m'accompagner. J'aimerais ici les remercier : dans l'ordre chronologique, j'exprime toute ma gratitude à Fakhreddine Ghommid, mon enseignant de mathématiques en terminale et en classes préparatoires aux grandes écoles sans qui, rien de tout cela n'aurait été possible. Il a su trouver les mots pour me remotiver, à un moment où j'étais en prise avec le doute et la perte de confiance en moi. Il m'a transmis sa passion des mathématiques, qui ne m'a plus quitté. Merci à Vincent Sitbon, enseignant en économie, qui a, je crois marqué tous ses élèves par sa qualité intellectuelle et sa pédagogie. Il a lui aussi énormément contribué à forger ma personnalité et à guider mon parcours. Je remercie les théoriciens des jeux Jérôme Renault, Benoit Duvocelle et Tristan Tomala de m'avoir transmis leur sujet avec passion ainsi que pour leurs conseils au moment de m'orienter en Master 2 et en thèse. Un grand merci à Filippo Santambrogio pour sa gestion du Master 2 d'optimisation, son écoute attentive et sa disponibilité.

Enfin, j'aimerais avoir quelques mots pour Guillaume Bonnet avec qui j'ai partagé ces trois (presque quatre) années. Guillaume, j'ai adoré échanger mes réflexions avec toi. J'ai une affection particulière pour ta manière de réfléchir et d'aborder les sujets d'un point de vue complètement neutre, tant mathématiques que sociétaux. Je crois que j'ai dû te rebattre les oreilles avec mes histoires de pétrole, de climat et d'écologie. Et pourtant, à chaque fois on reprenait les choses point par point, pour vérifier la structure et la véracité de nos raisonnements. Ces discussions structurées étaient vraiment passionnantes et source d'une grande joie intellectuelle. J'ai néanmoins un regret, celui de ne pas avoir travaillé avec toi. J'espère que nous aurons l'occasion d'y remédier, en nous penchant sur HJB ou autre !

Depuis le début de ma thèse, j'essaie de cultiver un profil assez particulier de mathématicien traitant des thématiques écologiques. Je crois que ces dernières peuvent être appréhendées de manière théorique mais qu'une expertise de terrain est aussi nécessaire, ne serait-ce que pour comprendre correctement les enjeux et être légitime sur ces sujets. C'est pour cette raison que je me suis impliqué dans divers projets en parallèle de ma thèse, qui m'ont nourri et fait grandir.

Quand je suis arrivé à l'École polytechnique, j'étais en grande réflexion sur les problématiques écologiques. Assez naturellement, je me suis tourné vers le binet Développement Durable (DDX) afin de discuter de ces thématiques. Je voudrais remercier les élèves qui y participent et le font vivre, je sais à quel point leur engagement est véritable et parfois difficile. J'aimerais particulièrement remercier Benoit Halgand, Victor François, Cédric Le Mouel, Luc Bedouelle et Maryam Bouhied pour leur investissement et leur accueil.

Puis, j'ai souhaité passer de la réflexion à l'action. C'est à ce moment que j'ai rencontré Lionel Montoliu, avec qui nous avons fondé le projet 0.6planet qui se donne pour objectif de prototyper et essayer un mode de vie compatible avec une empreinte écologique soutenable. Nous avons partagé une année intense pour donner naissance à ce projet, remplie de merveilleuses rencontres ! Merci à Lionel Montoliu, infatigable monteur de projet, à sa femme Monica Bassett ; merci à nos ingénieurs low-tech Xavier De Vesvrotte et à Jean-Baptiste Penant pour leur énorme investissement ! Merci à Nelly Bonnefous, qui doit avoir tous les numéros de téléphone utiles de la terre dans son répertoire. J'ai appris tellement de choses à vos côtés, que ce soit sur les dynamiques de groupe, sur la manière de construire un projet en partant de zéro, sur l'écologie ou encore sur moi-même. Vous êtes les architectes de demain.

À travers cette expérience, j'ai eu la chance d'encadrer Bastien Evra sur le sujet "Habitat et sobriété, une mesure de l'empreinte écologique du ménage français" dans le cadre de son projet d'accompagnement de troisième année d'école. Merci à lui et à Riwal Plougonven pour la confiance qu'ils m'ont accordée.

J'ai également encadré la mission de prospective d'Inès Abdessalem et Anastasia Millevoye : "Développement de maisons containers en France à l'horizon 2030" dans le cadre de leur Master 2 Management de la Technologie et de l'Innovation. Un immense merci à elles et à Lionel pour leur confiance.

Puisque les problématiques écologiques appellent à un changement de pratiques, d'organisation et plus généralement de paradigme dans toutes les strates de la société, la recherche n'échappe pas à cette règle. Certains chercheurs se posent la question de comprendre les enjeux écologiques contemporains et de savoir comment leurs activités devraient être modifiées pour (au moins) ne pas aggraver la situation. C'est pour cette raison qu'en juin 2020, Mathilde Boissier a décidé de créer la Commission de Développement Durable du laboratoire. J'aimerais la remercier chaleureusement pour son initiative et les précieux conseils qu'elle a pu me donner par ailleurs. Merci à ceux qui font vivre cette commission : Lucas Chesnel, Solange Pruilh, Giovanni Conforti et Célia Escribe qui nous a rejoint récemment. On aura mis le temps, mais on l'aura sorti ce bilan carbone !

J'aimerais remercier Alexis Tantet pour les discussions passionnantes que nous avons eues autour des sujets économie-énergie-climat à la suite de son cours. Ce sont des thématiques complexes, délicates, difficiles, qui touchent à la rationalité, à l'émotionnel et aux représentations du monde. C'est sur les conseils d'Alexis que j'ai contacté Peter Tankov avec qui j'ai monté mon projet postdoctoral. Merci à toi Peter d'avoir accepté d'encadrer ce projet, mais aussi pour ton investissement concernant les questions écologiques. Je pense que nous avons de beaux sujets devant nous !

Enfin, j'ai eu l'énorme opportunité de pouvoir, avec une équipe incroyable de doctorants du CMAP, organiser le premier congrès français de mathématiques appliquées pour les jeunes chercheuses et jeunes chercheurs (CJC-MA). Nous avons vécu une année intense d'organisation et riche d'enseignements. Merci à Matthieu Aussal d'avoir proposé ce projet ; à la Société de Mathématiques Appliquées et Industrielles (SMAI), le CMAP

et l'École polytechnique pour leurs soutiens. Je crois pouvoir dire que l'évènement a permis, à travers la convivialité, l'émulsion intellectuelle nécessaire au dynamisme de la communauté des jeunes mathématiciennes et mathématiciens. Merci à Constantin Philippenko et Thomas Bellotti de l'avoir présidé et vice-présidé ; à Guillaume Bonnet d'avoir été notre Webmaster ; à Claire Ecotière et Dominik Stantejsky d'avoir été respectivement trésorière et vice-trésorier ; à Louis Reboul et Clément Mantoux de vous être chargés de la communication et des partenariats ; à Apolline Louvet et Solange Pruihl d'avoir porté ce congrès sur vos épaules !

Avec mon binôme Baptiste Kerleguer, nous avons animé le comité scientifique de l'évènement. Merci Baptiste pour ton investissement, cela a été un réel plaisir de faire vivre ce comité avec toi. Merci à Amandine Véber, Matthieu Aussal, Aude Sportisse, Juliette Chevalier, Fédor Goncharov d'avoir accepté de faire partie de ce comité à nos côtés et de vous être prêtés au jeu. Merci à tous pour cette année hors du commun !

J'aimerais remercier mes proches, qui m'ont toujours soutenu en toutes circonstances et qui m'ont apporté l'équilibre nécessaire à la réalisation de ce travail.

Si je suis devenu la personne (et le jeune chercheur) que je suis, c'est grâce à vous deux : merci papa, merci maman. Vous n'avez jamais douté une seule seconde de mes capacités, et vous m'avez toujours soutenu dans mes choix. Je sais qu'avec Valentine nous avons été l'objet de toutes vos priorités depuis que nous sommes nés. Merci à toi Valentine pour ton soutien, je sais que ça n'a pas toujours été facile d'être ma soeur. Malgré tout, tu as su sortir de mes traces pour prendre un très beau chemin.

Malgré mon éloignement, je sais que je pourrai toujours compter sur vous. C'est tellement agréable de savoir que, quand je rentre à Dijon, j'ai une magnifique bande d'amis toujours prête à m'accueillir. Je n'arrive pas à être aussi présent que je le voudrais, mais merci Romain, Pierre(al), Julien<sup>2</sup>, Florian, Loris, Bryan et les autres pour votre amitié et votre compréhension.

*Last but not least*, merci aux membres de 3RJ, ma famille d'adoption depuis un an ! Vous m'avez rendu heureux. Heureux de rentrer tous les soirs pour vous retrouver, heureux de jouer de la musique avec vous, heureux de partager nos repas, nos joies, nos peines, nos doutes et nos espérances. Merci Adeline, Danièle, Joan, Laura<sup>2</sup> et Julien pour ces moments inoubliables.



# Contents

<b>1</b>	<b>Introduction (version française)</b>	<b>1</b>
1.1	Généralités sur les jeux à champ moyen . . . . .	2
1.1.1	Équilibres de Nash dans les jeux anonymes . . . . .	2
1.1.2	Problèmes de contrôle individuel . . . . .	3
1.1.3	Jeux à champ moyen et couplages . . . . .	6
1.1.4	Jeux à champ moyen potentiels . . . . .	9
1.1.5	Méthodes numériques . . . . .	11
<b>2</b>	<b>Introduction</b>	<b>13</b>
2.1	Generalities on mean field games . . . . .	14
2.1.1	Nash equilibria in anonymous games . . . . .	14
2.1.2	Individual control problems . . . . .	15
2.1.3	Mean field games and couplings . . . . .	18
2.1.4	Potential mean field games . . . . .	20
2.1.5	Numerical methods . . . . .	23
2.2	Contributions . . . . .	24
2.2.1	Chapter 1: Discrete-time mean field games with risk-averse agents . . . . .	24
2.2.2	Chapter 2: Discrete potential mean field games . . . . .	29
2.2.3	Chapter 3: Generalized conditional gradient and learning in potential mean field games . . . . .	35
<b>3</b>	<b>Discrete-time mean field games with risk averse-agents</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Problem Formulation . . . . .	42
3.2.1	Notations . . . . .	42
3.2.2	Coupled system . . . . .	43
3.2.3	Assumptions . . . . .	45
3.3	Interpretation of the coupled system . . . . .	45
3.3.1	Dynamic programming equation . . . . .	46
3.3.2	Kolmogorov equation . . . . .	49
3.4	Technical lemmas . . . . .	49
3.5	Existence result . . . . .	53
3.5.1	Dynamic Programming mapping . . . . .	53
3.5.2	Kolmogorov mapping . . . . .	56
3.5.3	Existence of equilibrium . . . . .	56
3.6	Connection with a finite player game . . . . .	57
3.6.1	Formulation of the game . . . . .	57
3.6.2	An approximate Nash equilibrium . . . . .	59
3.6.3	Proof of Theorem 3.6.4 . . . . .	60

3.7	Conclusion	64
<b>4</b>	<b>Discrete potential mean field games</b>	<b>67</b>
4.1	Introduction	67
4.2	Discrete mean field games	69
4.2.1	Notation	69
4.2.2	Coupled system	70
4.2.3	Further notation	72
4.3	Potential problem and convex formulation	72
4.3.1	Perspective functions	72
4.3.2	Potential problem	73
4.4	Duality	74
4.4.1	Duality result	74
4.4.2	A new dual problem	76
4.5	Connection between the MFG system and potential problems	77
4.6	Numerical methods	80
4.6.1	Notations	80
4.6.2	Primal-dual proximal algorithms	80
4.6.3	ADMM and ADM-G	83
4.6.4	Residuals	86
4.7	Numerical Results	86
4.7.1	Example 1	87
4.7.2	Example 2	90
4.8	Appendix	93
4.8.1	Projection on $Q$	94
4.8.2	Entropic proximity operator	94
<b>5</b>	<b>Generalized conditional gradient and learning in potential mean field games</b>	<b>97</b>
5.1	Introduction	97
5.2	Data and main assumptions	100
5.2.1	Notations	100
5.2.2	Coupled system and assumptions	101
5.2.3	Potentials	102
5.3	Estimates for the individual control problem	103
5.3.1	The individual problem as a stochastic optimal control problem	103
5.3.2	The individual problem as a PDE optimal control problem	105
5.4	Properties of the solution to the mean field game system	107
5.5	Generalized conditional gradient	109
5.6	Convergence Results	113
5.7	Appendix	118
5.7.1	Regularity of the Hamiltonian	118
5.7.2	A priori bounds for parabolic equations	119
5.7.3	Maximum principle	120
5.7.4	Existence and uniqueness of a classical solution to the Hamilton-Jacobi-Bellman equation	121

# Chapter 1

## Introduction (version française)

Les jeux à champ moyen (MFG) sont à la fois une théorie mathématique et un outil de modélisation. Ils ont été introduits en 2006 indépendamment par Jean-Michel Lasry et Pierre-Louis Lions dans [88, 89, 90] et Minyi Huang, Roland P. Malhamé, et Peter E. Caines dans [80]. Les jeux à champ moyen sont des modèles asymptotiques de jeux non coopératifs de joueurs rationnels, interagissant par un effet de champ moyen. Chaque acteur optimise son propre système dynamique, par rapport à une certaine fonctionnelle de coût. Cette dernière dépend de variables de couplage, elles-mêmes dépendantes de la distribution des états et des contrôles.

Cette thèse consiste en trois études consacrées à l'analyse et à la résolution numérique de différents modèles de jeux à champ moyen. Nous commençons par une introduction aux jeux à champ moyen dans la section 2.1. Dans la section 2.2 nous proposons une présentation détaillée des trois contributions :

**2.2.1 Jeux à champ moyen en temps discret avec agents averses au risque :** en utilisant le concept de mesure de risque composite, nous étudions un modèle MFG en temps discret impliquant des agents averses au risque . Nous montrons l'existence d'une solution via une approche point fixe. Nous montrons qu'une politique optimale du MFG est  $\varepsilon(N)$ -optimale pour un certain jeu à  $N$  joueurs et que la suite  $\varepsilon(N)$  converge vers zéro lorsque le nombre de joueurs tend vers l'infini.

**2.2.2 Jeux à champ moyen potentiels discrets :** nous étudions des MFG à potentiel en temps discret et en espace d'état fini (aussi appelé variationnel) avec des contraintes dures, c'est-à-dire avec des potentiels convexes, éventuellement non différentiables et avec un domaine borné. Nous définissons un problème primal et un problème dual, et nous montrons un résultat de dualité sous des conditions de qualification appropriées. Nous montrons l'existence d'une solution au système de jeu à champ moyen et montrons l'unicité lorsque les potentiels sont dérivables. Ensuite, nous implémentons deux familles de méthodes numériques : des méthodes proximales primales-duales (Chambolle-Pock et Chambolle-Pock-Bregman) et des méthodes basées sur le Lagrangien augmenté (ADMM et ADM-G). Nous comparons les performances de chaque méthode pour deux cas d'application : un modèle de congestion et un modèle de prix, tous deux avec des contraintes dures.

**2.2.3 Gradient conditionnel généralisé et apprentissage dans les jeux à champ moyen potentiels :** nous appliquons l'algorithme de gradient conditionnel généralisé pour les jeux à champ moyen potentiels dans un cadre EDP. Nous mettons en évidence le lien entre cet algorithme et une méthode itérative de meilleure réponse

pour résoudre des jeux appelée algorithme de jeu fictif. Cette dernière est une méthode d'apprentissage qui se déroule de la façon suivante : à chaque étape,

- (a) Optimisation : pour une croyance donnée de la stratégie des autres, trouver la meilleure réponse des joueurs ;
- (b) Apprentissage : mettre à jour la croyance en faisant la moyenne de toutes les meilleures réponses trouvées depuis le début de la procédure.

Dans ce cadre potentiel, la notion d'exploitabilité issue de la théorie des jeux peut être interprétée comme un écart primal-dual. On montre que pour la séquence d'apprentissage  $\delta_k = 2/(k+2)$ , le coût potentiel converge en  $O(1/k)$ , l'exploitabilité et les variables du problème (distribution, congestion, prix, fonction de valeur et termes de contrôle) convergent en  $O(1/\sqrt{k})$ , pour des normes appropriées.

## 1.1 Généralités sur les jeux à champ moyen

Nous introduisons dans cette section les principaux concepts de théorie des jeux à champ moyen qui seront approfondis dans la thèse. Par commodité, nous avons décidé d'utiliser le cadre du chapitre 5, qui est un modèle en temps continu avec un espace d'état continu. Il a la forme la plus standard (en comparaison avec la littérature actuelle), soulignons que les cadres des chapitres 3 et 4 sont différents.

Dans la section 1.1.1 nous rappelons d'abord la notion d'équilibre de Nash dans les jeux à  $N$ -joueurs et nous expliquons comment passer à la limite lorsque le nombre de joueurs tend vers l'infini. Dans la section 1.1.2, nous présentons un *sec:individual-control-problem* paramétré par des termes de couplage fixes  $\gamma$  et  $P$ , et dérivons un contrôle de rétroaction optimal. Dans la section 1.1.3, nous définissons un problème de jeu à champ moyen où chaque agent optimise le *sec:individual-control-problem* introduit dans la section 1.1.2 avec des termes de couplage dépendant maintenant de la distribution des états et des contrôles des joueurs. Nous détaillons la structure de point fixe du jeu et mettons en évidence quelques différences entre les modèles de congestion et de prix. Dans la section 1.1.4, nous présentons des jeux à champ moyen potentiels, qui sont des jeux à champ moyen avec une structure variationnelle. Cette classe est particulièrement intéressante pour montrer l'unicité de solution et l'application de méthodes numériques dont nous donnons un aperçu dans la section 1.1.5.

### 1.1.1 Équilibres de Nash dans les jeux anonymes

Les jeux à champ moyen sont des modèles limites pour des jeux avec un grand nombre de joueurs. Le but de cette première section est d'expliquer comment dériver un modèle limite pour un jeu à  $N$  joueurs et comment établir une connexion entre les deux ; plus précisément, nous expliquons comment construire une solution approchée du jeu à  $N$  joueurs à partir du modèle limite. Pour cela, nous considérons un modèle très simple.

**Jeu à  $N$  joueurs** Soit  $N > 0$  le nombre de joueurs et soit  $\mathcal{N} = \{1, \dots, N\}$  l'ensemble des joueurs. Soit  $\mathcal{C}$  un sous-ensemble borné de  $\mathbb{R}^d$  et soit  $\mathcal{P}(\mathcal{C})$  l'ensemble des mesures de probabilité sur  $\mathcal{C}$ . Soit  $c: \mathcal{C} \times \mathcal{P}(\mathcal{C}) \rightarrow \mathbb{R}$  une fonction objectif. On suppose que  $c$  est  $K$ -Lipschitz par rapport à sa deuxième variable, pour la distance de Rubinstein-Kantorovich, notée  $d_1$ . Chaque agent cherche à minimiser  $c$  en choisissant une stratégie  $x_i \in \mathcal{C}$  pour une mesure donnée  $m_{\bar{x}}^{N,i} := \frac{1}{N-1} \sum_{j \in \mathcal{N} \setminus \{i\}} \delta_{\bar{x}_j} \in \mathcal{P}(\mathcal{C})$  des autres. On dit

que  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_N) \in \mathcal{C}^N$  est un équilibre de Nash si pour tout  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, m_{\bar{x}}^{N,i}) \leq c(x, m_{\bar{x}}^{N,i}) \quad (1.1)$$

pour tout  $x \in \mathcal{C}$ . Un équilibre de Nash est un point tel qu'il n'y a pas de déviations unilatérales profitables. De façon équivalente, nous avons que

$$c(\bar{x}_i, m_{\bar{x}}^{N,i}) = \inf_{x \in \mathcal{C}} c(x, m_{\bar{x}}^{N,i}) = \inf_{m \in \mathcal{P}(\mathcal{C})} \int_{\mathcal{C}} c(x, m_{\bar{x}}^{N,i}) dm(x).$$

**Limite de champ moyen** Lorsque le nombre de joueurs est infini, on dit que  $\bar{m} \in \mathcal{P}(\mathcal{C})$  est un équilibre de Nash si et seulement si

$$\int_{\mathcal{C}} c(x, \bar{m}) d\bar{m}(x) = \inf_{x \in \mathcal{C}} c(x, \bar{m}) = \inf_{m \in \mathcal{P}(\mathcal{C})} \int_{\mathcal{C}} c(x, \bar{m}) dm(x). \quad (1.2)$$

De manière équivalente,  $\bar{m}$  est un équilibre de Nash si  $\text{supp}(\bar{m}) \subseteq \text{argmin}_{x \in \mathcal{C}} c(x, \bar{m})$ . Considérons une solution du modèle limite, c'est-à-dire un point  $\bar{m}$  satisfaisant (1.2). L'idée principale pour construire une solution approximative à (1.1) consiste à trouver  $\bar{x} = (\bar{x}_i)_{i \in \mathcal{N}} \in \mathcal{C}^N$  tel que pour tout  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, \bar{m}) \leq c(x, \bar{m}), \quad \forall x \in \mathcal{C} \quad (1.3)$$

et tel que

$$d_1(\bar{m}, m_{\bar{x}}^N) \leq \varepsilon(N), \quad \text{où} : m_{\bar{x}}^N = \frac{1}{N} \sum_{i \in \mathcal{N}} \delta_{\bar{x}_i} \in \mathcal{P}(\mathcal{C}).$$

En pratique  $\bar{x}$  peut être obtenu via des inégalités de concentration [58], qui assurent que  $\varepsilon(N) \rightarrow 0$  lorsque  $N \rightarrow \infty$ . Nous avons

$$m_{\bar{x}}^N = \frac{N-1}{N} m_{\bar{x}}^{N,i} + \frac{1}{N} \delta_{\bar{x}_i},$$

donc  $d_1(m_{\bar{x}}^N, m_{\bar{x}}^{N,i}) \leq C/N$  et donc,

$$d_1(\bar{m}, m_{\bar{x}}^{N,i}) \leq C/N + \varepsilon(N).$$

On déduit de la continuité Lipschitz de  $c$ , de l'inégalité (1.3) et l'inégalité ci-dessus que pour tout  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, \bar{m}) \leq c(x, \bar{m}) + 2K(C/N + \varepsilon(N)), \quad \forall x \in \mathcal{C},$$

ce qui prouve que  $\bar{x}$  est un équilibre de Nash approximatif pour le jeu à  $N$  joueurs.

### 1.1.2 Problèmes de contrôle individuel

Nous définissons maintenant le problème de contrôle individuel résolu par tout agent impliqué dans le problème de jeu à champ moyen. Dans cette introduction son problème est un problème de contrôle optimal stochastique en temps continu paramétré par une paire de congestion et de prix  $(\gamma, P)$ .

**Problème de contrôle stochastique** Soit  $(B_s)_{s \in [0, T]}$  un mouvement brownien et soit  $Y$  une variable aléatoire, indépendante de  $(B_s)_{s \in [0, T]}$ , avec une distribution de probabilité  $m_0 \in \mathcal{D}_1(\mathbb{T}^d) := \{m \in L^\infty(\mathbb{T}^d), m \geq 0, \int_{\mathbb{T}^d} m(x) dx = 1\}$ . Soit  $\mathbb{F}$  la filtration engendrée par le mouvement brownien  $(B_s)_{s \in [0, T]}$  et la variable aléatoire initiale  $Y$ . On note  $L_{\mathbb{F}}^2(\mathbb{R}^d)$  l'ensemble des processus stochastiques  $\nu$  sur  $[0, T]$  de valeur dans  $\mathbb{R}^d$  adapté à  $\mathbb{F}$  tel que  $\mathbb{E} \left[ \int_0^T |\nu_s|^2 ds \right] < +\infty$ . Pour tout terme de congestion Lipschitz  $\gamma$  et vecteur continu de prix  $P$ , on définit le coût individuel  $Z_{\gamma, P}: L_{\mathbb{F}}^2(\mathbb{R}^d) \rightarrow \mathbb{R}$ ,

$$Z_{\gamma, P}(\nu) = \mathbb{E} \left[ \int_0^T L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s^\nu, s) ds + g(X_T^\nu) \right],$$

où pour tout  $\nu \in L_{\mathbb{F}}^2(\mathbb{R}^d)$ , on note  $(X_s^\nu)_{s \in [0, T]}$  la solution de l'équation différentielle stochastique

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_0 = Y.$$

Le critère stochastique individuel est composé d'un coût courant individuel fortement convexe  $L: \mathbb{R}^d \rightarrow \mathbb{R}$ , un coût terminal lisse  $g \in \mathcal{C}^\infty(\mathbb{T}^d)$ , et  $m_0 \in \mathcal{C}^\infty(\mathbb{T}^d)$  est la distribution de probabilité de l'état initial  $Y$ . Le critère est également paramétré par la congestion  $\gamma$  et le vecteur des prix  $P$ . Le problème individuel est :

$$\inf_{\nu \in L_{\mathbb{F}}^2(\mathbb{R}^d)} Z_{\gamma, P}(\nu). \quad (1.4)$$

**Programmation dynamique** La théorie de la programmation dynamique nous permet de caractériser la solution du problème individuel présenté ci-dessus. L'idée est d'introduire la fonction valeur  $u[\gamma, P]: \mathbb{T}^d \times [0, T] \rightarrow \mathbb{R}$ , définie par

$$u[\gamma, P](x, t) = \inf_{\nu \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)} \mathbb{E} \left[ \int_t^T L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s, s) ds + g(X_T) \right],$$

paramétrée par  $\gamma$  et  $P$  et où  $X$  est la solution de

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_t = x. \quad (1.5)$$

Sous des hypothèses appropriées, la fonction  $u[\gamma, P]$  est la solution de viscosité à l'équation de Hamilton-Jacobi-Bellman

$$\begin{aligned} -\partial_t u - \Delta u + H(\nabla u + P) &= \gamma, & (x, t) \in Q, \\ u(x, T) &= g(x), & x \in \mathbb{T}^d, \end{aligned}$$

où  $Q$  est défini par  $Q = \mathbb{T}^d \times [0, T]$  et l'application  $H$ , appelée hamiltonienne, est liée à la transformée de Fenchel du coût courant  $L$ . C'est-à-dire, pour tout  $p \in \mathbb{R}^d$ ,

$$H(p) = L^*(-p) = \sup_{v \in \mathbb{R}^d} -\langle p, v \rangle - L(v).$$

Si  $L$  est strictement convexe (et donc au moins continu), le supremum est atteint à  $v = -H_p(p)$ , où  $H_p$  est le gradient de l'hamiltonien. Ensuite, à partir de la solution  $u[\gamma, P]$ , on peut dériver une fonction de rétroaction

$$v(x, t) = -H_p(\nabla u(x, t) + P(t)),$$

pour tout  $(x, t) \in Q$ .

**Vérification** La fonction de rétroaction  $v$  nous permet de construire une solution au problème de contrôle, via le système en boucle fermée

$$dX_s^v = v(X_s^v, s)ds + \sqrt{2}dB_s, \quad X_0 = Y. \quad (1.6)$$

Nous montrons ensuite que le processus stochastique  $\nu^v$ , défini par

$$\nu_s^v := v(X_s, s) = -H_p(\nabla u(X_s, s) + P(s)),$$

pour tout  $s \in [0, T]$ , est en effet optimal. Soit  $\nu \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$  et soit  $(X_s^\nu)_{s \in [0, T]}$  la solution de (1.5). Par la formule d'Itô nous avons

$$\begin{aligned} u(X_T^\nu, T) - u(X_0^\nu, 0) &= \int_0^T (\partial_t u(X_s^\nu, s) + \nabla u(X_s^\nu, s) \cdot \nu_s + \Delta u(X_s^\nu, s)) ds \\ &\quad + \int_0^T \nabla u(X_s^\nu, s) \cdot dB_s. \end{aligned}$$

En prenant l'espérance des deux côtés, on obtient

$$\mathbb{E}[u(X_0^\nu, 0) - u(X_T^\nu, T)] = (a) + (b) + (c),$$

où

$$\begin{aligned} (a) &= \mathbb{E} \left[ \int_0^T (L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s^\nu, s)) ds \right], \\ (b) &= -\mathbb{E} \left[ \int_0^T (\partial_t u(X_s^\nu, s) + \Delta u(X_s^\nu, s) - H(\nabla u(X_s^\nu, s) + P(s)) + \gamma(X_s^\nu, s)) ds \right], \\ (c) &= -\mathbb{E} \left[ \int_0^T (L(\nu_s) + H(\nabla u(X_s^\nu, s) + P(s)) + \langle P(s) + \nabla u(X_s^\nu, s), \nu_s \rangle) ds \right], \end{aligned}$$

pour tout  $\nu \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$ . Puisque  $u$  est solution de l'équation de Hamilton-Jacobi-Bellman, nous avons  $(b) = 0$ . De plus, par définition,  $u(x, T) = g(x)$  pour tout  $x \in \mathbb{T}^d$ . Ainsi

$$\mathbb{E}[u(Y, 0)] = Z_{\gamma, P}(\nu) + (c),$$

où nous avons utilisé la condition initiale  $X_0^\nu = Y$ . Par l'inégalité de Fenchel-Young, on a

$$L(\nu) + H(\nabla u(x, s) + P(s)) + \langle \nabla u(x, s) + P(s), \nu \rangle \geq 0,$$

pour tout  $(x, s) \in Q$  et  $\nu \in \mathbb{R}^d$ . L'inégalité est une égalité lorsque  $\nu = -H_p(\nabla u(x, s) + P(s))$ . En conséquence  $(c) \leq 0$  pour tout  $\nu \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$  et  $(c) = 0$  quand  $\nu = \nu^v$ . Nous avons donc

$$Z_{\gamma, P}(\nu^v) = \mathbb{E}[u(Y, 0)] \leq Z_{\gamma, P}(\nu),$$

ce qui prouve que le contrôle  $(\nu_s^v)_{s \in [0, T]}$  est optimal.

**Problème de contrôle d'équations aux dérivées partielles** Les espaces de Sobolev sont notés  $W^{n, q}(Q)$ , l'ordre de dérivation  $n$  étant éventuellement non entier (suivant la définition de [87, section II.2]). Nous fixons

$$W^{2, 1, q}(Q) = W^{1, q}(Q) \cap L^q(0, T; W^{2, q}(\mathbb{T}^d)), \quad W^{1, 0, q}(Q) = L^q(0, T; W^{1, q}(\mathbb{T}^d)).$$

Soit  $p > d + 2$ . Le problème de contrôle stochastique individuel a une formulation d'équation aux dérivées partielles donnée par

$$\inf_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,P}(m,v) := \int_Q (L(v) + \langle P, v \rangle + \gamma) m dx dt + \int_{\mathbb{T}^d} g m(T) dx,$$

où l'ensemble  $\mathcal{R}$  est défini comme l'ensemble des couples  $(m, v) \in W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$  satisfaisant l'équation de Fokker-Planck

$$\partial_t m - \Delta m + \nabla \cdot (vm) = 0, \quad m(0) = m_0.$$

Le lien entre les deux problèmes s'explique par le fait qu'étant donné une fonction de rétroaction  $v$ , la solution  $m$  de l'équation de Fokker-Planck nous fournit la distribution de probabilité de la solution du système en boucle fermée (1.6). Dans le modèle limite à infini d'agents qui sera introduit ensuite, l'équation de Fokker-Planck décrira l'évolution de la distribution des agents, en supposant qu'ils emploient tous la même rétroaction  $v$  et qu'ils sont soumis à des bruits indépendants.

### 1.1.3 Jeux à champ moyen et couplages

Jusqu'à présent, nous avons décrit un problème de contrôle individuel où les deux paramètres  $\gamma$  et  $P$  étaient fixés. Le but des jeux à champ moyen est de considérer un continuum de joueurs identiques, chacun optimisant son propre critère, tout en interagissant avec les autres à travers les termes de couplage.

**Un problème de point fixe** Dans la section 1.1.2 nous avons expliqué comment caractériser le contrôle optimal d'un agent représentatif grâce à la théorie de la programmation dynamique. Nous introduisons maintenant un problème de jeu à champ moyen, sous la forme d'un problème de point fixe, où les termes de couplage  $\gamma$  et  $P$  ne sont plus donnés, mais font partie du problème. Ils sont déduits de  $m$  et  $v$  via des fonctions de couplage. Soit  $f: \mathbb{T}^d \times \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}$  et  $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^d$  des applications Lipschitz et bornées. Le problème de jeu à champ moyen est donné par : trouver  $(\bar{m}, \bar{v}) \in \mathcal{R}$  tel que

$$(\bar{m}, \bar{v}) \in \arg \min_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,P}(m,v), \quad \gamma(x,t) = f(x, \bar{m}(t)), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x,t) \bar{m}(x,t) dx \right), \quad (1.7)$$

pour tout  $t \in [0, T]$ . Ce problème est un problème d'équilibre de Nash (voir le lien avec (1.2) dans la section 1.1.1). Il bénéficie d'une structure de point fixe : chaque  $(m, v)$  induit les termes de couplage  $\gamma$  et  $P$  via les applications de couplage  $f$  et  $\phi$  ; et toute paire de termes de couplage  $(\gamma, P)$  donne une paire  $(m, v)$  via l'optimisation du problème individuel. Avant de décrire le système de jeu à champ moyen associé, nous discutons des modèles de congestion et de prix.

**Modèles de congestion** Dans les modèles de congestion, l'interaction entre les acteurs se fait à travers le terme de congestion  $f$ . Cela correspond au cas où  $\phi = 0$  dans la section ci-dessus. Une situation naturelle à considérer est celle du mouvement d'une foule où chaque joueur cherche à éviter des régions encombrées, pénalisées via le couplage  $f$ . Le problème de jeu à champ moyen est donné par

$$(\bar{m}, \bar{v}) \in \arg \min_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,0}(m,v), \quad \gamma(x,t) = f(x, \bar{m}(t)),$$



pour tout  $t \in [0, T]$ . Dans cette situation l'état  $x \in \mathbb{T}^d$  d'un joueur représente par exemple sa localisation dans une pièce, une rue, etc. Le contrôle  $v$  est la vitesse du déplacement.

**Modèles de prix** Dans un modèle de prix, l'interaction se produit via le terme de prix  $P$ . Cela correspond au cas  $f(x, m) = 0$  pour tout  $x \in \mathbb{T}^d$ ,  $m \in \mathcal{D}_1(\mathbb{T}^d)$  dans la section ci-dessus. Une situation classique est donnée par le modèle de concurrence de Cournot : chaque acteur minimise son coût en vendant ou en achetant sur le marché. Le problème stochastique individuel est donné par

$$(\bar{m}, \bar{v}) \in \arg \min_{(m, v) \in \mathcal{R}} \mathcal{Z}_{0, P}(m, v), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x, t) \bar{m}(x, t) dx \right),$$

pour tout  $t \in [0, T]$ . Le vecteur des prix  $P(t)$  est fonction de la demande agrégée  $D(t) := \int_{\mathbb{T}^d} v(x, t) m(x, t) dx$  à l'instant  $t \in [0, T]$ . Ensuite, dans ce problème, un agent représentatif contrôle un niveau de stock  $x \in \mathbb{T}^d$  composé de  $d$  actifs différents. Dans la définition de la fonction de coût  $\mathcal{Z}_{0, P}$ , nous pouvons interpréter  $P(s)$  comme le prix (au temps  $s$ ) associé à  $\nu_s$ .

**Système de jeu de champ moyen** Nous sommes maintenant prêts à présenter le système de jeu à champ moyen où  $(u, v, m, \gamma, P)$  est l'inconnu associé au problème de jeu à champ moyen  $(\text{MFG}_{f, \phi})$ ,

$$\left\{ \begin{array}{ll} \text{(i)} & \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, & (x, t) \in Q, \\ u(x, T) = g(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(ii)} & v = -H_p(\nabla u + P), \quad (x, t) \in Q, \\ \text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, & (x, t) \in Q, \\ m(0, x) = m_0(x), & x \in \mathbb{T}^d, \end{cases} \quad (\text{MFG}_{f, \phi}) \\ \text{(iv)} & \gamma(x, t) = f(x, m(t)), \quad (x, t) \in Q, \\ \text{(v)} & P(t) = \phi \left( \int_{\mathbb{T}^d} v(x, t) m(x, t) \right), \quad t \in [0, T], \end{array} \right.$$

où  $u(x, t), m(x, t), \gamma(x, t) \in \mathbb{R}$ ,  $v(x, t) \in \mathbb{R}^d$ , et  $P(t) \in \mathbb{R}^d$ . Le système ci-dessus est un système couplé d'équations aux dérivées partielles :

- L'équation (i) est l'équation de Hamilton-Jacobi-Bellman, décrivant l'évolution de la fonction valeur du problème individuel.
- L'équation (ii) définit le contrôle optimal  $v$ .
- L'équation (iii) est l'équation de Fokker-Planck, décrivant l'évolution de la distribution d'état des agents. Le terme  $m_0$  est la distribution initiale des agents. Sauf mention contraire, les solutions s'entendent au sens des distributions.
- L'équation (iv) définit la congestion et l'équation (v) le prix.

Le lien entre (1.7) et  $(\text{MFG}_{f, \phi})$  est le suivant : toute solution  $(u, v, m, \gamma, P)$  au système de jeu à champ moyen  $(\text{MFG}_{f, \phi})$  est telle que  $(m, v)$  est la solution du problème

d'équilibre de Nash ; pour toute solution  $(m, v)$  solution au problème d'équilibre de Nash, définissant

$$\gamma(x, t) = f(x, \bar{m}(t)), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x, t) \bar{m}(x, t) dx \right), \quad u = u[\gamma, P],$$

le quintuplet  $(u, v, m, \gamma, P)$  est une solution du système de jeu à champ moyen  $(\text{MFG}_{f,\phi})$ .

**Structure en point fixe du système de jeu à champ moyen** Le système de jeu à champ moyen  $(\text{MFG}_{f,\phi})$  peut également être considéré comme un problème de point fixe. Ce point de vue est clé pour prouver l'existence d'une solution, grâce au théorème de point fixe de Schauder par exemple. Nous présentons ici les dépendances entre chacune des variables du système couplé et nous mettons en évidence les différences entre les modèles de congestion et de prix. Le graphique suivant illustre cette structure pour deux cas : les jeux à champ moyen classiques, c'est-à-dire les situations où  $\phi = 0$  ; et les jeux à champ moyen avec interactions par les contrôles qui correspondent au cas général considéré ici.

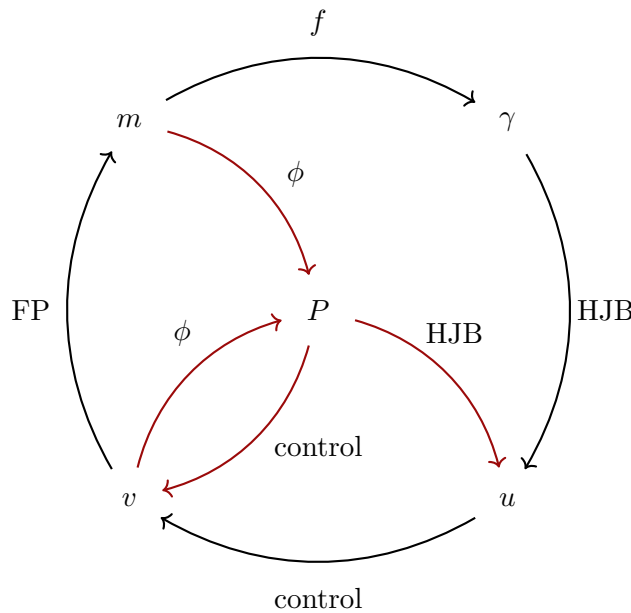


Figure 1.1: Unknowns dependency graph: fixed point structure of the mean field game problem.

Les dépendances inconnues, dans le contexte des jeux à champ moyen classiques, sont représentées par les flèches noires. La structure est assez simple : étant donné une congestion  $\gamma$ , on peut calculer la fonction de valeur  $u$  solution de l'équation de Hamilton-Jacobi-Bellman (HJB). Alors le gradient de  $u$  fournit le contrôle optimal  $v$ . Étant donné un contrôle optimal  $v$ , on peut calculer la solution  $m$  de l'équation de Fokker-Planck (FP). Enfin la congestion  $\gamma$  dépend de  $m$  via le mapping  $f$ .

Les jeux à champ moyen avec interactions par les contrôles ont une structure plus complexe : il existe des dépendances supplémentaires, représentées par les flèches rouges, dues à la variable prix. Étant donné un contrôle  $v$  et une distribution  $m$ , on peut calculer le prix  $P$  via l'application  $\phi$ . Ensuite, le prix  $P$  donne une valeur  $u$  via l'équation de Hamilton-Jacobi-Bellman et un contrôle associé  $v$ . Une preuve d'existence avec une représentation en point fixe peut être trouvée dans [21].

### 1.1.4 Jeux à champ moyen potentiels

Les jeux à champ moyen potentiels (également appelés variationnels dans la littérature) sont des jeux à champ moyen qui peuvent être interprétés comme les conditions du premier ordre de deux problèmes de contrôle appelés problèmes potentiels. Le premier problème est un problème de contrôle de l'équation de Fokker-Planck et le second un problème de contrôle optimal de l'équation de Hamilton-Jacobi-Bellman. Il y a plusieurs avantages à travailler avec des jeux à champ moyen potentiels. Ils offrent une manière différente (comparé aux méthodes de point fixe) de prouver l'existence d'une solution au système de jeu à champ moyen, qui peut être déduite des solutions des problèmes potentiels. Précisons que selon les modèles, cette approche peut être délicate à suivre (en particulier dans le cas de diffusion dégénérée, voir [34]). De plus, la forme potentielle permet d'utiliser des algorithmes d'optimisation pour résoudre le système de jeu à champ moyen, comme nous le verrons dans la section suivante.

Dans notre cadre, il suffit de supposer que  $f$  et  $\phi$  dérivent de potentiels pour obtenir une structure variationnelle du jeu. On dit que  $f$  et  $\phi$  dérivent de potentiels s'il existe deux applications convexes  $F: \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}$  et  $\Phi: \mathbb{R}^k \rightarrow \mathbb{R}$  tel que

$$\begin{aligned} F(m_2) - F(m_1) &= \int_0^1 \int_{\mathbb{T}^d} f(x, sm_2 + (1-s)m_1)(m_2(x) - m_1(x)) dx ds, \\ \phi(z) &= \nabla_z \Phi(z), \end{aligned}$$

pour tout  $m_1, m_2 \in \mathcal{D}_1(\mathbb{T}^d)$ , et tout  $z \in \mathbb{R}^k$ . On définit le critère suivant

$$\mathcal{J}(m, v) = \int_Q L(v)m dx dt + \int_0^T \left( F(m(t)) + \Phi \left( \int_{\mathbb{T}^d} v(x, t)m(x, t) dx \right) \right) dt + \int_{\mathbb{T}^d} gm(T) dx,$$

et le problème de contrôle suivant :

$$\inf_{(m, v) \in \mathcal{R}} \mathcal{J}(m, v). \quad (\text{P})$$

En raison des termes  $L(v)m$  et  $\Phi(\int vm)$ , le critère du problème ci-dessus n'est pas convexe. L'idée est d'étudier un problème convexe étroitement lié. On définit le critère

$$\tilde{\mathcal{J}}(m, w) = \int_Q \tilde{L}(m, w) dx dt + \int_0^T \left( F(m(t)) + \Phi \left( \int_{\mathbb{T}^d} v(x, t)m(x, t) dx \right) \right) dt + \int_{\mathbb{T}^d} gm(T) dx,$$

où  $\tilde{L}: Q \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  est la fonction perspective de  $L$ ,

$$\tilde{L}(m, w) = \begin{cases} mL(w/m), & \text{si } m > 0, \\ 0, & \text{si } m = 0 \text{ et } w = 0, \\ +\infty, & \text{sinon.} \end{cases}$$

Il est à noter que  $\tilde{L}$  est convexe et semi-continue inférieurement par rapport à  $(m, w)$ . On définit l'ensemble

$$\tilde{\mathcal{R}} = \{(m, w) \in W^{2,1,p}(Q) \times W^{1,0,\infty}(Q), \partial_t m - \Delta m + \nabla \cdot w = 0, m(0) = m_0\}.$$

Le critère  $\tilde{\mathcal{J}}$  et l'ensemble  $\tilde{\mathcal{R}}$  sont tous les deux convexes de sorte que le problème suivant est convexe,

$$\inf_{(m, w) \in \tilde{\mathcal{R}}} \mathcal{J}(m, v). \quad (\tilde{\text{P}})$$

A condition que les deux problèmes aient des solutions  $(\bar{m}, \bar{v})$  et  $(\bar{m}, \bar{w})$  et que  $\bar{m}(x, t) > 0$  pour tout  $(x, t) \in Q$ , les solutions des deux problèmes sont telles que  $\bar{v} = \bar{w}/\bar{m}$ . On définit le critère dual

$$\tilde{D}(u, \gamma, P) = \int_{\mathbb{T}^d} u(x, 0) m_0(x) dx - \int_0^T (F^*(\gamma(t)) + \Phi^*(P(t))) dt, \quad (1.8)$$

et on considère le problème dual

$$\sup_{(u, \gamma, P) \in \tilde{\mathcal{U}}} \tilde{D}(u, \gamma, P), \quad (\tilde{D})$$

où  $\tilde{\mathcal{U}}$  est l'ensemble de  $(u, \gamma, P) \in W^{2,1,p}(Q) \times L^\infty(Q) \times L^\infty(0, T; \mathbb{R}^k)$  satisfaisant

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^*P] &\leq \gamma, & (x, t) \in Q, \\ u(x, T) &\leq g(x), & x \in \mathbb{T}^d. \end{aligned}$$

**Dualité et lien avec le système de jeu à champ moyen** Ici, nous exposons et expliquons les liens existants entre les problèmes potentiels ci-dessus et le système de jeu à champ moyen. La présentation suivante n'est pas une preuve, mais plutôt un résumé des liens et résultats que l'on peut espérer être vrai sous des hypothèses appropriées.

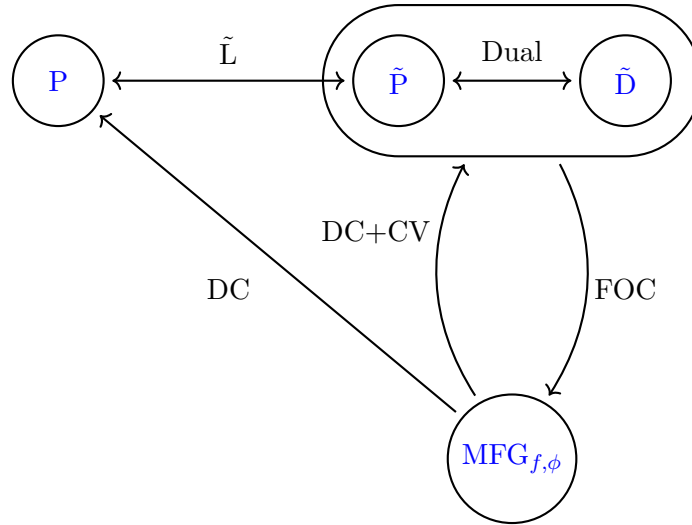


Figure 1.2: Résumé des liens entre  $(P)$ ,  $(\tilde{P})$ ,  $(\tilde{D})$ , et  $(MFG_{f,\phi})$ .

Soit  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$  une solution de  $(MFG_{f,\phi})$ , par souci de simplicité, nous supposons que  $\bar{m}(x, t) > 0$  pour tout  $(x, t) \in Q$ . On définit le changement de variable “à la Benamou-Brenier”  $\bar{w} = \bar{m}\bar{v}$  et on a  $\bar{v} = \bar{w}/\bar{m}$  par positivité uniforme de  $m$ . Nous présentons d’abord comment obtenir une solution à  $(\tilde{P})$ ,  $(\tilde{D})$  et  $(P)$ .

**DC:** Par un calcul direct (essentiellement par intégration par parties), toute solution de  $(MFG_{f,\phi})$  donne une solution de  $(P)$ .

**DC+CV:** Par un calcul direct (encore par intégration par parties) et un changement de variable  $\bar{w} = \bar{m}\bar{v}$ , toute solution de  $(MFG_{f,\phi})$  donne une solution  $(\bar{m}, \bar{w})$  à  $(\tilde{P})$  et une solution  $(\bar{u}, \bar{\gamma}, \bar{P})$  à  $(\tilde{D})$ .

Nous expliquons maintenant le lien entre  $(\mathbf{P})$  et  $(\tilde{\mathbf{P}})$ . Ce lien peut être utile car il est parfois plus facile de montrer l'unicité des solutions à  $(\mathbf{P})$  puis de déduire l'unicité des solutions à  $(\tilde{\mathbf{P}})$ .

**$\tilde{\mathbf{L}}$ :** Le lien entre  $(\mathbf{P})$  et  $(\tilde{\mathbf{P}})$  est donné par la fonction perspective et par le changement de variable “à la Benamou-Brenier”. Pour toute solution  $(\bar{m}, \bar{w})$  du problème  $(\tilde{\mathbf{P}})$ , on a que  $(\bar{m}, \bar{v} = \bar{w}/\bar{m})$  est la solution de  $(\mathbf{P})$ . Toute solution  $(\bar{m}, \bar{v})$  à  $(\mathbf{P})$  est telle que  $(\bar{m}, \bar{w} = \bar{m}\bar{v})$  est la solution au problème  $(\tilde{\mathbf{P}})$ . De plus, les valeurs des deux problèmes sont égales.

Enfin, nous expliquons le lien entre les problèmes primaux  $(\tilde{\mathbf{P}})$  et  $(\tilde{\mathbf{D}})$ . Nous expliquons également pourquoi les solutions à ces problèmes sont des solutions au système de jeu à champ moyen  $(\text{MFG}_{f,\phi})$ .

**Dual:** Les problèmes  $(\tilde{\mathbf{P}})$  et  $(\tilde{\mathbf{D}})$  sont formellement en dualité. Pour établir un tel résultat, une approche classique consiste à utiliser le théorème de Fenchel-Rockafellar [109]. Selon le cadre d'analyse, le choix des espaces de fonctions peut être une question délicate (voir [34] pour le cas dégénéré du second ordre). Dans le cadre discret, ce résultat est au cœur des méthodes numériques primales-duales.

**FOC:** Le système de jeu à champ moyen peut être interprété comme des conditions de premier ordre pour les problèmes potentiels  $(\tilde{\mathbf{P}})$  et  $(\tilde{\mathbf{D}})$ . On peut vérifier que pour toutes les solutions  $(\bar{m}, \bar{w})$  à  $(\tilde{\mathbf{P}})$  et pour toutes les solutions  $(\bar{u}, \bar{\gamma}, \bar{P})$ , le quintuplet  $(\bar{m}, \bar{v} = \bar{w}/\bar{m}, \bar{u}, \bar{\gamma}, \bar{P})$  est une solution à  $(\text{MFG}_{f,\phi})$ . Au chapitre 4, dans un cadre discret, nous utilisons l'inégalité de Fenchel-Young pour établir un tel résultat.

### 1.1.5 Méthodes numériques

La résolution numérique des jeux à champ moyen a reçu une attention particulière au cours des dernières années. Nous renvoyons le lecteur à [6, 91] pour une revue des méthodes numériques utilisés pour résoudre les problèmes de jeux à champ moyen. Nous proposons ici une présentation synthétique et une classification simple. Nous distinguons deux approches : une approche basée sur le système MFG et une approche Potentiel. La première approche consiste à résoudre directement le système de jeu à champ moyen via des méthodes d'apprentissage ou de Newton. La seconde approche consiste à minimiser un potentiel dont les solutions satisfont le système de jeu à champ moyen, via des méthodes primales-duales ou de points-selles.

Les méthodes d'apprentissage adoptent le point de vue d'un acteur représentatif. Les algorithmes d'apprentissage se déroulent comme suit : à chaque étape, pour une croyance donnée sur la stratégie des autres, un agent représentatif calcule sa meilleure réponse ; puis il apprend en faisant la moyenne de toutes les meilleures réponses trouvées depuis le début de la procédure d'apprentissage. La question d'apprentissage est centrale dans la théorie des jeux [60], et l'algorithme de jeu fictif a une interprétation en sciences sociales ou en ingénierie : les équilibres de Nash peuvent être atteints si tous les agents suivent leurs intérêts personnels, jouent au jeu suffisamment de fois et de manière répétée mettre à jour leurs croyances.

Le point de départ des méthodes de Newton est de considérer le système de jeu à champ moyen comme un problème de la forme  $f(x) = 0$  pour une fonction différentiable  $f$  des inconnues  $x$  (typiquement la fonction valeur, la distribution des agents, les termes de couplage). Pour résoudre un tel problème, un algorithme de Newton peut être utilisé par exemple comme dans [5]. Il est à noter que dans cette dernière référence, en raison de la dépendance entre chaque variable du problème, les auteurs sont en mesure de

Approche:	Classe:	Méthodes:
Système MFG	Apprentissage	Jeu fictif [35, 55, 77, 103] Descente miroire [76, 101] Apprentissage par renforcement [9, 50, 75, 117, 119]
	Newton	Newton algorithm [3, 4, 5, 96] Algorithme de Gauss-Newton [30, 31]
Potentiel	Primale	Gradient conditionnel (Frank-Wolfe) [62] Gradient conditionnel généralisé [24] Sinkhorn [16] Descente de gradient stochastique [42, 43] Algorithme d'Uzawa's [18]
	Point-selle	Lagrangienne [15, 17, 22, 97, 112] Méthodes primales-duales proximales [22, 27, 28]

Figure 1.3: Classification et résumé des méthodes numériques pour les jeux à champ moyen.

reformuler l'algorithme de Newton comme une méthode à point fixe pour les variables de couplage.

En toute généralité ces deux premières classes ne nécessitent aucune forme potentielle du jeu, même si en pratique, une forme potentielle où la monotonie des termes de couplage peut être supposée afin d'assurer l'unicité des solutions et la convergence de la méthode. Il est à noter que dans [5], les auteurs ne supposent pas la monotonie des termes de couplage.

Au contraire, l'approche potentielle est basée sur la minimisation des problèmes potentiels. Dans cette approche, nous pouvons distinguer deux classes de méthodes :

1. Méthodes primales (resp. dual), basées sur la minimisation (resp. maximisation) du problème primal (resp. dual). Dans cette catégorie, nous trouvons l'algorithme de gradient conditionnel, l'algorithme de gradient conditionnel généralisé, l'algorithme de Sinkhorn et les algorithmes de descente de gradient stochastique.
2. Les méthodes de point-selle sont des méthodes qui exploitent explicitement la structure primale-duale des problèmes potentiels. Dans cette catégorie, nous avons les méthodes lagrangiennes (ALG2, ADMM, ADM-G), les méthodes proximales primales-duales et l'algorithme d'Uzawa.

Parmi toutes ces références, seules [5, 22, 24, 96] sont liées aux jeux à champ moyen avec interactions par les contrôles.

# Chapter 2

## Introduction

Mean field games (abbreviated MFGs) are both a mathematical theory and a modeling tool. They were introduced in 2006 independently by Jean-Michel Lasry and Pierre-Louis Lions in [88, 89, 90] and Minyi Huang, Roland P. Malhamé, and Peter E. Caines in [80]. Mean field games are asymptotic models of non-cooperative games of rational players, interacting through a mean field effect. Each player optimizes its own dynamical system, with respect to some cost functional. The latter depends on coupling variables, which themselves depend on the state and control distribution.

This thesis consists in three studies devoted to the analysis and numerical resolution of different mean field game models. We start with an introduction to mean field games in Section 2.1. In Section 2.2 we provide a detailed presentation of the three contributions:

**2.2.1 Discrete-time mean field games with risk averse-agents:** using the concept of composite risk measure, we study a discrete-time MFG model involving risk-averse agents. We show the existence of a solution via a fixed point approach. We show that an optimal policy of the MFG is  $\varepsilon(N)$ -optimal for a certain  $N$ -player game. The sequence  $\varepsilon(N)$  converges to zero as the number of players tends to infinity.

**2.2.2 Discrete potential mean field games:** we study discrete time and finite state space potential (also called variational) MFGs with hard constraints, that is with convex potentials, possibly non-differentiable and with bounded domain. We define a primal and a dual problem, and we show a duality result under suitable qualification conditions. We show the existence of a solution to the mean field game system and show the uniqueness when the potentials are differentiable. Then we implement two families of numerical methods: primal-dual proximal methods (Chambolle-Pock and Chambolle-Pock-Bregman) and augmented Lagrangian based methods (ADMM and ADM-G). We compare the performance of each method for two application cases: a congestion model and a price model, both with hard constraints.

**2.2.3 Generalized conditional gradient and learning in potential mean field games:** we apply the generalized conditional gradient algorithm for potential mean field games in a PDE framework. We highlight the connection between this algorithm and a best response iterative method for solving games called fictitious play algorithm. The latter is a learning method which goes as follows: at each step,

- (a) optimize: for a given belief of the strategy of the others, find the best response of the players;

- (b) learn: update the belief by averaging all the best responses found from the beginning of the procedure.

In this potential framework the notion of exploitability from game theory can be interpreted as a primal-dual gap. We show that for the learning sequence  $\delta_k = 2/(k+2)$ , the potential cost converges in  $O(1/k)$ , the exploitability and the variables of the problem (distribution, congestion, price, value function and control terms) converge in  $O(1/\sqrt{k})$ , for specific norms.

## 2.1 Generalities on mean field games

We introduce in this section the main concepts of mean field game theory which will be further investigated in the thesis. For convenience, we have decided to use the framework of Chapter 5, which is a continuous time model with continuous state space. It has the most standard form (in comparison with the current literature), let us emphasize that the frameworks of Chapters 3 and 4 are different.

In Section 2.1.1 we first recall the notion of Nash equilibrium in  $N$ -player games and we explain how to pass to the limit when the number of players tends to infinity. In Section ?? we present an individual control problem parameterized by fixed coupling terms  $\gamma$  and  $P$ , and derive an optimal feedback control. In Section 2.1.3, we define a mean field game problem where each agent optimizes the individual control problem introduced in Section 2.1.2 with coupling terms now depending on the distribution of states and controls of the players. We detail the fixed point structure of the game and exhibit some differences between congestion and price models. In Section 2.1.4 we present potential mean field games, which are mean field games with a variational structure. This class is of particular interest to show uniqueness of solutions and design numerical methods. In Section 2.1.5 we give an overview of the numerical methods for mean field games investigated the literature.

### 2.1.1 Nash equilibria in anonymous games

Mean field games are limit models for games with a large number of players. The intention of this first section is to explain how to derive a limit model for an  $N$ -player game and how to establish a connection between the two; more precisely, we explain how to construct an approximate solution to the  $N$ -player game from the limit model. To this purpose, we consider a very simple model.

**N-player game** Let  $N > 0$  be the number of players and let  $\mathcal{N} = \{1, \dots, N\}$  be the set of players. Let  $\mathcal{C}$  be a bounded subset of  $\mathbb{R}^d$  and let  $\mathcal{P}(\mathcal{C})$  denote the set of probability measures on  $\mathcal{C}$ . Let  $c: \mathcal{C} \times \mathcal{P}(\mathcal{C}) \rightarrow \mathbb{R}$  be an objective function. We assume that  $c$  is  $K$ -Lipschitz with respect to its second variable, for the Rubinstein-Kantorovich distance, denoted  $d_1$ . Each agent aims at minimizing  $c$  by choosing a strategy  $x_i \in \mathcal{C}$  for a given measure  $m_{\bar{x}}^{N,i} := \frac{1}{N-1} \sum_{j \in \mathcal{N} \setminus \{i\}} \delta_{\bar{x}_j} \in \mathcal{P}(\mathcal{C})$  of the others. We say that  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_N) \in \mathcal{C}^N$  is a Nash equilibrium if for any  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, m_{\bar{x}}^{N,i}) \leq c(x, m_{\bar{x}}^{N,i}) \quad (2.1)$$

for any  $x \in \mathcal{C}$ . A Nash equilibrium is a point such that there are no profitable unilateral deviations. Equivalently we have that

$$c(\bar{x}_i, m_{\bar{x}}^{N,i}) = \inf_{x \in \mathcal{C}} c(x, m_{\bar{x}}^{N,i}) = \inf_{m \in \mathcal{P}(\mathcal{C})} \int_{\mathcal{C}} c(x, m_{\bar{x}}^{N,i}) dm(x).$$



**Mean field limit** When the number of players is infinite, we say that  $\bar{m} \in \mathcal{P}(\mathcal{C})$  is a Nash equilibrium if and only if

$$\int_{\mathcal{C}} c(x, \bar{m}) d\bar{m}(x) = \inf_{x \in \mathcal{C}} c(x, \bar{m}) = \inf_{m \in \mathcal{P}(\mathcal{C})} \int_{\mathcal{C}} c(x, \bar{m}) dm(x). \quad (2.2)$$

Equivalently,  $\bar{m}$  is a Nash equilibrium if  $\text{supp}(\bar{m}) \subseteq \text{argmin}_{x \in \mathcal{C}} c(x, \bar{m})$ . Let us consider a solution to the limit model, i.e. let  $\bar{m}$  satisfy (2.2). The main idea to construct an approximate solution to (2.1) consists in finding  $\bar{x} = (\bar{x}_i)_{i \in \mathcal{N}} \in \mathcal{C}^{\mathcal{N}}$  such that for all  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, \bar{m}) \leq c(x, \bar{m}), \quad \forall x \in \mathcal{C} \quad (2.3)$$

and such that

$$d_1(\bar{m}, m_{\bar{x}}^N) \leq \varepsilon(N), \quad \text{where: } m_{\bar{x}}^N = \frac{1}{N} \sum_{i \in \mathcal{N}} \delta_{\bar{x}_i} \in \mathcal{P}(\mathcal{C}).$$

In practice  $\bar{x}$  can be obtained via concentration inequalities [58], which ensure that  $\varepsilon(N) \rightarrow 0$  as  $N \rightarrow \infty$ . We have

$$m_{\bar{x}}^N = \frac{N-1}{N} m_{\bar{x}}^{N,i} + \frac{1}{N} \delta_{\bar{x}_i},$$

thus  $d_1(m_{\bar{x}}^N, m_{\bar{x}}^{N,i}) \leq C/N$  and therefore,

$$d_1(\bar{m}, m_{\bar{x}}^{N,i}) \leq C/N + \varepsilon(N).$$

We deduce from the Lipschitz continuity of  $c$ , inequality (2.3), and the above inequality that for any  $i \in \mathcal{N}$ ,

$$c(\bar{x}_i, \bar{m}) \leq c(x, \bar{m}) + 2K(C/N + \varepsilon(N)), \quad \forall x \in \mathcal{C},$$

which proves that  $\bar{x}$  is an approximate Nash equilibrium for the  $N$ -player game.

### 2.1.2 Individual control problems

We now define the individual control problem solved by any agent involved in the mean field game. In this introduction his/her problem is a continuous time stochastic optimal control problem parametrized by a pair of congestion and price  $(\gamma, P)$ .

**Stochastic control problem** Let  $(B_s)_{s \in [0, T]}$  denote a Brownian motion and let  $Y$  be a random variable, independent of  $(B_s)_{s \in [0, T]}$ , with probability distribution  $m_0 \in \mathcal{D}_1(\mathbb{T}^d) := \{m \in L^\infty(\mathbb{T}^d), m \geq 0, \int_{\mathbb{T}^d} m(x) dx = 1\}$ . Let  $\mathbb{F}$  denote the filtration generated by the Brownian motion  $(B_s)_{s \in [0, T]}$  and the initial random variable  $Y$ . We denote by  $L_{\mathbb{F}}^2(\mathbb{R}^d)$  the set of stochastic processes  $\nu$  on  $[0, T]$  with value in  $\mathbb{R}^d$  adapted to  $\mathbb{F}$  such that  $\mathbb{E} \left[ \int_0^T |\nu_s|^2 ds \right] < +\infty$ . For any Lipschitz congestion term  $\gamma$  and continuous vector of prices  $P$ , we define the individual cost  $Z_{\gamma, P}: L_{\mathbb{F}}^2(\mathbb{R}^d) \rightarrow \mathbb{R}$ ,

$$Z_{\gamma, P}(\nu) = \mathbb{E} \left[ \int_0^T L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s^\nu, s) ds + g(X_T^\nu) \right],$$

where for all  $\nu \in L_{\mathbb{F}}^2(\mathbb{R}^d)$ , we denote by  $(X_s^\nu)_{s \in [0, T]}$  the solution to the stochastic differential equation

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_0 = Y.$$

The individual stochastic criterion is composed of an individual strongly convex running cost  $L: \mathbb{R}^d \rightarrow \mathbb{R}$ , a smooth terminal cost  $g \in \mathcal{C}^\infty(\mathbb{T}^d)$ , and  $m_0 \in \mathcal{C}^\infty(\mathbb{T}^d)$  is the probability distribution of the initial state  $Y$ . The criterion is also parameterized by the congestion  $\gamma$  and the vector of prices  $P$ . The individual problem is:

$$\inf_{\nu \in L_{\mathbb{F}}^2(\mathbb{R}^d)} Z_{\gamma, P}(\nu). \quad (2.4)$$

**Dynamic programming** The dynamic programming theory allows us to characterize the solution of the individual problem introduced above. The idea is to introduce the value function  $u[\gamma, P]: \mathbb{T}^d \times [0, T] \rightarrow \mathbb{R}$ , defined by

$$u[\gamma, P](x, t) = \inf_{\nu \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)} \mathbb{E} \left[ \int_t^T L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s, s) ds + g(X_T) \right],$$

parametrized by  $\gamma$  and  $P$  and where  $X$  is the solution to

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_t = x. \quad (2.5)$$

It is well-known that, under suitable assumptions, the function  $u[\gamma, P]$  is the viscosity solution to the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\partial_t u - \Delta u + H(\nabla u + P) &= \gamma, & (x, t) \in Q, \\ u(x, T) &= g(x), & x \in \mathbb{T}^d, \end{aligned}$$

where  $Q$  is defined by  $Q = \mathbb{T}^d \times [0, T]$  and the map  $H$ , called Hamiltonian, is related to the Fenchel transform of the running cost  $L$ . That is to say, for any  $p \in \mathbb{R}^d$ ,

$$H(p) = L^*(-p) = \sup_{v \in \mathbb{R}^d} -\langle p, v \rangle - L(v).$$

If  $L$  is strictly convex (and thus at least continuous), the supremum is reached at  $v = -H_p(p)$ , where  $H_p$  is the gradient of the Hamiltonian. Then from the solution  $u[\gamma, P]$ , one can derive a feedback function

$$v(x, t) = -H_p(\nabla u(x, t) + P(t)),$$

for all  $(x, t) \in Q$ .

**Verification** The feedback function  $v$  enables us to construct a solution to the control problem, via the closed-loop system

$$dX_s^v = v(X_s^v, s) ds + \sqrt{2} dB_s, \quad X_0 = Y. \quad (2.6)$$

We next show that the stochastic process  $\nu^v$ , defined by

$$\nu_s^v := v(X_s, s) = -H_p(\nabla u(X_s, s) + P(s)),$$

for all  $s \in [0, T]$ , is indeed optimal. Let  $\nu \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$  and let  $(X_s^\nu)_{s \in [0, T]}$  be the solution to (2.5). By the Itô formula we have

$$\begin{aligned} u(X_T^\nu, T) - u(X_0^\nu, 0) &= \int_0^T (\partial_t u(X_s^\nu, s) + \nabla u(X_s^\nu, s) \cdot \nu_s + \Delta u(X_s^\nu, s)) ds \\ &\quad + \int_0^T \nabla u(X_s^\nu, s) \cdot dB_s. \end{aligned}$$

Taking the expectation of both sides, we obtain

$$\mathbb{E}[u(X_0^\nu, 0) - u(X_T^\nu, T)] = (a) + (b) + (c),$$

where

$$\begin{aligned} (a) &= \mathbb{E} \left[ \int_0^T (L(\nu_s) + \langle P(s), \nu_s \rangle + \gamma(X_s^\nu, s)) ds \right], \\ (b) &= -\mathbb{E} \left[ \int_0^T (\partial_t u(X_s^\nu, s) + \Delta u(X_s^\nu, s) - H(\nabla u(X_s^\nu, s) + P(s)) + \gamma(X_s^\nu, s)) ds \right], \\ (c) &= -\mathbb{E} \left[ \int_0^T (L(\nu_s) + H(\nabla u(X_s^\nu, s) + P(s)) + \langle P(s) + \nabla u(X_s^\nu, s), \nu_s \rangle) ds \right], \end{aligned}$$

for any  $\nu \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^d)$ . Since  $u$  is solution to the Hamilton-Jacobi-Bellman equation, we have  $(b) = 0$ . Moreover, by definition,  $u(x, T) = g(x)$  for any  $x \in \mathbb{T}^d$ . Thus

$$\mathbb{E}[u(Y, 0)] = Z_{\gamma, P}(\nu) + (c),$$

where we have used the initial condition  $X_0^\nu = Y$ . By the Fenchel-Young inequality, we have

$$L(\nu) + H(\nabla u(x, s) + P(s)) + \langle \nabla u(x, s) + P(s), \nu \rangle \geq 0,$$

for any  $(x, s) \in Q$  and  $\nu \in \mathbb{R}^d$ . The equality holds in the previous inequality when  $\nu = -H_p(\nabla u(x, s) + P(s))$ . As a consequence  $(c) \leq 0$  for any  $\nu \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^d)$  and  $(c) = 0$  when  $\nu = \nu^v$ . This yields

$$Z_{\gamma, P}(\nu^v) = \mathbb{E}[u(Y, 0)] \leq Z_{\gamma, P}(\nu),$$

which proves that the control  $(\nu_s^v)_{s \in [0, T]}$  is optimal.

**Partial differential equation control problem** Sobolev spaces are denoted by  $W^{n, q}(Q)$ , the order of derivation  $n$  being possibly non-integral (following the definition in [87, section II.2]). We set

$$W^{2, 1, q}(Q) = W^{1, q}(Q) \cap L^q(0, T; W^{2, q}(\mathbb{T}^d)), \quad W^{1, 0, q}(Q) = L^q(0, T; W^{1, q}(\mathbb{T}^d)).$$

Let  $p > d + 2$ . The individual stochastic control problem has a partial differential equation formulation given by

$$\inf_{(m, v) \in \mathcal{R}} \mathcal{Z}_{\gamma, P}(m, v) := \int_Q (L(v) + \langle P, v \rangle + \gamma) m dx dt + \int_{\mathbb{T}^d} gm(T) dx,$$

where the set  $\mathcal{R}$  is defined as the set of pairs  $(m, v) \in W^{2, 1, p}(Q) \times W^{1, 0, \infty}(Q)$  satisfying the Fokker-Planck equation

$$\partial_t m - \Delta m + \nabla \cdot (vm) = 0, \quad m(0) = m_0.$$

The connection between the two problems is explained by the fact that given a feedback function  $v$ , the solution  $m$  to the Fokker-Planck equation provides us with the probability distribution of the solution to the closed-loop system (2.6). In the limit model with infinitely many agents which will be introduced next, the Fokker-Planck equation will describe the evolution of the distribution of the agents, assuming that they all employ the same feedback  $v$  and that they are submitted to independent noises.

### 2.1.3 Mean field games and couplings

Until now we have described an individual control problem where the two parameters  $\gamma$  and  $P$  were fixed. The purpose of mean field games is to consider a continuum of identical players, each one optimizing its own criterion, while interacting with the others through the coupling terms.

**A fixed point problem** In Section 2.1.2 we have explained how to characterize the optimal control of a representative agent thanks to dynamic programming theory. We introduce now a mean field game problem, in the form of a fixed point problem, where the coupling terms  $\gamma$  and  $P$  are not given anymore, but are part of the problem. They are deduced from  $m$  and  $v$  via coupling functions. Let  $f: \mathbb{T}^d \times \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}$  and  $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^d$  be Lipschitz and bounded mappings. The mean field game problem is given by: find  $(\bar{m}, \bar{v}) \in \mathcal{R}$  such that

$$(\bar{m}, \bar{v}) \in \arg \min_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,P}(m, v), \quad \gamma(x, t) = f(x, \bar{m}(t)), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x, t) \bar{m}(x, t) dx \right), \quad (2.7)$$

for all  $t \in [0, T]$ . This problem is a Nash equilibrium problem (see the link with (2.2) in Section 2.1.1). It enjoys a fixed point structure: any  $(m, v)$  induces coupling terms  $\gamma$  and  $P$  via the coupling mappings  $f$  and  $\phi$ ; and any pair of coupling terms  $(\gamma, P)$  yields a pair  $(m, v)$  via the optimization of the individual problem. Before describing the associated mean field game system let us discuss congestion and price models.

**Congestion models** In congestion models, the interaction between the players occurs through the congestion term  $f$ . This corresponds to the case where  $\phi = 0$  in the above section. A natural situation to consider is the motion of a crowd where each player aims to avoid congested regions penalized via the coupling  $f$ . The mean field game problem is given by

$$(\bar{m}, \bar{v}) \in \arg \min_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,0}(m, v), \quad \gamma(x, t) = f(x, \bar{m}(t)),$$

for all  $t \in [0, T]$ . In this situation the state  $x \in \mathbb{T}^d$  of a player represents for example its location in a room, a street, etc. The control  $v$  is the velocity of the displacement.

**Price models** In a price model, the interaction occurs through the price term  $P$ . This corresponds to the case  $f(x, m) = 0$  for any  $x \in \mathbb{T}^d$ ,  $m \in \mathcal{D}_1(\mathbb{T}^d)$  in the above section. A classical situation is given by the Cournot competition model: each player minimizes its cost while selling or buying on the market. The individual stochastic problem is given by

$$(\bar{m}, \bar{v}) \in \arg \min_{(m,v) \in \mathcal{R}} \mathcal{Z}_{0,P}(m, v), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x, t) \bar{m}(x, t) dx \right),$$

for all  $t \in [0, T]$ . The vector of prices  $P(t)$  is a function of the aggregated demand  $D(t) := \int_{\mathbb{T}^d} v(x, t) m(x, t) dx$  at time  $t \in [0, T]$ . Then in this problem, a representative agent controls a stock level  $x \in \mathbb{T}^d$  composed of  $d$  different assets. In the definition of the cost function  $\mathcal{Z}_{0,P}$ , we can interpret  $P(s)$  as the price (at time  $s$ ) associated with  $\nu_s$ .

**Mean field game system** We are now ready to present the mean field game system with unknown  $(u, v, m, \gamma, P)$ , associated with the mean field game problem  $(\text{MFG}_{f,\phi})$ ,

$$\left\{ \begin{array}{ll} \text{(i)} & \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, & (x, t) \in Q, \\ u(x, T) = g(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(ii)} & v = -H_p(\nabla u + P), \quad (x, t) \in Q, \\ \text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, & (x, t) \in Q, \\ m(0, x) = m_0(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(iv)} & \gamma(x, t) = f(x, m(t)), \quad (x, t) \in Q, \\ \text{(v)} & P(t) = \phi \left( \int_{\mathbb{T}^d} v(x, t) m(x, t) \right), \quad t \in [0, T], \end{array} \right. \quad (\text{MFG}_{f,\phi})$$

where  $u(x, t), m(x, t), \gamma(x, t) \in \mathbb{R}$ ,  $v(x, t) \in \mathbb{R}^d$ , and  $P(t) \in \mathbb{R}^d$ . The above system is a coupled system of partial differential equations:

- Equation (i) is the Hamilton-Jacobi-Bellman equation, describing the evolution of the value function of the individual problem.
- Equation (ii) defines the optimal control  $v$ .
- Equation (iii) is the Fokker-Planck equation, describing the evolution of the state distribution of the agents. The term  $m_0$  is the initial distribution of the agents. Unless stated otherwise, solutions are understood in the sens of distributions.
- Equation (iv) defines the congestion and equation (v) the price.

The link between (2.7) and  $(\text{MFG}_{f,\phi})$  is the following: any solution  $(u, v, m, \gamma, P)$  to the mean field game system  $(\text{MFG}_{f,\phi})$  is such that  $(m, v)$  is solution to the Nash equilibrium problem; for any solution  $(m, v)$  solution to the Nash equilibrium problem, defining

$$\gamma(x, t) = f(x, \bar{m}(t)), \quad P(t) = \phi \left( \int_{\mathbb{T}^d} \bar{v}(x, t) \bar{m}(x, t) dx \right), \quad u = u[\gamma, P],$$

the quintuplet  $(u, v, m, \gamma, P)$  is a solution to the mean field game system  $(\text{MFG}_{f,\phi})$ .

**Fixed point structure of the mean field game system** The mean field game system  $(\text{MFG}_{f,\phi})$  can also be viewed as a fixed point problem. This point of view is a key tool for proving the existence of a solution, with the Schauder fixed-point theorem, for example. We present here the dependencies between each variables of the coupled system and we highlight the differences between congestion and price models. The following graphic illustrates this structure for two cases: classical mean field games, that is to say situations where  $\phi = 0$ ; and mean field games with interactions through the controls which corresponds to the general case considered here.

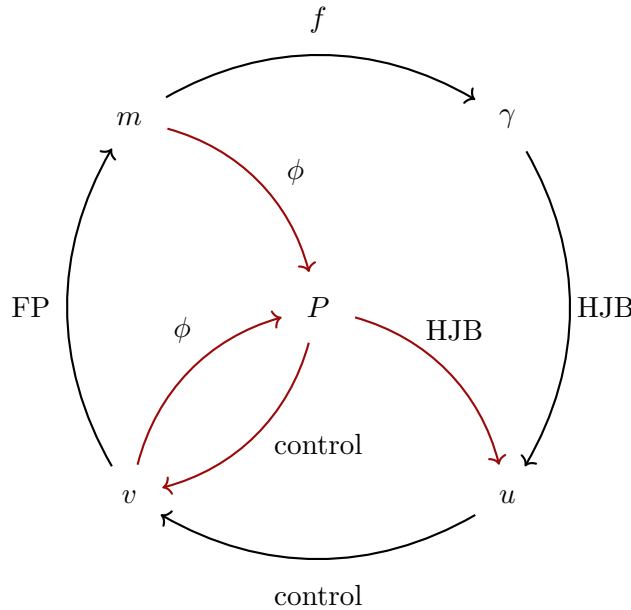


Figure 2.1: Unknowns dependency graph: fixed point structure of the mean field game problem.

The unknowns dependencies, in the context of classical mean field games, are represented by the black arrows. The structure is quite simple: given a congestion  $\gamma$ , one can compute the value function  $u$  solution to the Hamilton-Jacobi-Bellman (HJB) equation. Then the gradient of  $u$  provides the optimal control  $v$ . Given an optimal control  $v$ , one can compute the solution  $m$  to the Fokker-Planck (FP) equation. Finally the congestion  $\gamma$  depends on  $m$  via the mapping  $f$ .

Mean field games with interactions through the controls have a more complex structure: there are additional dependencies, represented by the red arrows, due to the price variable. Given a control  $v$  and a distribution  $m$ , one can compute the price  $P$  via the mapping  $\phi$ . Then the price  $P$  yields a value  $u$  via the Hamilton-Jacobi-Bellman equation and an associated control  $v$ . A proof of existence with a fixed point representation can be found in [21].

#### 2.1.4 Potential mean field games

Potential mean field games (also called variational in the literature) are mean field games which can be interpreted as the first order conditions of two control problems called potential problems. The first problem is a control problem of the Fokker-Planck equation and the second one an optimal control problem of the Hamilton-Jacobi-Bellman equation. There are several advantages in working with potential mean field games. They first offer a different way (from fixed point approaches) to prove the existence of a solution to the mean-field game, which can be deduced from solutions to the potential problems. Let us mention that depending on the model, this approach can be delicate to follow (in particular in the case of degenerate diffusion, see [34]). In addition, the potential form allows to use optimization algorithms to solve the mean field game system, as we will see in the next section.

In our framework, it is enough to assume that  $f$  and  $\phi$  derive from potentials to obtain a variational structure for the game. We say that  $f$  and  $\phi$  derive from potentials

if there exist two convex mappings  $F: \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}$  and  $\Phi: \mathbb{R}^k \rightarrow \mathbb{R}$  such that

$$\begin{aligned} F(m_2) - F(m_1) &= \int_0^1 \int_{\mathbb{T}^d} f(x, sm_2 + (1-s)m_1)(m_2(x) - m_1(x)) dx ds, \\ \phi(z) &= \nabla_z \Phi(z), \end{aligned}$$

for any  $m_1, m_2 \in \mathcal{D}_1(\mathbb{T}^d)$ , and any  $z \in \mathbb{R}^k$ . We define the following criterion

$$\mathcal{J}(m, v) = \int_Q L(v)m dx dt + \int_0^T \left( F(m(t)) + \Phi \left( \int_{\mathbb{T}^d} v(x, t)m(x, t) dx \right) \right) dt + \int_{\mathbb{T}^d} gm(T) dx,$$

and the following control problem:

$$\inf_{(m, v) \in \mathcal{R}} \mathcal{J}(m, v). \quad (\text{P})$$

Due to the terms  $L(v)m$  and  $\Phi(\int vm)$ , the criterion of the above problem is not convex. The idea is to study a closely related convex problem. We define the criterion

$$\tilde{\mathcal{J}}(m, w) = \int_Q \tilde{L}(m, w) dx dt + \int_0^T \left( F(m(t)) + \Phi \left( \int_{\mathbb{T}^d} v(x, t)m(x, t) dx \right) \right) dt + \int_{\mathbb{T}^d} gm(T) dx,$$

where  $\tilde{L}: Q \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  is the perspective function of  $L$ ,

$$\tilde{L}(m, w) = \begin{cases} mL(w/m), & \text{if } m > 0, \\ 0, & \text{if } m = 0 \text{ and } w = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that  $\tilde{L}$  is convex and lower semi-continuous with respect to  $(m, w)$ . We define the set

$$\tilde{\mathcal{R}} = \{(m, w) \in W^{2,1,p}(Q) \times W^{1,0,\infty}(Q), \partial_t m - \Delta m + \nabla \cdot w = 0, m(0) = m_0\}.$$

The criterion  $\tilde{\mathcal{J}}$  and the set  $\tilde{\mathcal{R}}$  are both convex so that the following problem is convex,

$$\inf_{(m, w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w). \quad (\tilde{\text{P}})$$

Provided that both problems have solutions  $(\bar{m}, \bar{v})$  and  $(\bar{m}, \bar{w})$  and that  $\bar{m}(x, t) > 0$  for any  $(x, t) \in Q$ , the solutions of the two problems are such that  $\bar{v} = \bar{w}/\bar{m}$ . We define the dual criterion

$$\tilde{\mathcal{D}}(u, \gamma, P) = \int_{\mathbb{T}^d} u(x, 0)m_0(x) dx - \int_0^T (F^*(\gamma(t)) + \Phi^*(P(t))) dt, \quad (2.8)$$

and consider the dual problem

$$\sup_{(u, \gamma, P) \in \tilde{\mathcal{U}}} \tilde{\mathcal{D}}(u, \gamma, P), \quad (\tilde{\text{D}})$$

where  $\tilde{\mathcal{U}}$  is the set of  $(u, \gamma, P) \in W^{2,1,p}(Q) \times L^\infty(Q) \times L^\infty(0, T; \mathbb{R}^k)$  satisfying

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^*P] &\leq \gamma, & (x, t) \in Q, \\ u(x, T) &\leq g(x), & x \in \mathbb{T}^d. \end{aligned}$$

**Duality and link with the mean field game system** Here we exhibit and explain the existing links between the above potential problems and the mean field game system. The following presentation is not a proof, rather a summary of the link that are expected to hold under suitable assumptions.

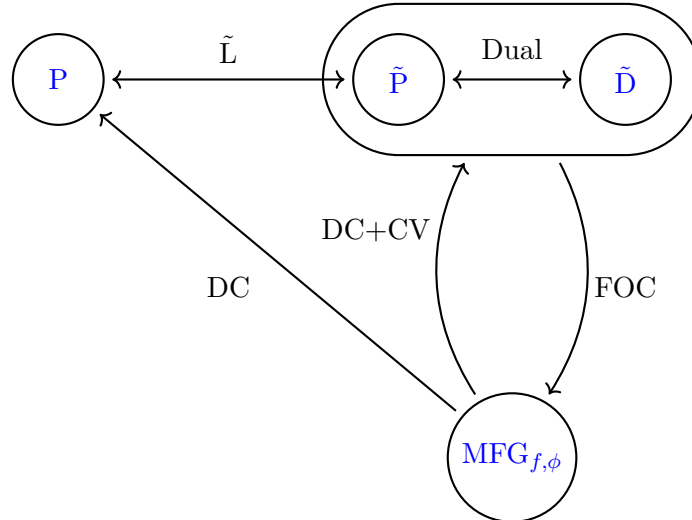


Figure 2.2: Summary of the links between  $(P)$ ,  $(\tilde{P})$ ,  $(\tilde{D})$ , and  $(MFG_{f,\phi})$ .

Let  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$  be a solution to  $(MFG_{f,\phi})$ , for the sake of simplicity we assume that  $\bar{m}(x, t) > 0$  for any  $(x, t) \in Q$ . We define the change of variable “à la Benamou-Brenier”  $\bar{w} = \bar{m}\bar{v}$  and we have  $\bar{v} = \bar{w}/\bar{m}$  by uniform positivity of  $m$ . We first present how to obtain a solution to  $(\tilde{P})$ ,  $(\tilde{D})$ , and  $(P)$ .

**DC:** By a direct calculus (essentially integration by parts) it is direct to show that any solution to  $(MFG_{f,\phi})$  yields a solution to  $(P)$ .

**DC+CV:** By a direct calculus (still integration by parts) and a change of variable  $\bar{w} = \bar{m}\bar{v}$ , any solution to  $(MFG_{f,\phi})$  yields a solution  $(\bar{m}, \bar{w})$  to  $(\tilde{P})$  and a solution  $(\bar{u}, \bar{\gamma}, \bar{P})$  to  $(\tilde{D})$ .

We now explain the link between  $(P)$  and  $(\tilde{P})$ . This link can be helpful since it is sometimes easier to show the uniqueness of a solution to  $(P)$  and then deduce the uniqueness of a solution to  $(\tilde{P})$ .

**$\tilde{L}$ :** The link between  $(P)$  and  $(\tilde{P})$  is given by the perspective function and the change of variable “à la Benamou-Brenier”. For any solution  $(\bar{m}, \bar{w})$  of problem  $(\tilde{P})$ , we have that  $(\bar{m}, \bar{v} = \bar{w}/\bar{m})$  is solution to  $(P)$ . Any solution  $(\bar{m}, \bar{v})$  to  $(P)$  is such that  $(\bar{m}, \bar{w} = \bar{m}\bar{v})$  is solution to problem  $(\tilde{P})$ . In addition the values of both problems are equal.

We finally explain the link between the primal problems  $(\tilde{P})$  and  $(\tilde{D})$ . We also explain why the solutions to those problems are solutions to the mean field game system  $(MFG_{f,\phi})$ .

**Dual:** Problems  $(\tilde{P})$  and  $(\tilde{D})$  are formally in duality. To establish such result, a classical approach is to use the Fenchel-Rockafellar Theorem [109]. Depending on the framework, the choice of the function spaces might be a delicate issue (see [34] for



the second-order degenerate case). In the discrete setting, this result is at the heart of primal-dual numerical methods.

**FOC:** The mean field game system can be interpreted as first-order conditions for the potential problems  $(\tilde{P})$  and  $(\tilde{D})$ . One can verify that for all solutions  $(\bar{m}, \bar{w})$  to  $(\tilde{P})$  and for all solutions  $(\bar{u}, \bar{\gamma}, \bar{P})$ , the quintuplet  $(\bar{m}, \bar{v} = \bar{w}/\bar{m}, \bar{u}, \bar{\gamma}, \bar{P})$  is a solution to  $(\text{MFG}_{f,\phi})$ . In chapter 4, in a discrete setting, we use the Fenchel-Young inequality to establish such result.

### 2.1.5 Numerical methods

The numerical resolution of mean field games has received significant attention over the last years. We refer the reader to the two surveys [6, 91] for an overview. Here we provide a synthetic presentation and a simple classification. We distinguish two approach: a MFG system based approach and a Potential approach. The first approach consists in solving directly the mean field game system via learning or Newton methods. The second approach consists in minimizing a potential whose solutions satisfy the mean field game system, via primal-dual or saddle-point methods.

Approach:	Class:	Methods:
MFG system	Learning	Fictitious play [35, 55, 77, 103] Reinforcement learning [9, 50, 75, 117, 119] Online mirror descent [76, 101]
	Newton	Newton algorithm [3, 4, 5, 96] Gauss-Newton algorithm [30, 31]
Potential	Primal	Conditional gradient (Frank-Wolfe) [62] Generalized conditional gradient [24] Sinkhorn [16] Stochastic gradient descent [42, 43] Uzawa's algorithm [18]
	Saddle-point	Lagrangian based [15, 17, 22, 97, 112] Primal-dual proximal methods [22, 27, 28]

Figure 2.3: Classification and summary of numerical methods for mean field games.

Learning methods are agent based methods: they adopt the point of view of a representative player. Learning algorithms go as follows: at each step, for a given belief on the strategy of the others, any representative agent computes its best response; then he/she learns by averaging all the best responses found from the beginning of the learning procedure. The learning question is central in game theory [60], and the fictitious play algorithm has an interpretations in social science or in engineering: Nash equilibria can be reached if all agents follow their personal interests, play the game sufficiently many times and repeatedly update their beliefs.

The starting point of the Newton methods is to consider the mean field game system as a problem of the form  $f(x) = 0$  for some differentiable function  $f$  of the unknowns  $x$  (typically the value function, the distribution of the agents, the coupling terms). To solve such problem, a Newton algorithm can be employed for example as in [5]. Notice that in the latter reference, due to the dependency between each variables of the problem, the

authors are able to recast the Newton algorithm as a fixed point method for the coupling variables.

In full generality these first two classes do not require any potential form of the game, even if in practice, a potential form or the monotony of the coupling terms can be assumed to ensure uniqueness of a solution and the convergence of the method. Notice that in [5], the authors do not assume monotonicity of the coupling terms.

On the contrary, the potential approach is based on the minimization of potential problems. In this approach we can distinguish between two class of methods:

1. Primal methods (resp. dual), based on the minimization (resp. maximization) of the primal (resp. dual) problem. In this category we find the conditional gradient algorithm, the generalized conditional gradient algorithm, Sinkhorn's algorithm and the stochastic gradient descent algorithms.
2. Saddle point methods are methods that explicitly exploit the primal-dual structure of the potential problems. In this category we have the Lagrangian methods (ALG2, ADMM, ADM-G) the primal-dual proximal methods, and Uzawa's algorithm.

Among all these references, only [5, 22, 24, 96] are related to mean field games with interactions through the controls.

## 2.2 Contributions

We now present the different chapters of the thesis. From now on, the notations are specific to each section. The following sections are structured as follows: summary of the framework and main objectives of the study, comparison with the literature, presentation of the model, presentation of the main results and the methodology, and finally some propositions of extensions or applications. The chapters of the thesis are independent and can be read in any order.

### 2.2.1 Chapter 1: Discrete-time mean field games with risk-averse agents

**Framework and objectives** The first chapter of the thesis is dedicated to the study of a discrete time mean field game model with risk averse agents. The risk aversion is modeled via coherent risk measures. In many economic situations, risk modeling is of interest, in particular in the banking industry [98]. This work is motivated by the development of mean field games in finance [36, 56], banking [40, 44], energy systems [8] or network systems [19], where risk and robust controls are central questions.

Here risk aversion is modeled with the help of composite risk measures (also called dynamic risk measures). Mathematically, a risk measure  $\rho$  is a map that assigns to a random variable  $U$  a real number, which is usually high when  $U$  is very volatile. In this way  $\rho$  can be used to model the reluctance of a player to face highly uncertain expenses. We refer to the seminal work by P. Artzner, F. Delbaen, J-M. Eber and D. Heath in [10]. We will make use of composite risk measures, the natural extension of risk measures to a multistage framework, see for example the article of A. Shapiro and A. Ruszczyński [111]; for an application to multistage portfolio selection one can refer to A. Shapiro [116].

In this chapter we study a population of identical agents which all optimize a linear discrete-time dynamical system (in a continuous state space). In the model, the associated cost function depends on a variable called belief, which is related to the behavior of the whole group, whence a coupling between a single agent and the population. Assuming that the population is very large, one can consider that an isolated representative agent has no impact on the belief. Therefore his/her behavior can be conveniently described by dynamic programming equations, in which the belief is a parameter. Assuming that all agents make use of the same feedback control, the distribution of their state is described by a Kolmogorov equation. Finally for any given distribution and control, the agents formulate a belief on the state action distribution.

**Comparison with the literature** Discrete-time and continuous-space mean field game models have been studied in different works. The framework that we propose in this article is close to the one of N. Saldi, T. Başar and M. Raginsky [113], in particular, we make use of similar weighted spaces. A few works have already investigated the issue of risk aversion. Most of them model risk sensitivity via exponential utility functions, see for example H. Tembine, Q. Zhu and T. Başar [118]. The case of robust mean field games is investigated in problem (P2) in the work of J. Moon and T. Başar [95]. From the risk-aversion modelization point of view, our work is closer to the reference [13], which investigates robust mean field games in the continuous time setting.

**The model** Let us fix  $T \in \mathbb{N}^*$ , let us set  $\mathcal{T} := \{0, \dots, T-1\}$  and  $\bar{\mathcal{T}} := \{0, \dots, T\}$ . We first define the running cost  $\ell: \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{B}_2 \rightarrow \mathbb{R}$  by

$$\ell(t, x, a, b) = \frac{1}{2}|a|^2 + \langle a, P(t, b) \rangle + F(t, x, b),$$

where  $F: \bar{\mathcal{T}} \times \mathbb{R}^d \times \mathcal{B}_2 \rightarrow \mathbb{R}$  is a congestion function and  $P: \mathcal{T} \times \mathcal{B}_2 \rightarrow \mathbb{R}^d$  is a price function. The set  $\mathcal{B}_2 := (\mathcal{P}_2(\mathbb{R}^{2d}))^T \times \mathcal{P}_2(\mathbb{R}^d)$  is the set of beliefs; here we use the notation  $\mathcal{P}_2(X)$  for the set of probability measures on  $X$  with finite second order moments. Let  $\bar{m} \in \mathcal{P}_2(\mathbb{R}^d)$  be an initial distribution and let  $(\nu(t))_{t \in \mathcal{T}} \in (\mathcal{P}_2(\mathbb{R}^d))^T$  be the individual noise distributions.

In this chapter we study the following risk averse mean field game system:

$$\left\{ \begin{array}{l} \text{(i)} \quad \begin{cases} u(t, x) = \inf_{a \in \mathbb{R}^d} \left( \ell(t, x, a, b) + \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} u(t+1, x+a+y) d\xi(y) \right), \\ u(T, x) = F(T, x, b), \end{cases} \\ \text{(ii)} \quad \alpha_t(x) = \arg \min_{a \in \mathbb{R}^d} \left( \ell(t, x, a, b) + \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} u(t+1, x+a+y) d\xi(y) \right), \\ \text{(iii)} \quad \begin{cases} m(t+1) = \nu(t) * [(id + \alpha_t) \# m(t)], \\ m(0) = \bar{m}, \end{cases} \\ \text{(iv)} \quad \mu(t) = (id, \alpha_t) \# m(t), \\ \text{(v)} \quad b := (\mu(0), \dots, \mu(T-1), m(T)), \end{array} \right. \quad (\rho\text{-MFG})$$

for any  $(t, x) \in \mathcal{T} \times \mathbb{R}^d$ . The unknowns of the problem are: the value function  $u$ , the feedback control  $\alpha$ , the distribution of states  $m$ , the joint distribution of states and controls  $\mu$  and the belief  $b$ .

We will describe in the next paragraph the optimal control problem leading to the dynamic programming equations (i) and (ii). These equations involve the sets  $\mathcal{M}_t$  defined by

$$\mathcal{M}_t := \left\{ \xi \in \mathcal{P}(\mathbb{R}^d), d\xi = Z d\nu(t), Z \in \mathcal{Z}_t \right\},$$

where  $(\mathcal{Z}_t)_{t \in \mathcal{T}}$  is a family of given nonempty and convex subsets

$$\mathcal{Z}_t \subseteq \left\{ Z \in L^\infty(\mathbb{R}^d), \int_{\mathbb{R}^d} Z(y) d\nu(t, y) = 1, Z \geq 0 \right\}, \quad \forall t \in \mathcal{T}.$$

Remark that in the particular case where the set  $\mathcal{M}_t$  is the singleton  $\{\nu(t)\}$ , we recover dynamic programming equations corresponding to the risk neutral case.

**Risk measures and individual control problem** We define here the individual control problem. We introduce first a general notation: given a tuple  $(y_0, y_1, \dots, y_T)$ , we write  $y_{[t]} = (y_0, \dots, y_t)$ . Let  $X_0$  and  $(Y_t)_{t \in \mathcal{T}}$  be  $(T + 1)$ -independent random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $\mathcal{L}(X_0) = \bar{m}$  and  $\mathcal{L}(Y_t) = \nu(t)$ . We define the filtration  $(\mathcal{F}_t)_{t \in \mathcal{T}}$ , where  $\mathcal{F}_0 := \sigma(X_0)$  is the sigma-algebra generated by  $X_0$ , and  $\mathcal{F}_{t+1} := \sigma(X_0, Y_{[t]})$ . We denote for any  $t \in \bar{\mathcal{T}}$  and any  $p \in [1, +\infty)$

$$\mathbb{L}_t^p(\Omega, \mathbb{R}^{d'}) := L^p(\Omega, \mathcal{F}_t, \mathbb{P}, \mathbb{R}^{d'}),$$

the space of  $\mathcal{F}_t$  measurable random variables with finite  $p$ -th order moment and value in  $\mathbb{R}^{d'}$ . When the dimension is  $d' = 1$ , we simplify the notation:  $\mathbb{L}_t^p := \mathbb{L}_t^p(\Omega, \mathbb{R})$ .

We fix a family of *one-step conditional risk mappings*  $(\rho_t)_{t \in \mathcal{T}}$ ,  $\rho_t: \mathbb{L}_{t+1}^1 \rightarrow \mathbb{L}_t^1$ , defined by

$$\rho_t(U_{t+1})(x_0, y_{[t-1]}) = \sup_{Z \in \mathcal{Z}_t} \int_{\Omega} U_{t+1}(x_0, y_{[t-1]}, Y_t(\omega)) Z(Y_t(\omega)) d\mathbb{P}(\omega), \quad (2.9)$$

where the random variables  $U_{t+1}$  and  $\rho_t(U_{t+1})$  are explicitly represented as measurable functions of  $(x_0, y_{[t]}) \in \mathbb{R}^{(t+2)d}$  and  $(x_0, y_{[t-1]}) \in \mathbb{R}^{(t+1)d}$ , respectively. By definition of  $\mathcal{M}_t$ , we have

$$\rho_t(U_{t+1})(x_0, y_{[t-1]}) = \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} U_{t+1}(x_0, y_{[t-1]}, y_t) d\xi(y_t).$$

Finally we construct the associated composite risk measure  $\rho: \mathbb{L}_T^1 \rightarrow \mathbb{R}$ ,

$$\rho(U) := \mathbb{E}[\rho_0 \circ \dots \circ \rho_{T-1}(U)].$$

Now we are ready to describe the individual risk averse control problem of each agent for fixed belief  $b \in \mathcal{B}_2$ . Given a control  $A \in \mathcal{A} := \mathcal{A}_0 \times \dots \times \mathcal{A}_{T-1}$ , with  $\mathcal{A}_t = \mathbb{L}_t^2(\Omega, \mathbb{R}^d)$  for any  $t \in \mathcal{T}$ , the evolution of the state of the representative player is given by

$$X_{t+1} = X_t + A_t + Y_t, \quad \forall t \in \mathcal{T}.$$

Note that by induction,  $X_t \in \mathbb{L}_t^2(\Omega, \mathbb{R}^d)$  for any  $t \in \bar{\mathcal{T}}$ . The problem of the representative agent is given by

$$\inf_{A \in \mathcal{A}} \mathcal{J}(A, b) := \rho \left( \sum_{t=0}^{T-1} \ell(t, X_t, A_t, b) + F(T, X_T, b) \right).$$

The dynamic programming equations (i-ii) allow to characterize the solution to the above problem. To motivate this assertion, observe that the individual control problem can be recasted in the following nested form:

$$\begin{aligned} \inf_{A \in \mathcal{A}} \mathcal{J}(A, b) = & \mathbb{E} \left[ \inf_{A_0 \in \mathcal{A}_0} \ell(0, X_0, A_0, b) + \rho_0 \left( \inf_{A_1 \in \mathcal{A}_1} \ell(1, X_1, A_1, b) + \dots \right. \right. \\ & \left. \left. + \rho_{T-2} \left( \inf_{A_{T-1} \in \mathcal{A}_{T-1}} \ell(T-1, X_{T-1}, A_{T-1}, b) + \rho_{T-1} \left( F(T, X_T, b) \right) \right) \dots \right) \right]. \end{aligned}$$

Recalling that for any  $t \in \mathcal{T}$ , the conditional risk measure  $\rho_t$  can be written as a supremum by (2.9), one can show (Proposition 3.3.5) that

$$\inf_{A \in \mathcal{A}} \mathcal{J}(A, b) = \mathbb{E} [u(0, X_0)].$$

In addition the unique solution  $\alpha$  to equation (ii) provides the unique solution

$$\bar{A} := (\alpha_0(\bar{X}_0), \dots, \alpha_{T-1}(\bar{X}_{T-1}))$$

to the individual control problem, where  $(\bar{X}_t)_{t \in \bar{\mathcal{T}}}$  is the solution to the closed-loop system

$$\bar{X}_{t+1} = \bar{X}_t + \alpha_t(\bar{X}_t) + Y_t, \quad \mathcal{L}(\bar{X}_0) = m_0, \quad \forall t \in \mathcal{T}.$$

We can finally give an interpretation of the remaining equations in the model. Equation (iii) yields the probability distribution of the state variable  $(\bar{X}_t)_{t \in \bar{\mathcal{T}}}$ , solution to the above closed-loop system. Equation (iv) yields the probability distribution of the pair  $(\bar{X}_t, \bar{A}_t)$ . Finally, the belief  $b$  is defined through equation (v).

**Methodology and main results** Our first result is an existence result, obtained with a standard fixed point approach (see for example the approach followed in the proof of [32, Theorem 3.1]). Under suitable assumptions on the data (essentially Lipschitz continuity of the congestion and price mappings, but no monotonicity assumptions) we have the following:

- **Existence:** There exists a solution  $(u, \alpha, m, \mu, b)$  to the mean field game system.

Our approach for proving the existence of a solution consists in formulating the system ( $\rho$ -MFG) as a fixed point equation. For this purpose, we consider two mappings. The first one, that we call dynamic programming mapping, assigns to a belief  $b$  the solutions  $u^*(b)$  and  $\alpha^*(b)$  to equations ( $\rho$ -MFG,i) and ( $\rho$ -MFG,ii), respectively. The second one, the Kolmogorov mapping, assigns to a feedback control  $\alpha$  the triplet  $(m^*(\alpha), \mu^*(\alpha), b^*(\alpha))$ , where  $m^*(\alpha)$ ,  $\mu^*(\alpha)$ , and  $b^*(\alpha)$  are the solutions to ( $\rho$ -MFG,iii), ( $\rho$ -MFG,iv), and ( $\rho$ -MFG,v), respectively. They allow to reformulate the system ( $\rho$ -MFG) as an equivalent fixed point equation

$$b = b^* \circ \alpha^*(b).$$

Then the result is obtained by a Schauder's fixed point argument. The main difficulties to obtain such result concern: the choice of functional spaces for the variables of the problem and the proof of the continuity of the mapping  $\alpha^*$ .

Now we turn to our second main result, concerning the link between the mean field game model and an  $N$ -player game. Let us first define the  $N$ -player game. Let  $\mathcal{N} = \{1, \dots, N\}$  be the set of players. For any vector  $(x^1, \dots, x^N)$  we denote

$$\begin{aligned} \mathbf{x} &= (x^1, \dots, x^N), \\ \mathbf{x}^{-i} &= (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^N). \end{aligned}$$

We consider independent random variables  $(X_0^i)_{i \in \mathcal{N}}$  and  $(Y_t^i)_{i \in \mathcal{N}, t \in \mathcal{T}}$ , such that  $\mathcal{L}(X_0^i) = \bar{m}$  and  $\mathcal{L}(Y_t^i) = \nu(t)$ . We define the filtration  $(\mathcal{F}_t)_{t \in \mathcal{T}}$  as follows:  $\mathcal{F}_0 := \sigma(\mathbf{X}_0)$  is the sigma-algebra generated by  $\mathbf{X}_0$ ,  $\mathcal{F}_{t+1} := \sigma(\mathbf{X}_0, \mathbf{Y}_{[t]})$ . We denote  $L_t^2(\Omega, \mathbb{R}^d) := L^2(\Omega, \mathcal{F}_t, \mathbb{P}, \mathbb{R}^d)$  the space of  $\mathcal{F}_t$  measurable random variables with finite 2-nd order moment and value in  $\mathbb{R}^d$ . For any  $t \in \mathcal{T}$ , we consider the control set

$$\mathcal{A}_t := L_t^2(\Omega, \mathbb{R}^d), \quad \mathcal{A} := \mathcal{A}_0 \times \cdots \times \mathcal{A}_{T-1}.$$

The control of player  $i \in \mathcal{N}$  is an adapted stochastic process  $A^i \in \mathcal{A}$ , whose associated trajectory  $(X_t^i[A^i])_{t \in \mathcal{T}}$  is defined by the following state equation

$$X_{t+1}^i = X_t^i + A_t^i + Y_t^i.$$

For any  $\mathbf{A} = (A^i, \mathbf{A}^{-i}) \in \mathcal{A}^N$ , the cost of each player  $i \in \mathcal{N}$  is given by

$$\mathcal{J}^{i,N}(A^i, \mathbf{A}^{-i}) := \rho^i \left( \sum_{t=0}^{T-1} \ell(t, X_t^i[A^i], A_t^i, b_{\mathbf{A}}^N) + F(T, X_T^i[A^i], b_{\mathbf{A}}^N) \right),$$

where  $\rho^i$  is the individual risk measure (see Section 3.6.1, in particular notice that players are risk averse with respect to their individual noise only) and  $b_{\mathbf{A}}^N$  is the empirical belief (that is to say the  $N$ -player counterpart of the belief  $b$ ).

Now we construct an approximate solution to the game, deduced from a fixed solution  $(\bar{u}, \bar{\alpha}, \bar{m}, \bar{\mu}, \bar{b})$  to the mean field game system. We denote  $(\bar{X}_t^i)_{t \in \mathcal{T}}$  the solution to the closed-loop system

$$X_{t+1}^i = X_t^i + \bar{\alpha}_t(X_t^i) + Y_t^i.$$

We define the control  $\bar{A}^i \in \mathcal{A}$  by  $\bar{A}_t^i = \bar{\alpha}_t(\bar{X}_t^i)$  and we set  $\bar{\mathbf{A}} = (\bar{A}^1, \dots, \bar{A}^N)$ .

- **$\varepsilon$ -Nash equilibrium in the  $N$ -player game:** There exists a constant  $C > 0$ , independent of  $N$ , such that the  $N$ -uplet  $\bar{\mathbf{A}}$  defined above is an  $\varepsilon$ -Nash equilibrium, that is to say

$$\mathcal{J}^{i,N}(\bar{A}^i, \bar{\mathbf{A}}^{-i}) \leq \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) + \varepsilon,$$

with  $\varepsilon := CN^{-\tau/2}$  and  $\tau > 0$  depending only on the dimension  $d$  of the state space. In addition we have that

$$|\mathcal{J}^{i,N}(\bar{\mathbf{A}}) - \mathcal{J}^i(\bar{A}^i, \bar{b})| \leq CN^{-\tau/2},$$

where  $\mathcal{J}^i(A^i, \bar{b}) := \rho^i \left( \sum_{t=0}^{T-1} \ell(t, X_t^i[A^i], A_t^i, \bar{b}) + F(T, X_T^i[A^i], \bar{b}) \right)$ .

The proof consists in two main steps. In the first step, we estimate the distance between the empirical belief  $b_{\mathbf{A}}^N$  of the  $N$ -player game and the belief  $\bar{b}$  of the mean field model. This is done via concentration inequalities [58]. This allows us to estimate the sensibility of the criterion  $\mathcal{J}^{i,N}(A^i, \mathbf{A}^{-i})$  with respect to its second variable, in a second step.

**Extension and applications** Let us discuss some possible extensions and applications of this work.

- In our analysis we require the uniqueness of the feedback  $\alpha$ , in particular to apply the Schauder's fixed point theorem. Recall that  $\alpha$  is characterized via equation (ii). In order to prove the uniqueness of the optimization problem involved in (ii), we need to ensure that the value function is convex, which can be achieved by

assuming that the congestion terms and the running cost are convex. We think that the non-convex case can be treated following the ideas developed in [113] and would rely on Kakutani's fixed point theorem and would require that the set of strategies is compact.

- ▶ The risk measures that we consider here do not have the most general structure possible. In our setting, the sets  $\mathcal{M}_t$  are fixed. In [110], these sets depend on the current state and control (see in particular Sections 4 and 5). In this more general context, it is still possible to derive a dynamic programming principle for the underlying optimal control problem (see [110, Theorem 2]). However, the convexity of the value function, which plays an important role in our analysis, is lost in such a setting.
- ▶ Another possible extension is to consider common noise in our framework. Two cases could be distinguished, depending on whether or not the agents are risk averse to the common noise. In the risk neutral and continuous case, this extension is classical and leads to the study of a master equation (an infinite dimensional Hamilton-Jacobi equation posed over the space of probability measures). We refer the reader to [33, 39, 86]. But up to our knowledge, there is no work available about a master equation in discrete time and continuous space setting. A first step, which does not necessitate to introduce the master equation, could be to consider a common noise described by a finite scenario tree. This would require to make the variables of the game depend on the scenarios.
- ▶ A last direction concerns numerical methods. Up to our knowledge, no potential formulation is available in this framework. As a consequence, only learning techniques do apply: one could for example combine the stochastic dual dynamic programming method [100] to solve the risk averse dynamic programming equation with the fictitious play algorithm.

### 2.2.2 Chapter 2: Discrete potential mean field games

**Framework and objectives** In the second chapter of the thesis, we propose and study a discrete time, discrete state space potential (also called variational) mean field game problem with hard constraints. In our framework, the dynamical system of each agent is a Markov chain, with controlled probability transitions. The congestion  $\gamma$  is linked to the distribution of the agents via the subdifferential of a proper convex and l.s.c. potential  $F$ . The price  $P$  is related to the joint law of states and controls of the agents via the subdifferential of a proper convex and l.s.c. potential  $\phi$ . We consider hard constraints, which means that the potentials  $F$  and  $\phi$  can take the value  $+\infty$  and thus induce constraints on the distribution of the agents. Our results are motivated by the study of discrete models as such but can be applied to discretized MFGs.

The chapter is organized in three parts:

1. The first part is devoted to the theoretical analysis of the MFG system. We first introduce a potential problem, involving the Kolmogorov equation, closely related to another optimization problem, which is convex. This second problem is obtained via a change of variable, similar to the one commonly employed in the continuous setting (e.g. in [14]). Under a suitable qualification condition, we establish a duality result between this problem and an optimal control problem involving the dynamic programming equation. We show the existence of solutions to these problems and

finally we show the existence of a solution to the MFG system. A uniqueness result is proved (when  $F$  and  $\phi$  are differentiable).

2. The second part is devoted to the numerical resolution of the MFG system. We focus on two families of methods: primal-dual methods and augmented Lagrangian methods. These two classes exploit the duality structure of the potential problems and can deal with hard constraints. They have already been applied to continuous MFGs, see for example the survey article [6].
3. In the third part we propose and solve numerically two hard constraints problems: a congestion mean field game problem and a ‘‘Cournot’’ mean field game. Following our analysis we define a notion of residuals allowing us to compare the empirical convergence of each method in a common setting.

**Comparison with the literature** Few publications deal with fully discrete models; in a seminal work, D. Gomes, J. Mohr, and R. R. Souza [65] have studied the existence of a Nash equilibrium via a fixed point approach and investigated the long-term behavior of the game. Up to our knowledge, this study is the first to analyze discrete and potential mean field games. In the continuous case, four classes (summarized in the figure 2.4) of variational mean field games can be identified. Our model is general enough to be seen as the discrete counterpart of these four cases.

	Soft	Hard
$F$	Case 1 ( $\gamma = \nabla F$ ): [17, 34, 89, 106]	Case 2 ( $\gamma \in \partial F$ ): [38, 93, 114, 115]
$\phi$	Case 3 ( $P = \nabla \phi$ ): [21, 67, 68, 70, 72]	Case 4 ( $P \in \partial \phi$ ): [66]

Figure 2.4: Literature for soft and hard interactions in potential mean field games.

- Case 1: MFGs with monotone congestion terms ( $F$  is differentiable,  $\phi = 0$ ).
- Case 2: MFGs with density constraints ( $F$  has a bounded domain,  $\phi = 0$ ). These models are of particular interest for describing crowd motions. The coupling variable  $\gamma$  has there an incentive role.
- Case 3: MFGs with Cournot interactions ( $F = 0$ ,  $\phi$  is differentiable). In this situation, each agent optimally chooses a quantity to be sold at each time step of the game. Interactions with the other players occur through the gradient of  $\phi$  which maps the mean strategy (the market demand) to a market price.
- Case 4: MFGs with price formation ( $F = 0$ ,  $\phi$  has a bounded domain). These models incorporate a hard constraint on the demand. The price variable is the associated Lagrange multiplier and has an incentive role.

**The model** Let  $T \in \mathbb{N}^*$  be the duration of the game. Let  $\mathcal{T} = \{0, \dots, T - 1\}$  and  $\bar{\mathcal{T}} = \{0, \dots, T\}$ . Let  $S = \{0, \dots, n - 1\}$  denote the state space and  $\Delta(S) = \{\pi: S \rightarrow [0, 1] \mid \sum_{x \in S} \pi(x) = 1\}$  be the simplex over  $S$ . Given a finite set  $X$ , we denote by  $\mathbb{R}(X)$  the set of maps from  $X$  to  $\mathbb{R}$  (i.e.  $\mathbb{R}(X) = \mathbb{R}^X$ ). With this notation at hand, we define



the following spaces:

$$\begin{aligned}\mathcal{R} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T} \times S^2), & \mathcal{U} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T}), \\ \mathcal{C} &= \mathcal{R} \times \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T}), & \mathcal{K} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathcal{U}.\end{aligned}$$

We fix an initial distribution  $m_0 \in \Delta(S)$  and four maps: a running cost  $\ell$ , a potential price function  $\phi$ , a potential congestion cost  $F$ , and a displacement cost  $\alpha$ ,

$$\begin{aligned}\ell: \mathcal{T} \times S \times \mathbb{R}(S) &\rightarrow \mathbb{R} \cup \{+\infty\}, & \phi: \mathcal{T} \times \mathbb{R} &\rightarrow \mathbb{R} \cup \{+\infty\}, \\ F: \bar{\mathcal{T}} \times \mathbb{R}(S) &\rightarrow \mathbb{R} \cup \{+\infty\}, & \alpha: \mathcal{T} \times S^2 &\rightarrow \mathbb{R}.\end{aligned}$$

**Assumption 2.2.1** (Convexity). *For any  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ , the maps  $\ell(t, x, \cdot)$ ,  $F(s, \cdot)$ , and  $\phi(t, \cdot)$  are proper, convex and lower semi-continuous. In addition  $\text{dom}(\ell(t, x, \cdot)) \subseteq \Delta(S)$ .*

We define the individual cost  $c: \mathcal{T} \times S \times S \times \Delta(S) \rightarrow \mathbb{R}$ ,

$$c_{\gamma, P}(t, x, y, \rho) = \ell(t, x, \rho) + \gamma(t, x) + \alpha(t, x, y)P(t),$$

which is the cost of moving from state  $x \in S$  to state  $y \in S$  at time  $t \in \mathcal{T}$  when the control  $\rho \in \Delta(S)$  is implemented by the agent, given the coupling terms  $(\gamma, P)$ . When an agent is at state  $x$  at time  $t$  and plays the control  $\rho$ , he reaches the state  $y$  at time  $t + 1$  with probability  $\rho(y)$ .

The mean field game problem under study is the following: find a quintuplet  $(m, \pi, u, \gamma, P)$  such that for any  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ ,

$$\left\{ \begin{array}{l} \text{(i)} \quad \begin{cases} u(t, x) = \inf_{\rho \in \Delta(S)} \sum_{y \in S} \rho(y) \left( c_{\gamma, P}(t, x, y, \rho) + u(t+1, y) \right), \\ u(T, x) = \gamma(T, x), \end{cases} \\ \text{(ii)} \quad \pi(t, x, \cdot) \in \arg \min_{\rho \in \Delta(S)} \sum_{y \in S} \rho(y) \left( c_{\gamma, P}(t, x, y, \rho) + u(t+1, y) \right), \\ \text{(iii)} \quad \begin{cases} m(t+1, x) = \sum_{y \in S} m(t, y) \pi(t, y, x), \\ m(0, x) = m_0(x), \end{cases} & (\partial\text{-MFG}) \\ \text{(iv)} \quad \gamma(s, \cdot) \in \partial F(s, m(s, \cdot)), \\ \text{(v)} \quad P(t) \in \partial \phi \left( t, \sum_{(x, y) \in S^2} m(t, x) \pi(t, x, y) \alpha(t, x, y) \right). \end{array} \right.$$

The unknowns of the system are  $((m, \pi), (u, \gamma, P)) \in \mathcal{R} \times \mathcal{K}$ . They can be described as follows:  $u(t, x)$  is the value function of the agents;  $\pi(t, x, y)$  denotes the probability transition from  $x \in S$  to  $y \in S$ , for agents located at  $x$  at time  $t$ ;  $m(t, x)$  denotes the proportion of agents located at  $x \in S$  at time  $t \in \bar{\mathcal{T}}$ ;  $\gamma$  and  $P$  are the coupling terms of the MFG:  $\gamma(t, x)$  is a congestion term incurred by agents located at  $x \in S$  at time  $t \in \bar{\mathcal{T}}$  and  $P(t)$  is a price variable.

**The potential problems** We introduce the operator  $\mathbf{A}: \mathbb{R}(\mathcal{T} \times S^2) \rightarrow \mathbb{R}(\mathcal{T})$  and the operator  $\mathbf{S}: \mathbb{R}(\mathcal{T} \times S^2) \rightarrow \mathbb{R}(\bar{\mathcal{T}} \times S)$ , defined by

$$\mathbf{A}[w](t) = \sum_{(x,y) \in S^2} w(t,x,y) \alpha(t,x,y), \quad \mathbf{S}[w](s,x) = \begin{cases} \sum_{y \in S} w(s-1,y,x) & \text{if } s > 0, \\ 0 & \text{if } s = 0. \end{cases}$$

To study the mean field game system, we define four problems. We define the following minimization problem

$$\begin{cases} \min_{(m,\pi) \in \mathcal{R}} \mathcal{J}(m,\pi) := \sum_{(t,x) \in \mathcal{T} \times S} \ell[\pi](t,x) m(t,x) + \sum_{t \in \mathcal{T}} \phi[\mathbf{A}w](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m](s), \\ \text{s.t.:} \begin{cases} m(t+1,x) = \sum_{y \in S} \pi(t,y,x) m(t,y), & (t,x) \in \mathcal{T} \times S, \\ m(0,x) = m_0(x), & x \in S, \end{cases} \end{cases} \quad (\text{P})$$

where  $\ell$  is the Nemytskii operator associated to  $\ell$  and  $\tilde{\ell}$  is the perspective function associated to  $\ell$ . We also define the convex optimization problem (that we call primal problem)

$$\begin{cases} \min_{(m,w) \in \mathcal{R}} \tilde{\mathcal{J}}(m,w) := \sum_{(t,x) \in \mathcal{T} \times S} \tilde{\ell}[m,w](t,x) + \sum_{t \in \mathcal{T}} \phi[\mathbf{A}w](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m](s), \\ \text{s.t.:} \begin{cases} m(t+1,x) = \sum_{y \in S} w(t,y,x), & (t,x) \in \mathcal{T} \times S, \\ m(0,x) = m_0(x), & x \in S. \end{cases} \end{cases} \quad (\tilde{\text{P}})$$

We define the following dual problem

$$\begin{cases} \max_{(u,\gamma,P) \in \mathcal{U}} \tilde{\mathcal{D}}(u,\gamma,P) := \langle \bar{m}_0, u(0) \rangle - \sum_{t \in \mathcal{T}} \phi^*[P](t) - \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}^*[\gamma](s), \\ \text{s.t.:} \begin{cases} u(t,x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t,x) \leq \gamma(t,x), & (t,x) \in \mathcal{T} \times S, \\ u(T,x) = \gamma(T,x), & x \in S. \end{cases} \end{cases} \quad (\tilde{\text{D}})$$

Finally, let  $\mathbf{U}[\gamma,P]$  be the solution to the dynamic programming equation ( $\partial$ -MFG,i), for fixed coupling terms  $(\gamma,P)$ . We define the problem

$$\max_{(\gamma,P) \in \mathcal{U}} \tilde{\mathcal{D}}(\gamma,P) := \mathcal{D}(\mathbf{U}[\gamma,P], \gamma, P). \quad (\text{D})$$

**Methodology and main results** A first part of the results concerns existence/uniqueness of a solution to the mean field game system. Let us start with the following:

- **Existence and uniqueness:** There exists a solution  $(m,\pi,u,\gamma,P)$  to ( $\partial$ -MFG). If  $F$  and  $\phi$  are differentiable then  $(u,\gamma,P)$  is uniquely defined. If  $F$  and  $\phi$  are strictly convex then ( $\partial$ -MFG) has a unique solution.

Then we explain the link between all the optimization problems and the mean field game system. The following results hold under the following qualification assumption. For any  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \mathcal{K}$  and  $\pi \in \mathbf{dom}(\ell)$  we define  $m_1[\varepsilon, \pi]$  the solution to the following perturbed Kolmogorov equation

$$m_1(t+1,x) = \sum_{y \in S} m_1(t,y) \pi(t,y,x) - \varepsilon_1(t+1,x), \quad m_1(0) - \varepsilon_1(0) = \bar{m}_0.$$

We also define, for any  $(t,x,y) \in \mathcal{T} \times S \times S$ ,

$$\begin{aligned} w[\varepsilon, \pi](t,x,y) &= m_1[\varepsilon, \pi](t,x) \pi(t,x,y) \\ m_2[\varepsilon, \pi](t,x) &= m_1[\varepsilon, \pi](t,x) + \varepsilon_2(t,x) \\ D[\varepsilon, \pi](t) &= \sum_{(x,y) \in S^2} w[\varepsilon, \pi](t,x,y) \alpha(t,x,y) + \varepsilon_3(t). \end{aligned}$$

**Assumption 2.2.2** (Qualification). *There exists  $\alpha > 0$  such that for any  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$  in  $\mathcal{K}$  with  $\|\varepsilon\| \leq \alpha$ , there exists  $\pi \in \mathbf{dom}(\ell)$  such that*

$$m_1[\varepsilon, \pi] \geq 0, \quad m_2[\varepsilon, \pi] \in \mathbf{dom}(\mathbf{F}), \quad D[\varepsilon, \pi] \in \mathbf{dom}(\phi).$$

We provide a graphical representation to summarize the results. The labels on the arrows correspond to one or several results proved in the chapter, that we have recast here for presentation purpose.

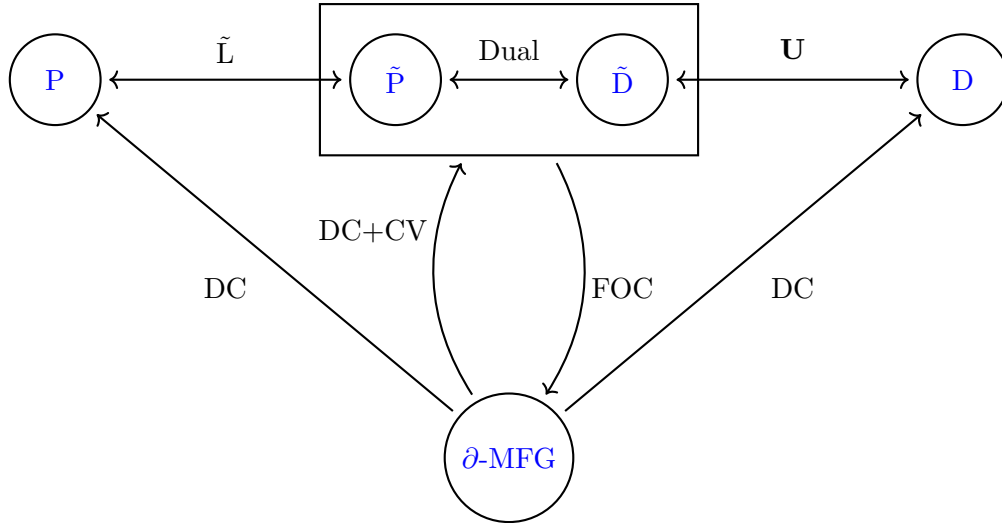


Figure 2.5: Summary of the results obtained in Chapter 4.

- **Dual:** Problem  $(\tilde{\mathbf{P}})$  and  $(\tilde{\mathbf{D}})$  are in duality.
- **DC+CV:** Let  $(m, \pi, u, \gamma, P)$  be a solution to  $(\partial\text{-MFG})$ . Then  $(m, w := m\pi)$  is solution to  $(\tilde{\mathbf{P}})$  and  $(u, \gamma, P)$  is solution to  $(\tilde{\mathbf{D}})$ .
- **FOC:** Let  $(m, w)$  be a solution to  $(\tilde{\mathbf{P}})$  and  $(u, \gamma, P)$  be a solution to  $(\tilde{\mathbf{D}})$ . Let  $\pi$  be such that for any  $(t, x) \in \mathcal{T} \times \mathcal{S}$ ,  $\pi(t, x, \cdot) = w(t, x, \cdot)/m(t, x)$  if  $m(t, x) > 0$  and

$$\pi(t, x, \cdot) \in \arg \min_{\rho \in \Delta(\mathcal{S})} \ell(t, x, \rho) + \sum_{y \in \mathcal{S}} \rho(y) (P(t) \alpha(t, x, y) + u(t+1, y))$$

if  $m(t, x) = 0$ . Then  $(m, \pi, u, \gamma, P)$  is a solution to  $(\partial\text{-MFG})$ .

- **L-tilde:** We have  $\text{val}(\mathbf{P}) = \text{val}(\tilde{\mathbf{P}})$ . By the ‘‘Benamou-Brenier’’ change of variable, any solution to Problem  $(\mathbf{P})$  provides a solution to Problem  $(\tilde{\mathbf{P}})$  and vice-versa.
- **DC:** Let  $(m, \pi, u, \gamma, P)$  be a solution to  $(\partial\text{-MFG})$  then  $(m, \pi)$  is solution to  $(\mathbf{P})$ , and  $(\gamma, P)$  is solution to  $(\mathbf{D})$ .
- **U:** Problems  $(\tilde{\mathbf{D}})$  and  $(\mathbf{D})$  have the same value. Moreover, for any solution  $(u, \gamma, P)$  to  $(\tilde{\mathbf{D}})$ ,  $(\gamma, P)$  is a solution to  $(\mathbf{D})$ ; conversely, for any solution  $(\gamma, P)$  to  $(\mathbf{D})$  (there exists at least one),  $(\mathbf{U}[\gamma, P], \gamma, P)$  is a solution to  $(\tilde{\mathbf{D}})$ .

A second part of the results concerns numerical aspects. We investigate the numerical resolution of the primal and dual problems in the hard constraint case for both congestion and price problems. We investigate different numerical methods: primal-dual proximal algorithms (called Chambolle-Pock and Chambolle-Pock-Bregman) [45, 46], alternating

direction method of moments (ADMM)[61, 64], and Gaussian back alternating direction method (ADM-G) [78]. The numerical methods are saddle-point methods which rely on the following primal-dual formulation:

$$\begin{aligned} \inf_{(m_1, w, m_2, D) \in \mathcal{C}} \mathcal{F}(m_1, w, m_2, D) + \mathcal{G}(\mathcal{A}(m_1, w, m_2, D)) = \\ \inf_{(u, \gamma, P) \in \mathcal{K}} \mathcal{F}^*(-\mathcal{A}^*(u, \gamma, P)) + \mathcal{G}^*(u, \gamma, P), \end{aligned}$$

where the maps  $\mathcal{F}: \mathcal{C} \rightarrow \mathbb{R} \cup \{+\infty\}$  and  $\mathcal{G}: \mathcal{K} \rightarrow \mathbb{R} \cup \{+\infty\}$  and the operator  $\mathcal{A}: \mathcal{C} \rightarrow \mathcal{K}$  are defined by

$$\begin{aligned} \mathcal{F}(m_1, w, m_2, D) &= \sum_{(t,x) \in \mathcal{T} \times \mathcal{S}} \tilde{\ell}[m_1, w](t, x) + \sum_{t \in \mathcal{T}} \phi[D](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m_2](s), \\ \mathcal{G}(y_1, y_2, y_3) &= \chi(y_1 + \bar{m}_0) + \chi(y_2) + \chi(y_3), \\ \mathcal{A}(m_1, w, m_2, D) &= (\mathbf{S}w - m_1, m_1 - m_2, \mathbf{A}w - D). \end{aligned}$$

The idea is to introduce additional variables to decouple the perspective function from the potential mapping in the criterion  $\mathcal{F}$ . This formulation is important in our numerical implementation. Indeed our numerical methods require to compute the proximal operator of  $\mathcal{F}$  at each step. This way, the proximal operator of  $\mathcal{F}$  can be decomposed as the concatenation of three tractable proximal operators. The coupling between the variables is recovered in the term  $\mathcal{G}(\mathcal{A}(m_1, w, m_2, D))$ . The counterpart of this decoupling is the increase of the number of successive minimization steps that have to be done in ADMM. Then convergence guarantee are lost for the latter algorithm since we are not in the ALG2 framework. However, we implement a close method called ADM-G which offers a convergence guarantee. We summarize in the following table the convergence guarantee for each method.

Method	Convergence guarantee
ADMM	No
ADM-G	Yes
Chambolle-Pock	Primal dual gap convergence $O(1/k)$
Chambolle-Pock-Bregman	Primal dual gap convergence $O(1/k)$

Figure 2.6: Convergence guarantee and execution time.

The concrete examples we solved can be found in Section 4.6.

**Applications** Let us discuss some possible applications of this work. In the following applications, one should carefully show that the qualification condition (presented above) holds, in particular when the transition probabilities are constrained.

- A direct application of this work concerns mean field game planning problems. In those problems, we are interested in moving a crowd of identical and rational agents from an initial spatial density to a desired target density  $m_T \in \Delta(S)$  at a

final time horizon  $T > 0$ . The interested reader can refer to the studies [3, 69, 99] on the subject. The mean field game planning problem is given by ( $\partial$ -MFG) with  $F(T, m(T, \cdot)) := \chi(m(T, \cdot) - m_T(\cdot))$ . In this situation, the variable  $\gamma(T, \cdot)$  plays an incentive role to ensure the constraint to be satisfied at final time  $T$ .

In fact, one could also prescribe at any time the density  $m(t, \cdot)$  via a mapping  $F(t, m(t, \cdot)) := \chi(m(t, \cdot) - m_t(\cdot))$  with  $m_t \in \Delta(S)$  for any  $t > 0$ .

- The same idea applies for the price. Instead of considering the price as a function of the aggregated behaviors, one can prescribe a quantity  $\bar{D}(s)$  that should be satisfied by the aggregation of the behaviors  $\mathbf{A}w(s)$  for any  $s > 0$  (see [66] for an application to the electricity market). This condition can be encoded with the function  $\phi(s, \mathbf{A}w(s)) := \chi(\mathbf{A}w(s) - \bar{D}(s))$ . In this situation, the price variable  $P(s)$  only play an incentive role to ensure the constraint to be satisfied at time  $s$ .

### 2.2.3 Chapter 3: Generalized conditional gradient and learning in potential mean field games

**Framework and objectives** In this chapter, we study the generalized conditional gradient algorithm to solve potential mean field game problems. We consider the continuous and finite time framework formulated in [21], consisting of a Hamilton-Jacobi-Bellman equation, a Fokker-Planck equation, and other coupling equations. The generalized conditional gradient method is an extension of the conditional gradient method, also called Frank-Wolfe algorithm. We show that the generalized conditional gradient method can be interpreted as a learning procedure called fictitious play. More precisely, each step of the generalized conditional gradient method amounts to compute the best-response of the representative agent, for a predicted value of the coupling terms of the game. This perspective allows us to:

1. Borrow and apply classical tools from the conditional gradient theory and derive, under suitable assumptions, convergence rates for the potential cost, the different variables generated by the fictitious play algorithm, and the exploitability;
2. Show that the notion of exploitability from game theory is equivalent to a notion of primal-dual gap (called primal gap certificate in the Frank-Wolfe literature).

The main results of this chapter deal with the well-posedness of the generalized conditional gradient algorithm and its convergence to the solution of the problem. The well-posedness is established with the help of suitable regularity estimates for the Hamilton-Jacobi-Bellman equation and the Fokker-Planck equation. We show that for the learning sequence  $\delta_k = 2/(k + 2)$ , the potential cost converges in  $O(1/k)$ , the exploitability and the variables of the problem (distribution, congestion, price, value function and control terms) converge in  $O(1/\sqrt{k})$ , for specific norms.

**Comparison with the literature** Up to our knowledge, this is the first contribution to consider the generalized conditional gradient algorithm in the mean field games literature, in a PDE setting. However, in a discrete setting, the conditional gradient method has been studied in the recent reference [62]. In the latter reference, the link with the fictitious play algorithm is also established.

Similarly to [62], we use the standard convergence results of the conditional gradient method to prove that the potential cost converges at a rate  $O(1/k)$  and the exploitability at a rate  $O(1/\sqrt{k})$ , when  $\delta_k = 2/(k + 2)$ .

In comparison with [62], the main novelty of our work (besides the different analytical framework) is the proof of convergence of all variables of the game: the coupling terms (price and congestion), the distribution of the agents, and their value function, at a rate  $O(1/\sqrt{k})$ . A key tool for the proof of convergence is a kind of quadratic growth property satisfied by the potential cost, which itself follows from the (assumed) strong convexity of the running cost of the agents.

Let us mention that we also provide convergence rates for the case  $\delta_k = 1/(k+1)$  which is more standard in the fictitious play algorithm:  $O(\ln(k)/k)$  for the potential cost,  $O(\sqrt{\ln(k)/k})$  for the exploitability and the different variables of the game.

**The model** In this chapter, we follow the framework developed in [21]. Following the latter reference, there exists a unique classical solution to  $(m, v, u, \gamma, P)$  to

$$\left\{ \begin{array}{ll} \text{(i)} & \begin{cases} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + P] = \gamma, & (x, t) \in Q, \\ u(x, T) = g(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(ii)} & v = -\mathbf{H}_p[\nabla u + P], & (x, t) \in Q, \\ \text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, & (x, t) \in Q, \\ m(0, x) = m_0(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(iv)} & \gamma(x, t) = f(x, t, m(t)), & (x, t) \in Q, \\ \text{(v)} & P(t) = \phi\left(\int_{\mathbb{T}^d} v(x, t)m(x, t)\right), & t \in [0, T], \end{array} \right. \quad (\text{MFG})$$

where  $Q := \mathbb{T}^d \times [0, T]$ . The unknown  $(m, v, u, \gamma, P)$  is such that  $m(x, t) \in \mathbb{R}$ ,  $v(x, t) \in \mathbb{R}^d$ ,  $u(x, t) \in \mathbb{R}$ ,  $\gamma(x, t) \in \mathbb{R}$ , and  $P(t) \in \mathbb{R}^d$ , for any  $(x, t) \in Q$ . The equation (MFG,i) is a Hamilton-Jacobi-Bellman equation and describes the evolution of the value function as time goes backward. Equation (MFG,ii) defines the optimal control  $v$ , which is given by the gradient  $H_p$  of the Hamiltonian. Equation (MFG,iii) is a Fokker-Planck equation, describing the evolution of the state distribution of the agents. Equation (MFG,iv) defines the congestion  $\gamma$  and equation (MFG,v) the price  $P$ .

**Generalized conditional gradient** The generalized conditional gradient algorithm is a generalization of a method called conditional gradient or Frank-Wolfe algorithm, first developed in [59]. The method is designed to minimize convex objective function under a bounded and convex constraint: consider the optimization problem

$$\min_{x \in K} f(x) = f_1(x) + f_2(x),$$

where  $K$  is a convex and compact subset of  $\mathbb{R}^n$  of finite diameter  $D$ ,  $f_1$  is a (possibly non-smooth) convex function and  $f_2$  a continuous differentiable function with  $L$ -Lipschitz gradient. We consider the mapping  $h: K \times K \rightarrow \mathbb{R}$  defined by

$$h(x, y) = f_1(y) - f_1(x) + \langle \nabla f_2(x), y - x \rangle.$$

The mapping  $h$  can be seen as a first-order Taylor approximation of  $f(y) - f(x)$ , where only  $f_2$  is linearized. Let  $(\delta_k)_{k \in \mathbb{N}} \in [0, 1]$  a sequence of step sizes. The method generates iteratively two sequences  $(\bar{x}_k)_{k \in \mathbb{N}}$  and  $(x_k)_{k \in \mathbb{N}}$  in  $K$ . At iteration  $k$ ,  $\bar{x}_k$  is available and

$(x_k, \bar{x}_{k+1})$  is obtained as follows:

$$\begin{aligned} x_k &= \arg \min_{y \in K} h(\bar{x}_k, y), \\ \bar{x}_{k+1} &= (1 - \delta_k) \bar{x}_k + \delta_k x_k. \end{aligned}$$

The purpose of our study is to adapt the generalized conditional gradient method presented above to compute the solution of the mean field game problem. As mentioned in the Section 2.1.5, several methods are available to solve potential mean field games. Here we present a new approach, and exhibit its links with the fictitious play algorithm. In our context, the objective function to be minimized is the following potential cost

$$\inf_{(m,w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w) := \int_Q \tilde{\mathbf{L}}[m, w] dx dt + \int_0^T \left( \mathbf{F}[m] + \Phi \left[ \int_{\mathbb{T}^d} w \right] \right) dt + \int_{\mathbb{T}^d} gm(T) dx \quad (\tilde{\mathfrak{P}})$$

where  $\tilde{\mathcal{R}}$  is the set of all  $(m, w) \in W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$  such that

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot w &= 0, & (x, t) \in Q, \\ m(x, 0) &= m_0(x), & x \in \mathbb{T}^d. \end{aligned}$$

We define a linearized control problem (similar to the one defined in subsection 2.1.2, up to a change of variable)

$$\inf_{(m', w') \in \tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma, P}(m', w') = \int_Q \tilde{\mathbf{L}}[m', w'] dx dt + \underbrace{\int_Q \gamma m' dx dt + \int_0^T \langle P, w' \rangle dt}_{(\star)} + \int_{\mathbb{T}^d} gm'(T) dx,$$

for any  $(\gamma, P)$ . The key remark is that if  $\gamma(x, t) = f(x, t, m(t))$  and  $P(t) = \phi(t, Aw(t))$  for any  $(x, t) \in Q$ , then the  $(\star)$  part of the individual cost can be interpreted as a linearization of the potentials  $\mathbf{F}$  and  $\Phi$  at  $(m, w) \in W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$ . Then we design a generalized conditional gradient method based on this observation. We define the semi-linearized cost

$$h((m, w), (m', w')) = \tilde{\mathcal{Z}}_{\gamma, P}(m', w') - \tilde{\mathcal{Z}}_{\gamma, P}(m, w),$$

of the potential cost  $\tilde{\mathcal{J}}$ , where  $\gamma(x, t) = f(x, t, m(t))$  and  $P(t) = \phi(t, Aw(t))$  for any  $(x, t) \in Q$ . The generalized conditional gradient method goes as follows:

---

**Algorithm 1** Generalized conditional gradient

---

Choose  $(\bar{m}_0, \bar{w}_0) \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \bar{w}_0 \in \mathcal{C}^{1+\alpha, \alpha}(Q, \mathbb{R}^d)$  and choose a sequence  $(\delta_k)_{k \in \mathbb{N}} \in [0, 1]$ .

**for**  $0 \leq k < N$  **do**

    Find the solution  $(m_k, w_k)$  to

$$\min_{(m,w) \in \tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w)). \quad (2.10)$$

    Actualise  $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1 - \delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$ .

**end for**

**return**  $(\bar{m}_N, \bar{w}_N)$ .

---

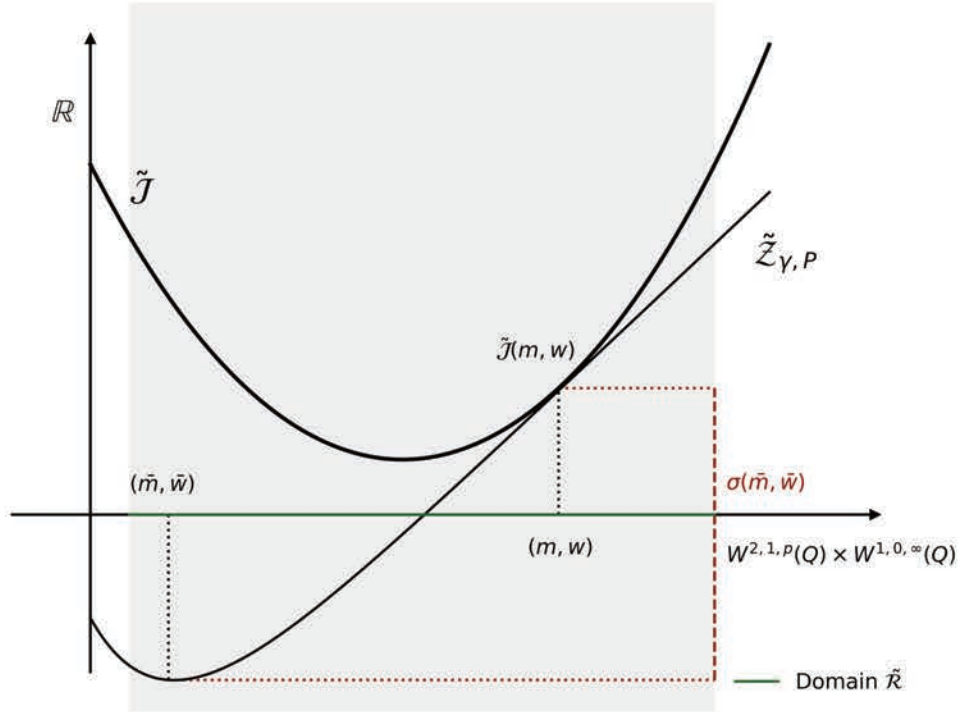


Figure 2.7: Illustration of the potential cost  $\tilde{J}$ , the individual cost  $\tilde{Z}_{\gamma,P}$  and the exploitability  $\sigma$ .

**Connection with the fictitious play method** As we explained in Subsection 2.1.2, the problem of minimization of  $\tilde{Z}_{\gamma,P}(\cdot)$  on  $\tilde{\mathcal{R}}$  is equivalent to the individual stochastic control problem (2.4). The quantity  $\tilde{Z}_{\gamma,P}(m,w)$  is the cost associated with the feedback  $v = w/m$ . Let  $\sigma_k$  be defined by

$$\sigma_k = - \min_{(m,w) \in \tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w)) = \tilde{Z}_{\gamma_k, P_k}(\bar{m}_k, \bar{w}_k) - \tilde{Z}_{\gamma_k, P_k}(m_k, w_k).$$

The real number  $\sigma_k$  is called exploitability: this is greatest reduction of cost that a representative agent can expect, assuming that the coupling terms  $\gamma_k$  and  $P_k$  remain constant. Let us mention that the notion of exploitability has received a growing attention in the mean field game and machine learning literatures [50, 51, 62, 101, 102, 103]. In particular, the exploitability is used as a metric to evaluate the convergence of learning schemes for mean field games.

We provide now an interpretation of the generalized gradient algorithm as a learning procedure called fictitious play. A definition and a study of the latter learning algorithm in the context of mean field games can be found in [35, 77]. Each iteration  $k$  of Algorithm 1 relies on the following steps:

1. Given  $(\bar{m}_k, \bar{w}_k)$  compute the coupling terms  $\gamma_k(x, t) = f(x, t, \bar{m}_k(t))$  and  $P_k(t) = \phi(t, \int_{\mathbb{T}^d} \bar{w}_k(t))$  for any  $(x, t) \in Q$ . In words, the agents make a prediction of the congestion term and the price at equilibrium, based on the belief  $(\bar{m}_k, \bar{w}_k)$ .
2. Find the value function  $u_k$  solution to the Hamilton-Jacobi-Bellman equation parametrized by  $(\gamma_k, P_k)$ ,

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + P_k] &= \gamma_k & (x, t) \in Q, \\ u(x, T) &= g(x) & x \in \mathbb{T}^d. \end{aligned}$$



Then compute the optimal control  $v_k = -\mathbf{H}_p[\nabla u_k + P_k]$ , given the value function  $u_k$  and the price  $P_k$ . This step can be interpreted as follows: for a given belief on the distributions of the coupling terms  $\gamma_k$  and  $P_k$ , a representative agent computes its best response  $v_k$ .

3. Find the solution  $m_k$  to the Fokker-Planck equation for the given drift  $v_k$

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (v_k m) &= 0, & (x, t) \in Q, \\ m(x, 0) &= m_0(x), & x \in \mathbb{T}^d, \end{aligned}$$

and compute the associated distribution of controls  $w_k = m_k v_k$ . The obtained pair  $(m_k, w_k)$  is the solution to (2.10).

4. The actualization step of  $(\bar{m}_{k+1}, \bar{w}_{k+1})$  can be interpreted as a learning step. The learning rule consists in averaging the past realizations of the distribution of states and controls at a rate determined by the sequence  $(\delta_k)_{k \in \mathbb{N}}$ .

**Methodology and main results** Besides the connection between generalized condition gradient and fictitious play, our main results deal with the convergence of the method. The first result concern the convergence of the primal cost and is almost a direct consequence of the convexity of the primal cost and the convexity and boundedness of the constraints set. The second result concern the convergence of the variables of the problem and relies on a sort of quadratic growth property satisfied by the potential cost, which itself followed from the (assumed) strong convexity of the running cost of the agents. The results presented here concern the case where the learning rate is  $\delta_k = 2/(k+2)$ .

- There exists  $C > 0$  such that  $\epsilon_k \leq C/k$ , where  $(\epsilon_k)_{k \in \mathbb{N}}$  is the sequence of primal gaps defined by

$$(\bar{m}, \bar{w}) = \arg \min_{(m, w) \in \bar{\mathcal{R}}} \tilde{\mathcal{J}}(m, w), \quad \epsilon_k = \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \tilde{\mathcal{J}}(\bar{m}, \bar{w}),$$

for all  $k \in \mathbb{N}$ .

- There exists  $C > 0$  such that  $\sigma(\bar{m}_k, \bar{w}_k) \leq C/\sqrt{k}$  and

$$\begin{aligned} \|\bar{v}_k - \bar{v}\|_{L^2(Q; \mathbb{R}^d)} + \|\bar{m}_k - \bar{m}\|_{L^\infty(0, T; L^2(\mathbb{T}^d))} + \|\bar{w}_k - \bar{w}\|_{L^2(Q; \mathbb{R}^d)} \\ + \|P_k - \bar{P}\|_{L^2(0, T; \mathbb{R}^k)} + \|\gamma_k - \bar{\gamma}\|_{L^\infty(Q)} + \|u_k - \bar{u}\|_{L^\infty(Q)} \leq C/\sqrt{k}, \end{aligned}$$

for all  $k \in \mathbb{N}$ . Here  $(\bar{u}, \bar{v}, \bar{m}, \bar{\gamma}, \bar{P})$  is the unique solution to (MFG) and  $\bar{w} = \bar{m}\bar{v}$ .

**Extensions** We now present possible research directions and extensions for this work:

1. Many improvements of the conditional gradient method have been proposed and investigated in the literature, see for example [81]. For example, one can replace the learning rule  $\delta_k = 2/(k+2)$  by a line search method. We could investigate the applicability of these variants to the generalized conditional gradient method in general and to potential mean field games in particular.
2. The connection between generalized condition gradient method and fictitious play could be established in different settings, in particular, for the Lagrangian formulation of first order mean field games. We expect that the convergence of the gap and the exploitability could be proved. Some convergence results have already been obtained in this setting in [35].



## Chapter 3

# Discrete-time mean field games with risk averse-agents

### 3.1 Introduction<sup>1</sup>

The class of mean field games problem was introduced by J-M. Lasry and P-L. Lions in [88, 89, 90] and M. Huang, R. Malhamé, and P. Caines in [80], to study interactions among a large population of players. Many developments and applications have been proposed this last decade, in particular in economics modeling and finance; one can refer for example to Y. Achdou and al. [1], O. Guéant, J-M. Lasry and P-L. Lions [73], and P. Cardaliaguet and C.-H. Lehalle [37]. Economic models "à la Cournot", considering interactions between the agents via a price variable, have recently received particular attention, let us mention the works of A. Bensoussan and P. J. Graber [67], J. F. Bonnans, S. Hadikanloo, and L. Pfeiffer [21], Z. Kobeissi [83], and P. J. Graber, V. Ignazio, and A. Neufeld [68].

The specificity of the mean field game of this article is the risk aversion of the involved agents. Here risk aversion is modeled with the help of composite risk measures (also called dynamic risk measures). Mathematically, a risk measure  $\rho$  is a map that assigns to a random variable  $U$  a real number, which is usually high when  $U$  is very volatile. In this way  $\rho$  can be used to model the reluctance of a player to face highly uncertain expenses. We refer to the seminal work by P. Artzner, F. Delbaen, J-M. Eber and D. Heath in [10]. We will make use of composite risk measures, the natural extension of risk measures to a multistage framework, see for example the article of A. Shapiro and A. Ruszczyński [111]; for an application to multistage portfolio selection one can refer to A. Shapiro [116].

Let us describe more precisely our coupled system and the obtained results. The coupled system describes a population of identical agents which all optimize a linear discrete-time dynamical system (in a continuous state space). In the model, the associated cost function depends on a variable called belief, which is related to the behavior of the whole group, whence a coupling between a single agent and the population. Assuming that the population is very large, one can consider that an isolated representative agent has no impact on the belief. Therefore his/her behavior can be conveniently described by dynamic programming equations (in which the belief is a parameter). Mathematically, the belief is the probability distribution of the states and controls of all agents at the different time steps of the game; it is described via the Kolmogorov equation. Our first result is an existence result, obtained with a standard fixed point approach.

---

<sup>1</sup>The corresponding article [23] was published in ESAIM: COCV.

In our second result, we show that an optimal feedback control for the mean field game yields an  $\varepsilon$ -Nash equilibrium for an  $N$ -player dynamic game, where  $\varepsilon \rightarrow 0$  as  $N \rightarrow \infty$ . The proof of this result is based on an estimate of the expectation of the Wasserstein distance between the empirical measure of i.i.d. variables and the law of these variables, obtained by N. Fournier and A. Guillin [58, Theorem 1]. The approach that we follow was proposed by M. Huang, P. Caines, and R. Malhamé in [79].

Discrete-time and continuous-space mean field game models have been studied in different works. The framework that we propose in this article is close to the one of N. Saldi, T. Başar and M. Raginsky [113], in particular, we make use of similar weighted spaces. A few works have already investigated the issue of risk aversion. Most of them model risk sensitivity via exponential utility functions, see for example H. Tembine, Q. Zhu and T. Başar [118]. The case of robust mean field games is investigated in problem (P2) in the work of J. Moon and T. Başar [95]. In many economic situations, risk modeling is of interest, in particular in the banking industry [98]. Our approach can also be relevant in situations where mean field games are used to design telecommunication systems or smart grids; see C. Bertucci et al. [19] and C. Alasseur, I. Ben Tahar and A. Matoussi [8]. For example, in the latter reference, it could be interesting to take into account the risk of individual no-energy situations or collective black-out situations via robust control.

The article is structured as follows. In Section 3.2 we introduce notations, assumptions, and the system of coupled equations. In Section 3.3 we interpret this system as a mean field game system with risk averse agents. In Section 3.4 we establish general technical results that will be helpful in Section 3.5, where we prove the existence of a solution to the coupled system. Finally in Section 3.6 we investigate the connection between the coupled system and an  $N$ -player game.

## 3.2 Problem Formulation

### 3.2.1 Notations

We set  $\mathcal{T} := \{0, \dots, T-1\}$  and  $\bar{\mathcal{T}} := \{0, \dots, T\}$  with  $T \in \mathbb{N}^*$ . For any  $t \in \bar{\mathcal{T}}$  and any vector  $(x_0, \dots, x_t)$  we denote

$$x_{[t]} := (x_0, \dots, x_t).$$

We denote

$$id : \mathbb{R}^d \rightarrow \mathbb{R}^d,$$

the identity mapping.

### Functions

Let  $C$ -Lip denote the set of Lipschitz functions of modulus  $C$  on  $\mathbb{R}^d$ . We define the  $p$ -polynomially weighted space

$$\mathcal{G}_p^C := \left\{ f : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}, |f(x)| \leq C(|x|^p + 1) \right\},$$

where the dimension  $d'$  depends on the context, with associated norm

$$\|f\|_{\mathcal{G},p} := \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{1 + |x|^p}.$$

Let  $\mathcal{Q}_p^C \subset \mathcal{G}_p^C$  denote the set of convex mappings  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfying

$$-C \leq f(x) \leq C(1 + |x|^p), \quad \forall x \in \mathbb{R}^d. \quad (3.1)$$

### Probability measures

Let  $\mathcal{P}(\mathbb{R}^d)$  denote the set of probability measures on  $\mathbb{R}^d$ . Given  $p \in [1, +\infty)$ , we define the set of finite  $p$ -th order moment measures

$$\mathcal{P}_p(\mathbb{R}^d) := \left\{ m \in \mathcal{P}(\mathbb{R}^d), \int_{\mathbb{R}^d} |x|^p dm(x) < +\infty \right\},$$

that we endow with the Rubinstein-Kantorovitch distance, defined by

$$d_1(\mu, \nu) := \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} \phi(x) d(\mu - \nu)(x),$$

for any  $\mu$  and  $\nu \in \mathcal{P}_1(\mathbb{R}^d)$  (see [120, Particular case 5.15] for more details). We recall that by the Hölder inequality,  $\mathcal{P}_p(\mathbb{R}^d) \subseteq \mathcal{P}_1(\mathbb{R}^d)$  for any  $p > 1$ . Given  $C > 0$ , we define

$$\mathcal{P}_p^C(\mathbb{R}^d) := \left\{ m \in \mathcal{P}_p(\mathbb{R}^d), \int_{\mathbb{R}^d} |x|^p dm(x) \leq C \right\}.$$

We also consider the following sets of beliefs

$$\mathcal{B}_2 := (\mathcal{P}_2(\mathbb{R}^{2d}))^T \times \mathcal{P}_2(\mathbb{R}^d), \quad \mathcal{B}_2^C := (\mathcal{P}_2^C(\mathbb{R}^{2d}))^T \times \mathcal{P}_2^C(\mathbb{R}^d),$$

endowed with the Rubinstein-Kantorovitch distances for the product topology, also denoted  $d_1$ .

For any  $m$  and  $\nu \in \mathcal{P}(\mathbb{R}^d)$ , we define the convolution product  $\nu * m$  by

$$\int_{\mathbb{R}^d} h(x) d(\nu * m)(x) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(y + z) d\nu(y) dm(z), \quad (3.2)$$

for any bounded Borel map  $h \in \mathbb{R}^d \rightarrow \mathbb{R}$ . For any  $m \in \mathcal{P}(\mathbb{R}^d)$  and for any Borel map  $g: \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$ , we define the image measure  $g\#m \in \mathcal{P}(\mathbb{R}^{d'})$  by

$$\int_{\mathbb{R}^d} (h \circ g)(x) dm(x) = \int_{\mathbb{R}^d} h(y) dg\#m(y), \quad (3.3)$$

for any bounded Borel map  $h \in \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$ .

### 3.2.2 Coupled system

Let us first introduce the data of the problem. We consider

- a congestion function  $F: \bar{\mathcal{T}} \times \mathbb{R}^d \times \mathcal{B}_2 \rightarrow \mathbb{R}$
- a price function  $P: \mathcal{T} \times \mathcal{B}_2 \rightarrow \mathbb{R}^d$
- an initial distribution  $\bar{m} \in \mathcal{P}_2(\mathbb{R}^d)$
- individual noise distributions  $(\nu(t))_{t \in \mathcal{T}} \in (\mathcal{P}_2(\mathbb{R}^d))^T$ .

The running cost  $\ell: \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{B}_2 \rightarrow \mathbb{R}$  is defined by

$$\ell(t, x, a, b) = \frac{1}{2}|a|^2 + \langle a, P(t, b) \rangle + F(t, x, b).$$

For modeling risk aversion, we consider a family of subsets  $(\mathcal{Z}_t)_{t \in \mathcal{T}}$  such that

$$\mathcal{Z}_t \subseteq \left\{ Z \in L^\infty(\mathbb{R}^d), \int_{\mathbb{R}^d} Z(y) d\nu(t, y) = 1, Z \geq 0 \right\}, \quad \forall t \in \mathcal{T}.$$

For any  $t \in \mathcal{T}$ , we define

$$\mathcal{M}_t := \left\{ \xi \in \mathcal{P}(\mathbb{R}^d), d\xi = Z d\nu(t), Z \in \mathcal{Z}_t \right\}. \quad (3.4)$$

For any  $t \in \mathcal{T}$ ,  $\mathcal{Z}_t$  is assumed to be nonempty and convex, thus  $\mathcal{M}_t$  is a nonempty and convex subset of  $\mathcal{P}(\mathbb{R}^d)$ .

We propose to study a *risk averse mean field game* (MFG), taking the form of the following coupled system:

$$\left\{ \begin{array}{l} \text{(i)} \quad \begin{cases} u(t, x) = \inf_{a \in \mathbb{R}^d} \left( \ell(t, x, a, b) + \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} u(t+1, x+a+y) d\xi(y) \right), \\ u(T, x) = F(T, x, b), \end{cases} \\ \text{(ii)} \quad \alpha_t(x) = \arg \min_{a \in \mathbb{R}^d} \left( \ell(t, x, a, b) + \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} u(t+1, x+a+y) d\xi(y) \right), \\ \text{(iii)} \quad \begin{cases} m(t+1) = \nu(t) * [(id + \alpha_t) \# m(t)], \\ m(0) = \bar{m}, \end{cases} \\ \text{(iv)} \quad \mu(t) = (id, \alpha_t) \# m(t), \\ \text{(v)} \quad b := (\mu(0), \dots, \mu(T-1), m(T)), \end{array} \right. \quad (\text{MFG})$$

for any  $(t, x) \in \mathcal{T} \times \mathbb{R}^d$ . The five unknowns in the above system are

- the value function  $u \in (\mathcal{G}_2)^{T+1}$
- the feedback control  $\alpha \in (\mathcal{G}_1 \cap 1\text{-Lip})^T$
- the distribution of states  $m \in (\mathcal{P}_2)^{T+1}$
- the joint distribution of states and controls  $\mu \in (\mathcal{P}_2(\mathbb{R}^{2d}))^T$
- the belief  $b \in \mathcal{B}_2$ .

Let us describe briefly the coupled system; we will justify it more in detail in Section 3.3. Equation (MFG,i) is a dynamic programming equation associated with a discrete-time optimal control problem for a representative agent. The belief  $b$  appears as a parameter of the equation, since a single agent has no impact on it. The corresponding optimal feedback control  $\alpha$  is then given by (MFG,ii). Now, assuming that all agents make use of the feedback control  $\alpha$ , the distribution of their state  $m$  is described by the Kolmogorov equation (MFG,iii) with initial condition  $\bar{m}$ .

Our approach for proving the existence of a solution consists in formulating the system (MFG) as a fixed point equation. For this purpose, we consider two mappings. The first one, that we call dynamic programming mapping, assigns to a belief  $b$  the solutions  $u^*(b)$  and  $\alpha^*(b)$  to equations (MFG,i) and (MFG,ii), respectively. The second one, the Kolmogorov mapping, assigns to a feedback control  $\alpha$  the triplet  $(m^*(\alpha), \mu^*(\alpha), b^*(\alpha))$ , where  $m^*(\alpha)$ ,  $\mu^*(\alpha)$ , and  $b^*(\alpha)$  are the solutions to (MFG,iii), (MFG,iv), and (MFG,v), respectively. These two mappings will be investigated in Section 3.5. They allow to reformulate the system (MFG) as an equivalent fixed point equation

$$b = b^* \circ \alpha^*(b).$$

### 3.2.3 Assumptions

We state now the assumptions on the data of the problem, in force all along the article. Note that for the results of Section 3.6 (dealing with the  $N$ -player dynamic game), we will need a slightly stronger assumption on the mapping  $F$ .

We make use of the same constant  $C$  to formulate the different assumptions. In the sequel, the constant  $C$  denotes a generic constant depending only on those involved in the assumptions and  $T$ ; its value can change from an inequality to the next one.

**Assumption 3.2.1.** *There exists  $C > 0$  such that  $\bar{m} \in \mathcal{P}_2^C(\mathbb{R}^d)$  and such that for any  $t \in \mathcal{T}$ ,  $\nu(t) \in \mathcal{P}_2^C(\mathbb{R}^d)$ .*

**Assumption 3.2.2.** *There exists  $C > 0$  such that for any  $t \in \mathcal{T}$  and for any  $Z \in \mathcal{Z}_t$ ,*

$$\|Z\|_\infty \leq C,$$

and there exists  $Z' \in \mathcal{Z}_t$  such that

$$Z' \geq \frac{1}{C} \quad \text{a.e.}$$

**Remark 3.2.3.** *Assumption 3.2.2 implies the existence of  $C > 0$  such that*

$$\mathcal{M}_t \subseteq \mathcal{P}_2^C(\mathbb{R}^d), \quad \forall t \in \mathcal{T}. \quad (3.5)$$

*The results obtained in Section 3.5 only require (3.5) to hold. The full Assumption 3.2.2 will be used in Section 3.6.*

**Assumption 3.2.4.** *There exists  $C > 0$  such that for any  $t \in \mathcal{T}$  and for any  $b_1$  and  $b_2 \in \mathcal{B}_2$ ,*

- (i)  $F(t, \cdot, b_1) \in \mathcal{Q}_2^C$ ,
- (ii)  $\|F(t, \cdot, b_1) - F(t, \cdot, b_2)\|_{\mathcal{G},2} \leq C d_1(b_1, b_2)$ ,
- (iii)  $|P(t, b_1) - P(t, b_2)| \leq C d_1(b_1, b_2)$ ,
- (iv)  $|P(t, b_1)| \leq C$ .

**Remark 3.2.5.** *In economics or in finance, prices typically depend on the aggregated demand or supply. One could consider for example*

$$P(t, b) := \psi \left( t, \int_{\mathbb{R}^{2d}} \alpha d\mu(t, x, \alpha) \right),$$

where  $\psi: \mathcal{T} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ . In this case, if  $\psi$  is a  $C$ -Lipschitz mapping then for any  $b_1$  and  $b_2 \in \mathcal{B}_2$ , one has that

$$|P(t, b_1) - P(t, b_2)| \leq C \left| \int_{\mathbb{R}^{2d}} \alpha d(\mu_1 - \mu_2)(t, x, \alpha) \right| \leq C d_1(\mu_1, \mu_2) \leq C d_1(b_1, b_2),$$

which implies Assumption 3.2.4 (iii). Assumption 3.2.4 (iv) also holds if  $|\psi| \leq C$ .

## 3.3 Interpretation of the coupled system

In Subsection 3.3.1 we describe the risk averse optimal control problem associated with (MFG,i-ii). In Subsection 3.3.2 we justify the Kolmogorov equation (MFG,iii).

### 3.3.1 Dynamic programming equation

#### Risk measures

Let  $X_0$  and  $(Y_t)_{t \in \mathcal{T}}$  be  $(T + 1)$ -independent random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $\mathcal{L}(X_0) = \bar{m}$  and  $\mathcal{L}(Y_t) = \nu(t)$ . We define the filtration  $(\mathcal{F}_t)_{t \in \mathcal{T}}$ , where  $\mathcal{F}_0 := \sigma(X_0)$  is the sigma-algebra generated by  $X_0$ , and  $\mathcal{F}_{t+1} := \sigma(X_0, Y_{[t]})$ . We denote for any  $t \in \mathcal{T}$  and any  $p \in [1, +\infty)$

$$\mathbb{L}_t^p(\Omega, \mathbb{R}^{d'}) := L^p(\Omega, \mathcal{F}_t, \mathbb{P}, \mathbb{R}^{d'}),$$

the space of  $\mathcal{F}_t$  measurable random variables with finite  $p$ -th order moment and value in  $\mathbb{R}^{d'}$ . When the dimension is  $d' = 1$ , we simplify the notation:  $\mathbb{L}_t^p := \mathbb{L}_t^p(\Omega, \mathbb{R})$ .

**Definition 3.3.1.** *Given  $t \in \mathcal{T}$ , we say that a mapping  $\rho_t: \mathbb{L}_{t+1}^1 \rightarrow \mathbb{L}_t^1$  is a one-step conditional risk mapping if it satisfies the following conditions:*

- **(M) Monotonicity:** *For any  $U$  and  $U' \in \mathbb{L}_{t+1}^1$  such that  $U \leq U'$ , we have*

$$\rho_t(U) \leq \rho_t(U'), \quad \text{a.s.}$$

- **(C) Convexity:** *For any  $U$  and  $U' \in \mathbb{L}_{t+1}^1$ , for any  $\alpha \in [0, 1]$ , we have*

$$\rho_t(\alpha U + (1 - \alpha)U') \leq \alpha \rho_t(U) + (1 - \alpha)\rho_t(U'), \quad \text{a.s.}$$

- **(TI) Translation Invariance:** *For any  $U \in \mathbb{L}_{t+1}^1$  and for any  $V \in \mathbb{L}_t^1$ , we have*

$$\rho_t(U + V) = \rho_t(U) + V, \quad \text{a.s.}$$

- **(PH) Positive Homogeneity:** *For any  $\alpha \geq 0$ , for any  $U \in \mathbb{L}_{t+1}^1$ , we have*

$$\rho_t(\alpha U) = \alpha \rho_t(U), \quad \text{a.s.}$$

Quoting [110], the conditional risk mapping  $\rho_t(U_{t+1})$  can be interpreted as a fair one-time  $\mathcal{F}_t$ -measurable charge we would be willing to incur at time  $t$  instead of the random futur cost  $U_{t+1}$ .

We fix now a family of *one-step conditional risk mapping*  $(\rho_t)_{t \in \mathcal{T}}$ ,  $\rho_t: \mathbb{L}_{t+1}^1 \rightarrow \mathbb{L}_t^1$ , defined by

$$\rho_t(U_{t+1})(x_0, y_{[t-1]}) = \sup_{Z \in \mathcal{Z}_t} \int_{\Omega} U_{t+1}(x_0, y_{[t-1]}, Y_t(\omega)) Z(Y_t(\omega)) d\mathbb{P}(\omega), \quad (3.6)$$

where the random variables  $U_{t+1}$  and  $\rho_t(U_{t+1})$  are explicitly represented as measurable functions of  $(x_0, y_{[t]}) \in \mathbb{R}^{(t+2)d}$  and  $(x_0, y_{[t-1]}) \in \mathbb{R}^{(t+1)d}$ , respectively. Recalling the definition of  $\mathcal{M}_t$  (3.4), we have

$$\rho_t(U_{t+1})(x_0, y_{[t-1]}) = \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} U_{t+1}(x_0, y_{[t-1]}, y_t) d\xi(y_t).$$

We set

$$\mathcal{Q}_{t+1} := \{Q = Z(Y_t) \text{ a.s., } Z \in \mathcal{Z}_t\}$$

so that  $\rho_t$  can be expressed in the following form:

$$\rho_t(U_{t+1}) = \sup_{Q_{t+1} \in \mathcal{Q}_{t+1}} \mathbb{E}[U_{t+1} Q_{t+1} | \mathcal{F}_t].$$

Finally we construct the associated composite risk measure  $\rho: \mathbb{L}_T^1 \rightarrow \mathbb{R}$ ,

$$\rho(U) := \mathbb{E}[\rho_0 \circ \dots \circ \rho_{T-1}(U)],$$

which also satisfies **(M)**, **(C)**, **(TI)**, and **(PH)**.



**Remark 3.3.2.** Given a probability space  $(\Omega', \mathcal{F}', \mathbb{P}')$  and given  $\alpha \in (0, 1]$ , the conditional value at risk (also called expected shortfall or average value at risk) of a random variable  $U \in L^1(\Omega', \mathcal{F}', \mathbb{P}')$  is defined by

$$\text{CV@R}_\alpha(U) := \inf_{W \in L^1(\Omega', \mathcal{F}', \mathbb{P}')} W + \alpha^{-1} \mathbb{E}[(U - W)_+],$$

where  $x_+ = \max\{0, x\}$  denotes the positive part of any  $x \in \mathbb{R}$ . It has the following dual representation (see [57, Lemma 4.51 and Theorem 4.52]):

$$\text{CV@R}_\alpha(U) = \sup \left\{ \mathbb{E}[UZ] \mid Z \in L^\infty(\Omega', \mathbb{F}', \mathbb{P}'), Z \in [0, \alpha^{-1}] \text{ a.s.}, \mathbb{E}[Z] = 1 \right\}.$$

Therefore, a natural extension of the conditional value at risk to the framework of the article is given by

$$\rho_t(U_{t+1}) = \sup_{Z \in \mathcal{Z}_t} \mathbb{E}[U_{t+1}Z(Y_t)|\mathcal{F}_t],$$

where

$$\mathcal{Z}_t := \left\{ Z \in L^\infty(\mathbb{R}^d) \mid Z \in [0, \alpha^{-1}] \text{ a.e.}, \int_{\mathbb{R}^d} Z(y) d\nu(t, y) = 1 \right\}.$$

This particular definition of  $\mathcal{Z}_t$  satisfies Assumption 3.2.2. We refer to [57, Definition 11.8] and [47, Subsection 2.3.1] for extensions of the conditional value at risk to general filtrations in a discrete-time setting.

**Remark 3.3.3.** The risk measure that we have constructed does not have the most general structure possible. In our setting, the sets  $\mathcal{M}_t$  are fixed. In [110], these sets depend on the current state and control (see in particular Sections 4 and 5). In this more general context, it is still possible to derive a dynamic programming principle for the underlying optimal control problem (see [110, Theorem 2]). However, the convexity of the value function, which plays an important role in our analysis, is lost in such a setting.

### Control problem

We consider the following set of controls for any  $t \in \mathcal{T}$ ,

$$\mathcal{A}_t = \mathbb{L}_t^2(\Omega, \mathbb{R}^d), \quad \mathcal{A} := \mathcal{A}_0 \times \cdots \times \mathcal{A}_{T-1}.$$

Given a control  $A \in \mathcal{A}$ , the evolution of the state of the representative player is given by

$$X_{t+1} = X_t + A_t + Y_t, \quad \forall t \in \mathcal{T}. \quad (C)$$

The initial condition is the random variable  $X_0$  fixed previously. We call the variable  $(X_t)_{t \in \bar{\mathcal{T}}}$  associated state with  $A$ . In the notation, we do not make explicit the dependence of  $(X_t)_{t \in \bar{\mathcal{T}}}$  with respect to  $A$ , which is always clear from the context. Note that by induction,  $X_t \in \mathbb{L}_t^2(\Omega, \mathbb{R}^d)$  for any  $t \in \bar{\mathcal{T}}$ .

For a given belief  $b \in \mathcal{B}_2$ , the risk averse multistage cost of the representative agent is given by

$$\mathcal{J}(A, b) := \rho \left( \sum_{t=0}^{T-1} \ell(t, X_t, A_t, b) + F(T, X_T, b) \right). \quad (3.7)$$

The corresponding problem is

$$\inf_{A \in \mathcal{A}} \mathcal{J}(A, b). \quad (P)$$

In what follows, we show how equations (MFG,i) and (MFG,ii) allow to characterize the unique solution to (P). Let us recall that  $b$  is fixed in this subsection. Let us denote by  $u \in (\mathcal{G}_2)^{T+1}$  the solution to (MFG,i) and let us denote by  $\alpha \in (\mathcal{G}_1 \cap 1\text{-Lip})^T$  the solution to (MFG,ii). The existence and uniqueness of these solutions will be independently established in Lemma 3.5.1 and Lemma 3.5.2.

**Lemma 3.3.4.** *There exists a unique control  $\bar{A} \in \mathcal{A}$  with associated state  $\bar{X}$  such that for all  $t \in \mathcal{T}$ ,*

$$\bar{A}_t = \alpha_t(\bar{X}_t), \quad a.s. \quad (3.8)$$

*Proof.* Let  $(\bar{X}_t)_{t \in \bar{\mathcal{T}}}$  be the solution to the closed-loop system

$$\bar{X}_{t+1} = \bar{X}_t + \alpha_t(\bar{X}_t) + Y_t, \quad \forall t \in \mathcal{T}. \quad (3.9)$$

It is easy to verify by induction that for all  $t \in \bar{\mathcal{T}}$ , the random variable  $\bar{X}_t$  is  $\mathcal{F}_t$ -measurable and has a bounded second-order moment. Indeed,  $\alpha_t$  is Lipschitz-continuous, thus has a linear growth; therefore, if  $\bar{X}_t$  has a bounded second-order moment, then  $\alpha_t(\bar{X}_t)$  also has a bounded second-order moment. We define now  $\bar{A}$  by

$$\bar{A}_t = \alpha_t(\bar{X}_t). \quad (3.10)$$

Since  $\bar{X}_t$  is adapted to  $\mathcal{F}_t$ , we also have that  $\bar{A}_t$  is  $\mathcal{F}_t$ -measurable. As we already pointed out,  $\alpha_t(\bar{X}_t)$  has a bounded second-order moment. This proves that  $\bar{A} \in \mathcal{A}$ . Finally, it is clear that by (3.9) and (3.10), the pair  $(\bar{A}, \bar{X})$  satisfies the state equation (C).

Let us justify the uniqueness of  $\bar{A}$ . Let  $\tilde{A} \in \mathcal{A}$  be such that  $\tilde{A}_t = \alpha_t(\tilde{X}_t)$ , where  $\tilde{X}$  is the associated state. Then,  $\tilde{X}$  is a solution to the closed-loop system (3.9). Therefore  $\tilde{X} = \bar{X}$  and finally  $\tilde{A}_t = \alpha_t(\tilde{X}_t) = \alpha_t(\bar{X}_t) = \bar{A}_t$ . The lemma is proved.  $\square$

The following proposition states the optimality of the control  $\bar{A}$ .

**Proposition 3.3.5.** *We have*

$$\inf_{A \in \mathcal{A}} \mathcal{J}(A, b) = \mathbb{E}[u(0, X_0)] = \int_{\mathbb{R}^d} u(0, x) dm(0, x), \quad (3.11)$$

where  $u$  solves the dynamic programming equation (MFG,i). Moreover, the control  $\bar{A}$  defined in Lemma 3.3.4 is the unique solution to Problem (P).

*Proof.* The proof is directly adapted from [110, Theorem 2]. As a consequence of the translation invariance property (TI), the problem (P) can be expressed in a nested form

$$\begin{aligned} \inf_{A \in \mathcal{A}} \mathcal{J}(A, b) = & \mathbb{E} \left[ \inf_{A_0 \in \mathcal{A}_0} \ell(0, X_0, A_0, b) + \rho_0 \left( \inf_{A_1 \in \mathcal{A}_1} \ell(1, X_1, A_1, b) + \dots \right. \right. \\ & \left. \left. + \rho_{T-2} \left( \inf_{A_{T-1} \in \mathcal{A}_{T-1}} \ell(T-1, X_{T-1}, A_{T-1}, b) + \rho_{T-1} \left( F(T, X_T, b) \right) \right) \dots \right) \right]. \end{aligned} \quad (3.12)$$

By (MFG,i), we have  $u(T, X_T) = F(T, X_T, b)$  almost surely. We also have  $X_T = X_{T-1} + A_{T-1} + Y_{T-1}$ , as a consequence of the state equation (C). Therefore, the innermost subproblem in (3.12) is given by

$$\inf_{A_{T-1} \in \mathcal{A}_{T-1}} \ell(T-1, X_{T-1}, A_{T-1}, b) + \rho_{T-1}(u(T, X_{T-1} + A_{T-1} + Y_{T-1})). \quad (3.13)$$

Since  $X_{T-1}, A_{T-1} \in \mathcal{F}_{T-1}$ , the unique solution to subproblem (3.13) is  $A_{T-1} = \alpha_{T-1}(X_{T-1})$ . Moreover, the value of subproblem (3.13) is  $u(T-1, X_{T-1})$ . Proceeding iteratively for all times  $t \in \mathcal{T}$ , we conclude that (3.11) holds and that any solution  $A$  to problem (P) with associated state  $X$  satisfies  $A_t = \alpha_t(X_t)$ . Therefore, by Lemma 3.3.4,  $\bar{A}$  is the unique solution to (P). The proof is complete.  $\square$

### 3.3.2 Kolmogorov equation

**Lemma 3.3.6.** *Let  $\alpha: \mathcal{T} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a continuous vector field. Suppose that the state equation (C) is of the feedback form*

$$X_{t+1} = X_t + \alpha_t(X_t) + Y_t.$$

*Then for any  $t \in \bar{\mathcal{T}}$ ,  $m(t) = \mathcal{L}(X_t) \in \mathcal{P}(\mathbb{R}^d)$  is characterized by the Kolmogorov equation (MFG,iv).*

*Proof.* Let  $\phi$  be a bounded Borel test function. For any  $t \in \mathcal{T}$ , by independence of  $X_t$  and  $Y_t$  we have

$$\begin{aligned} \mathbb{E}[\phi(X_{t+1})] &= \mathbb{E}[\phi(X_t + \alpha_t(X_t) + Y_t)] \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x + \alpha_t(x) + y) dm(t, x) d\nu(t, y). \end{aligned}$$

By definition of the push-forward (3.3) we obtain

$$\int_{\mathbb{R}^d} \phi(x + \alpha_t(x) + y) dm(t, x) = \int_{\mathbb{R}^d} \phi(z + y) d(id + \alpha_t)\#m(t, z).$$

By definition of convolution (3.2) we have

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(z + y) d\nu(t, y) d(id + \alpha_t)\#m(t, z) = \int_{\mathbb{R}^d} \phi(x) d(\nu(t) * [(id + \alpha_t)\#m(t)])(x),$$

as was to be proved.  $\square$

## 3.4 Technical lemmas

This section contains independent technical lemmas. The reader only interested in the main results of the article can skip it.

**Lemma 3.4.1.** *Let  $p \in [1, +\infty)$  and let  $C > 0$ . For any  $m_1$  and  $m_2$  in  $\mathcal{P}_p^C(\mathbb{R}^d)$ , the probability measure  $m_1 * m_2$  lies in  $\mathcal{P}_p^{2pC}(\mathbb{R}^d)$ . In addition, given  $m_0 \in \mathcal{P}_p^C(\mathbb{R}^d)$ , the mapping  $\mathcal{P}_p^C(\mathbb{R}^d) \ni m \mapsto m_0 * m$  is non-expansive for the distance  $d_1$ .*

*Proof.* Let  $m_1$  and  $m_2$  in  $\mathcal{P}_p^C(\mathbb{R}^d)$ . We have

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^p d(m_1 * m_2)(x) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |y + z|^p dm_1(y) dm_2(z) \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 2^{p-1}(|y|^p + |z|^p) dm_1(y) dm_2(z) \leq 2^p C. \end{aligned}$$

Thus  $m_1 * m_2 \in \mathcal{P}_p^{2pC}(\mathbb{R}^d)$ . Moreover, given  $m_0 \in \mathcal{P}_p^C(\mathbb{R}^d)$ , we have

$$\begin{aligned} d_1(m_0 * m_1, m_0 * m_2) &= \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} \phi(x) d(m_0 * m_1 - m_0 * m_2)(x) \\ &= \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \phi(y + z) dm_0(y) \right) d(m_1 - m_2)(z). \end{aligned}$$

Since the mapping  $z \mapsto \int_{\mathbb{R}^d} \phi(y + z) dm_0(y)$  is non-expansive, we further obtain that

$$d_1(m_0 * m_1, m_0 * m_2) \leq d_1(m_1, m_2),$$

which concludes the proof.  $\square$

**Lemma 3.4.2.** *Let  $p \in [1, +\infty)$  and let  $C > 0$ . For any  $m \in \mathcal{P}_p^C(\mathbb{R}^d)$  and for any Borel map  $g \in \mathcal{G}_1^C$ , the probability measure  $g\#m$  lies in  $\mathcal{P}_p^q(\mathbb{R}^d)$ , with  $q = 2^{p-1}C^p(1+C)$ . In addition, the inequality*

$$d_1(g_1\#m_1, g_2\#m_2) \leq (1+C)\|g_1 - g_2\|_{\mathcal{G},1} + Cd_1(m_1, m_2) \quad (3.14)$$

holds for any  $m_1$  and  $m_2$  in  $\mathcal{P}_p^C(\mathbb{R}^d)$  and for any Borel maps  $g_1$  and  $g_2$  in  $\mathcal{G}_1^C \cap C\text{-Lip}$ .

*Proof.* Let  $m \in \mathcal{P}_p^C(\mathbb{R}^d)$  and let  $g \in \mathcal{G}_1^C$  be a Borel map. By (3.3) we have

$$\int_{\mathbb{R}^d} |x|^p dg\#m(x) = \int_{\mathbb{R}^d} |g(x)|^p dm(x) \leq \|g\|_{\mathcal{G},1}^p \int_{\mathbb{R}^d} (1+|x|)^p dm(x) \leq q.$$

Consider  $(g_1, m_1)$  and  $(g_2, m_2)$  in  $\mathcal{G}_1^C \times \mathcal{P}_p^C(\mathbb{R}^d)$ . We have

$$\begin{aligned} d_1(g_1\#m_1, g_2\#m_2) &= \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} \phi(x) d(g_1\#m_1 - g_2\#m_2)(x) \\ &= \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} (\phi \circ g_1(x) - \phi \circ g_2(x)) dm_2(x) + \int_{\mathbb{R}^d} \phi \circ g_1(x) d(m_1 - m_2)(x) \\ &\leq \|g_1 - g_2\|_{\mathcal{G},1} \int_{\mathbb{R}^d} (1+|x|) dm_2(x) + \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} \phi \circ g_1(x) d(m_1 - m_2)(x) \\ &\leq (1+C)\|g_1 - g_2\|_{\mathcal{G},1} + C \sup_{\phi \in 1\text{-Lip}} \int_{\mathbb{R}^d} C^{-1}\phi \circ g_1(x) d(m_1 - m_2)(x). \end{aligned}$$

Observing that  $C^{-1}\phi \circ g_1 \in 1\text{-Lip}$ , we deduce inequality (3.14).  $\square$

Given a convex function  $u: \mathbb{R}^d \rightarrow \mathbb{R}$ , we define the Moreau envelope  $V_u$  and the proximal operator  $\text{prox}_u$  of  $u$  as follows:

$$V_u(x) := \min_{y \in \mathbb{R}^d} \frac{1}{2}|x - y|^2 + u(y), \quad \text{prox}_u(x) := \arg \min_{y \in \mathbb{R}^d} \frac{1}{2}|x - y|^2 + u(y). \quad (3.15)$$

In the proofs, we will occasionally consider the map  $g_u: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ , defined by

$$g_u(x, y) := \frac{1}{2}|x - y|^2 + u(y).$$

**Proposition 3.4.3.** *Let  $u: \mathbb{R}^d \rightarrow \mathbb{R}$  be a convex function. Then  $\text{prox}_u$  and  $(\text{id} - \text{prox}_u)$  are non-expansive.*

*Proof.* Direct consequence of [48, Proposition 12.27].  $\square$

**Lemma 3.4.4.** *Let  $R > 0$  and let  $u \in \mathcal{Q}_2^R$  (the set was defined in (3.1)). Then  $|\text{prox}_u|^2 \in \mathcal{G}_2^{C_1(R)}$  and  $|\text{prox}_u| \in \mathcal{G}_1^{(C_1(R))^{1/2}}$ , where  $C_1(R) := 8R + 2$ .*

*Proof.* Let  $u \in \mathcal{Q}_2^R$ . By Proposition 3.4.3, the map  $\text{prox}_u$  is non-expansive. Thus

$$|\text{prox}_u(x)| \leq |\text{prox}_u(0)| + |x|. \quad (3.16)$$

In addition, from the definition of the proximal operator (3.15), we have

$$\frac{1}{2}|\text{prox}_u(0)|^2 + u(\text{prox}_u(0)) \leq u(0).$$

Since  $u \in \mathcal{Q}_2^R$ , we deduce that  $|\text{prox}_u(0)|^2 \leq 4R$ . We further obtain with (3.16) that

$$|\text{prox}_u(x)|^2 \leq 2(|x|^2 + |\text{prox}_u(0)|^2) \leq C_1(R)(1 + |x|^2), \quad (3.17)$$

as was to be proved. Taking the square root of (3.17), we infer that  $|\text{prox}_u| \in \mathcal{G}_1^{C_1(R)^{1/2}}$ .  $\square$

**Lemma 3.4.5.** *Let  $R > 0$  and let  $u \in \mathcal{Q}_2^R$ . Then  $V_u \in \mathcal{Q}_2^{C_2(R)}$ , where*

$$C_2(R) := (R + 1)(1 + C_1(R)).$$

*Proof.* Let  $u \in \mathcal{Q}_2^R$ . Clearly  $V_u$  is convex as the infimum with respect to  $y \in \mathbb{R}^d$  of the jointly convex map  $(x, y) \mapsto g_u(x, y)$ . For any  $x \in \mathbb{R}^d$ , we have

$$V_u(x) = \frac{1}{2}|x - \text{prox}_u(x)|^2 + u(\text{prox}_u(x)),$$

by definition of  $V_u$  and  $\text{prox}_u$ . Since  $u \in \mathcal{Q}_2^R$ , we further obtain that

$$-R \leq V_u(x) \leq |x|^2 + |\text{prox}_u(x)|^2 + R(1 + |\text{prox}_u(x)|^2).$$

Applying Lemma 3.4.4, we finally obtain that  $V_u \in \mathcal{Q}_2^{C_2(R)}$ .  $\square$

**Lemma 3.4.6.** *Let  $R > 0$ . For any  $u$  and  $v$  in  $\mathcal{Q}_2^R$ , the inequality*

$$\|\text{prox}_u - \text{prox}_v\|_{\mathcal{G},1} \leq C_3(R)\|u - v\|_{\mathcal{G},2}^{1/2} \quad (3.18)$$

*holds, where  $C_3(R) := \sqrt{2(1 + C_1(R))}$ .*

*Proof.* Let  $u$  and  $v$  in  $\mathcal{Q}_2^R$ . Observing that  $g_u$  and  $g_v$  are 1-strongly convex with respect to their second argument, we have

$$\begin{aligned} \frac{1}{2}|\text{prox}_u(x) - \text{prox}_v(x)|^2 &\leq g_u(x, \text{prox}_v(x)) - g_u(x, \text{prox}_u(x)), \\ \frac{1}{2}|\text{prox}_u(x) - \text{prox}_v(x)|^2 &\leq g_v(x, \text{prox}_u(x)) - g_v(x, \text{prox}_v(x)). \end{aligned}$$

Summing up the two inequalities, we obtain that

$$\begin{aligned} |\text{prox}_u(x) - \text{prox}_v(x)|^2 &\leq v(\text{prox}_u(x)) - u(\text{prox}_u(x)) + u(\text{prox}_v(x)) - v(\text{prox}_v(x)) \\ &\leq (2 + |\text{prox}_u(x)|^2 + |\text{prox}_v(x)|^2)\|u - v\|_{\mathcal{G},2}. \end{aligned} \quad (3.19)$$

By Lemma 3.4.4,

$$2 + |\text{prox}_u(x)|^2 + |\text{prox}_v(x)|^2 \leq 2 + 2C_1(R)(1 + |x|^2) \leq C_3(R)^2(1 + |x|^2). \quad (3.20)$$

Combining (3.19) and (3.20) and taking the square root, we obtain (3.18).  $\square$

**Lemma 3.4.7.** *Let  $R > 0$ . For any  $u$  and  $v$  in  $\mathcal{Q}_2^R$ , we have*

$$\|V_u - V_v\|_{\mathcal{G},2} \leq C_4(R)\|u - v\|_{\mathcal{G},2}, \quad (3.21)$$

*where  $C_4(R) := 1 + C_1(R)$ .*

*Proof.* Let  $u$  and  $v$  in  $\mathcal{Q}_2^R$ . Recalling the definitions of  $g_u$  and  $g_v$ , we have

$$V_u(x) - V_v(x) \leq g_u(x, \text{prox}_v(x)) - g_v(x, \text{prox}_v(x)) = u(\text{prox}_v(x)) - v(\text{prox}_v(x)).$$

Lemma 3.4.4 yields

$$V_u(x) - V_v(x) \leq (1 + \|\text{prox}_v(x)\|_{\mathcal{G},1}^2)\|u - v\|_{\mathcal{G},2} \leq (1 + C_1(R)(1 + |x|^2))\|u - v\|_{\mathcal{G},2}.$$

Exchanging  $u$  and  $v$ , we deduce (3.21).  $\square$

**Lemma 3.4.8.** *Let  $R > 0$ . For any  $u \in \mathcal{Q}_2^R$  and for any  $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ ,*

$$|V_u(x) - V_u(y)| \leq C_5(R)(1 + |x| + |y|)|x - y|, \quad (3.22)$$

where  $C_5(R) := 1 + \sqrt{C_1(R)}$ .

*Proof.* Let  $u \in \mathcal{Q}_2^R$ . We have

$$\begin{aligned} V_u(x) - V_u(y) &\leq g_u(x, \text{prox}_u(y)) - g_u(y, \text{prox}_u(y)) \\ &= \frac{1}{2}|x - \text{prox}_u(y)|^2 - \frac{1}{2}|y - \text{prox}_u(y)|^2 \\ &\leq \frac{1}{2}|x + y - 2\text{prox}_u(y)| \cdot |x - y|. \end{aligned}$$

We further obtain with Lemma 3.4.4 that

$$|x + y - 2\text{prox}_u(y)| \leq |x| + |y| + 2\sqrt{C_1(R)}(1 + |y|) \leq 2(1 + \sqrt{C_1(R)})(1 + |x| + |y|).$$

Combining the two obtained inequalities and exchanging  $x$  and  $y$ , we obtain (3.22).  $\square$

**Lemma 3.4.9.** *Let  $R > 0$  and let  $\mathcal{M}$  be a subset of  $\mathcal{P}_2^R(\mathbb{R}^d)$ . Given  $u \in \mathcal{Q}_2^R$ , consider the mapping  $\Upsilon[u](x)$  defined for any  $x \in \mathbb{R}^d$  by*

$$\Upsilon[u](x) := \sup_{\xi \in \mathcal{M}} \int_{\mathbb{R}^d} u(x + y) d\xi(y).$$

Then  $\Upsilon[u] \in \mathcal{Q}_2^{C_6(R)}$ , where  $C_6(R) = 2R(1 + R)$ . Moreover, the map  $\mathcal{Q}_2^R \ni u \mapsto \Upsilon[u]$  is Lipschitz continuous with modulus  $2(1 + R)$ .

*Proof.* Let  $u \in \mathcal{Q}_2^R$ . For any  $\xi \in \mathcal{M}$ , the map  $\mathbb{R}^d \ni x \mapsto \int_{\mathbb{R}^d} u(x + y) d\xi(y)$  is convex, as can be easily verified. Thus  $\Upsilon[u](x)$  is convex with respect to  $x$ , as a supremum of convex maps. Moreover, for any  $x \in \mathbb{R}^d$ , we have

$$-R \leq \Upsilon[u](x) \leq \sup_{\xi \in \mathcal{M}} \int_{\mathbb{R}^d} 2R(1 + |x|^2 + |y|^2) d\xi(y) \leq 2R(1 + |x|^2 + R).$$

This proves that  $\Upsilon[u] \in \mathcal{Q}_2^{C_6(R)}$ . Consider now  $v \in \mathcal{Q}_2^R$ . We have

$$\begin{aligned} |\Upsilon[u](x) - \Upsilon[v](x)| &\leq \sup_{\xi \in \mathcal{M}} \left| \int_{\mathbb{R}^d} (u(x + y) - v(x + y)) d\xi(y) \right| \\ &\leq \|u - v\|_{\mathcal{G},2} \left( \sup_{\xi \in \mathcal{M}} \int_{\mathbb{R}^d} (1 + |x + y|^2) d\xi(y) \right). \end{aligned} \quad (3.23)$$

For any  $\xi \in \mathcal{M}$ , we further have

$$\int_{\mathbb{R}^d} (1 + |x + y|^2) d\xi(y) \leq 1 + 2|x|^2 + 2R \leq 2(1 + R)(1 + |x|^2). \quad (3.24)$$

Combining (3.23) and (3.24), we deduce that

$$\|\Upsilon[u] - \Upsilon[v]\|_{\mathcal{G},2} \leq 2(1 + R)\|u - v\|_{\mathcal{G},2},$$

as was to be proved.  $\square$

### 3.5 Existence result

In this section we prove the main existence result. We first investigate the continuity of the dynamic programming mapping and the continuity of the Kolmogorov mapping introduced in Subsection 3.2.2.

#### 3.5.1 Dynamic Programming mapping

In this section we show that for any given belief  $b \in \mathcal{B}_2$ , equations (MFG,i) and (MFG,ii) have unique solutions  $u$  and  $\alpha$ . We also investigate their dependence with respect to  $b$ . These equations can be equivalently formulated as follows, with an additional variable  $\bar{u} \in (\mathcal{G}_2)^{T+1}$ :

$$\bar{u}(t+1, x) = \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} u(t+1, x+y) d\xi(y), \quad (3.25)$$

$$u(t, x) = \inf_{a \in \mathbb{R}^d} \frac{1}{2} |a|^2 + \langle a, P(t, b) \rangle + F(t, x, b) + \bar{u}(t+1, x+a), \quad (3.26)$$

$$\alpha_t(x) = \operatorname{argmin}_{a \in \mathbb{R}^d} \frac{1}{2} |a|^2 + \langle a, P(t, b) \rangle + F(t, x, b) + \bar{u}(t+1, x+a), \quad (3.27)$$

$$u(T, x) = F(T, x, b), \quad (3.28)$$

for all  $t \in \mathcal{T}$  and for all  $x \in \mathbb{R}^d$ . The first step of our analysis consists in rewriting these equations in a functional form, with the help of the Moreau envelope and the proximal operator (introduced in (3.15)).

**Lemma 3.5.1.** *Let  $b \in \mathcal{B}_2$ . Let  $u \in (\mathcal{G}_2)^{T+1}$ , let  $\bar{u} \in (\mathcal{G}_2)^{T+1}$ , and let  $\alpha \in (\mathcal{G}_1 \cap 1\text{-Lip})^T$ . Then, for any  $t \in \mathcal{T}$  and for any  $x \in \mathbb{R}^d$ , equations (3.25)-(3.27) hold true if and only if*

$$\bar{u}(t+1, x) = \Upsilon[u(t+1, \cdot)](x), \quad (3.29)$$

$$u(t, x) = V_{\bar{u}(t+1, \cdot)}(x - P(t, b)) + F(t, x, b) - \frac{1}{2} |P(t, b)|^2, \quad (3.30)$$

$$\alpha_t(x) = (\operatorname{prox}_{\bar{u}(t+1, \cdot)} - \operatorname{id})(x - P(t, b)) - P(t, b). \quad (3.31)$$

*Proof.* Equality (3.29) is obviously equivalent to (3.25), by the definition of  $\Upsilon$ . By the change of variable  $y = x + a$ , the dynamic programming equation (3.26) can be reformulated as follows:

$$\begin{aligned} u(t, x) &= \inf_{y \in \mathbb{R}^d} \left( \frac{1}{2} |y - x|^2 + \langle (y - x), P(t, b) \rangle + F(t, x, b) + \bar{u}(t+1, y) \right) \\ &= \inf_{y \in \mathbb{R}^d} \left( \frac{1}{2} |y - (x - P(t, b))|^2 + \bar{u}(t+1, y) \right) + F(t, x, b) - \frac{1}{2} |P(t, b)|^2. \end{aligned} \quad (3.32)$$

This proves the equivalence between (3.26) and (3.30). Moreover, since  $\bar{u}(t+1, \cdot)$  is convex, the right-hand side of (3.32) has a unique minimizer given by

$$y^* := \operatorname{prox}_{\bar{u}(t+1, \cdot)}(x - P(t, b))$$

and therefore, the unique minimizer in the right-hand side of (3.27) is  $y^* - x$ , which proves the equivalence between (3.27) and (3.31). The lemma is proved.  $\square$

**Lemma 3.5.2.** *Let  $b \in \mathcal{B}_2$ . There exists a unique triplet  $(u, \bar{u}, \alpha) \in (\mathcal{G}_2)^{T+1} \times (\mathcal{G}_2)^{T+1} \times (\mathcal{G}_1 \cap 1\text{-Lip})^T$  such that (3.25)-(3.28) holds true. Moreover, for any  $t \in \bar{\mathcal{T}}$ , we have*

$$u(t, \cdot) \in \mathcal{Q}_2^{C_u}, \quad (3.33)$$

and for any  $t \in \mathcal{T}$ , we have

$$\alpha_t(\cdot) \in \mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip}, \quad (3.34)$$

for some positive constants  $C_\alpha$  and  $C_u$  independent of  $t$  and  $b$ .

*Proof.* Since  $u(T, \cdot)$  is uniquely defined by the terminal condition (3.28),  $\bar{u}(T, \cdot)$  is uniquely defined by (3.29) (with  $t = T - 1$ ). Then  $u(T - 1, \cdot)$  and  $\alpha_{T-1}(\cdot)$  are uniquely defined by (3.30) and (3.31) (with  $t = T - 1$ ) and so on, until  $t = 0$ .

Let us prove (3.33) by backward induction. The terminal condition  $u(T, \cdot) = F(T, \cdot, b)$  and Assumption 3.2.4 (i) imply that  $u(T, \cdot) \in \mathcal{Q}_2^C$ , for some constant  $C > 0$  (independent of  $b$ ). Let us take  $t \in \mathcal{T}$  and let us suppose that  $u(t + 1, \cdot) \in \mathcal{Q}_2^C$ . Then by Lemma 3.4.9 and relation 3.5, we have  $\bar{u}(t, \cdot) \in \mathcal{Q}_2^C$ . Recall that by Lemma 3.5.1, we have

$$u(t, \cdot) = V_{\bar{u}(t+1, \cdot)}(\cdot - P(t, b)) + F(t, \cdot, b) - \frac{1}{2}|P(t, b)|^2. \quad (3.35)$$

By Assumptions 3.2.4 (i) and (iv),  $F(t, \cdot, b) - \frac{1}{2}|P(t, b)|^2 \in \mathcal{Q}_2^C$ . Using again Assumption 3.2.4 (iv) and Lemma 3.4.5, we obtain that  $V_{\bar{u}(t+1, \cdot)}(\cdot - P(t, b)) \in \mathcal{Q}_2^C$ . Therefore, the right-hand side of (3.35) lies in  $\mathcal{Q}_2^C$  and finally,  $u(t, \cdot) \in \mathcal{Q}_2^C$ , where  $C$  is independent of  $b$ .

Let us prove (3.34). By Lemma 3.5.1, we have

$$\alpha_t(\cdot) = (\text{prox}_{\bar{u}(t+1, \cdot)} - \text{id})(\cdot - P(t, b)) - P(t, b). \quad (3.36)$$

We already know that  $\bar{u}(t + 1, \cdot) \in \mathcal{Q}_2^C$ . Moreover, by Assumption 3.2.4 (iv),  $P(t, b)$  is bounded. Therefore, by Lemma 3.4.4,  $\text{prox}_{\bar{u}(t+1, \cdot)}(\cdot - P(t, b)) \in \mathcal{G}_1^C$ . Then it is easy to show that  $\alpha_t(\cdot, b) \in \mathcal{G}_1^C$ , where again,  $C$  does not depend on  $b$ . Finally,  $\alpha(t, \cdot)$  is non-expansive as a consequence of (3.36) and Proposition 3.4.3. The lemma is proved.  $\square$

From now on, we denote by  $(u^*(\cdot, \cdot, b), \bar{u}^*(\cdot, \cdot, b), \alpha^*(\cdot, b))$  the unique solution to (MFG,i)-(MFG,ii).

**Lemma 3.5.3.** *There exists  $C > 0$  such that for any  $(t, b_1, b_2) \in \mathcal{T} \times \mathcal{B}_2 \times \mathcal{B}_2$ ,*

$$\|u^*(t, \cdot, b_1) - u^*(t, \cdot, b_2)\|_{\mathcal{G}, 2} \leq C d_1(b_1, b_2), \quad (3.37)$$

$$\|\bar{u}^*(t, \cdot, b_1) - \bar{u}^*(t, \cdot, b_2)\|_{\mathcal{G}, 2} \leq C d_1(b_1, b_2). \quad (3.38)$$

*Proof.* In the proof, all constants  $C$  are independent of  $b_1$  and  $b_2$ . We proceed by backward induction. By Assumption 3.2.4 (iii) and by the terminal condition  $u^*(T, \cdot, b) = F(T, \cdot, b)$ , inequality (3.37) holds true for  $t = T$ . Let  $t \in \mathcal{T}$ . Suppose that

$$\|u^*(t + 1, \cdot, b_1) - u^*(t + 1, \cdot, b_2)\|_{\mathcal{G}, 2} \leq C d_1(b_1, b_2),$$

for some positive constant  $C > 0$  independent of  $b_1$  and  $b_2$ . By Remark 3.2.3 and Lemma 3.4.9, we deduce that

$$\|\bar{u}^*(t + 1, \cdot, b_1) - \bar{u}^*(t + 1, \cdot, b_2)\|_{\mathcal{G}, 2} \leq C d_1(b_1, b_2). \quad (3.39)$$

By Lemma 3.5.1, we have

$$u^*(t, x, b_1) - u^*(t, x, b_2) = a_1(t, x, b_1, b_2) + a_2(t, x, b_1, b_2) + a_3(t, x, b_1, b_2), \quad (3.40)$$



where

$$\begin{aligned} a_1(t, x, b_1, b_2) &:= V_{\bar{u}^*(t+1, \cdot, b_1)}(x - P(t, b_1)) - V_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_1)), \\ a_2(t, x, b_1, b_2) &:= V_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_1)) - V_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_2)), \\ a_3(t, x, b_1, b_2) &:= F(t, x, b_1) - F(t, x, b_2) + \frac{1}{2}(|P(t, b_2)|^2 - |P(t, b_1)|^2). \end{aligned}$$

It remains to bound  $a_1(t, \cdot, b_1, b_2)$ ,  $a_2(t, \cdot, b_1, b_2)$ , and  $a_3(t, \cdot, b_1, b_2)$  in  $\mathcal{G}_2^C$ . We deduce from Lemma 3.4.7, Assumption 3.2.4 (iv), and estimate (3.39), that

$$\begin{aligned} |a_1(t, x, b_1, b_2)| &\leq \|V_{\bar{u}^*(t+1, \cdot, b_1)} - V_{\bar{u}^*(t+1, \cdot, b_2)}\|_{\mathcal{G}, 2}(1 + |x - P(t, b_1)|^2) \\ &\leq C\|\bar{u}^*(t+1, \cdot, b_1) - \bar{u}^*(t+1, \cdot, b_2)\|_{\mathcal{G}, 2}(1 + |x|^2) \\ &\leq Cd_1(b_1, b_2)(1 + |x|^2). \end{aligned}$$

Then by Lemma 3.4.8 and Assumption 3.2.4 (iv), we have

$$\begin{aligned} |a_2(t, x, b_1, b_2)| &\leq C(1 + |x - P(t, b_1)| + |x - P(t, b_2)|)|P(t, b_2) - P(t, b_1)| \\ &\leq C(1 + |x|)d_1(b_1, b_2) \\ &\leq C(1 + |x|^2)d_1(b_1, b_2). \end{aligned}$$

Finally by Assumption 3.2.4 (ii-iv), we have

$$\begin{aligned} |a_3(t, x, b_1, b_2)| &\leq \|F(t, \cdot, b_1) - F(t, \cdot, b_2)\|_{\mathcal{G}, 2}(1 + |x|^2) + C|P(t, b_1) - P(t, b_2)| \\ &\leq C(1 + |x|^2)d_1(b_1, b_2). \end{aligned}$$

Then combining (3.40) and the three estimates of  $a_1$ ,  $a_2$ , and  $a_3$ , we obtain that

$$\|u^*(t, \cdot, b_1) - u^*(t, \cdot, b_2)\|_{\mathcal{G}, 2} \leq Cd_1(b_1, b_2),$$

which concludes the proof.  $\square$

**Lemma 3.5.4.** *There exists  $C > 0$  such that for any  $(t, b_1, b_2) \in \mathcal{T} \times \mathcal{B}_2 \times \mathcal{B}_2$ ,*

$$\|\alpha_t^*(\cdot, b_1) - \alpha_t^*(\cdot, b_2)\|_{\mathcal{G}, 1} \leq C \left( d_1(b_1, b_2)^{1/2} + d_1(b_1, b_2) \right). \quad (3.41)$$

*Proof.* Let  $(t, b_1, b_2) \in \mathcal{T} \times \mathcal{B}_2 \times \mathcal{B}_2$ . By Lemma 3.5.1, we have

$$\alpha_t^*(x, b_1) - \alpha_t^*(x, b_2) = a_4(t, x, b_1, b_2) + a_5(t, x, b_1, b_2), \quad (3.42)$$

where

$$\begin{aligned} a_4(t, x, b_1, b_2) &= \text{prox}_{\bar{u}^*(t+1, \cdot, b_1)}(x - P(t, b_1)) - \text{prox}_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_1)), \\ a_5(t, x, b_1, b_2) &= \text{prox}_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_1)) - \text{prox}_{\bar{u}^*(t+1, \cdot, b_2)}(x - P(t, b_2)). \end{aligned}$$

Using successively Lemma 3.4.6, Assumption 3.2.4 (iv), and estimate (3.38), we obtain

$$\begin{aligned} |a_4(t, x, b_1, b_2)| &\leq \|\text{prox}_{\bar{u}^*(t+1, \cdot, b_1)} - \text{prox}_{\bar{u}^*(t+1, \cdot, b_2)}\|_{\mathcal{G}, 1}(1 + |x - P(t, b_1)|) \\ &\leq C\|\bar{u}^*(t+1, \cdot, b_1) - \bar{u}^*(t+1, \cdot, b_2)\|^{1/2}(1 + |x|) \\ &\leq C\|d_1(b_1, b_2)\|^{1/2}(1 + |x|). \end{aligned}$$

Moreover, since  $\text{prox}_{\bar{u}^*(t+1, \cdot, b_2)}$  is non-expansive, we have with Assumption 3.2.4 (iii) that

$$|a_5(t, x, b_1, b_2)| \leq |(x - P(t, b_1)) - (x - P(t, b_2))| \leq d_1(b_1, b_2).$$

Combining the two obtained estimates of  $a_4$  and  $a_5$  with (3.42), we obtain (3.41).  $\square$

### 3.5.2 Kolmogorov mapping

We study now the Kolmogorov mapping

$$(\mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip})^T \ni \alpha \mapsto (m^*, \mu^*, b^*)(\alpha),$$

where  $(m^*, \mu^*, b^*)$  is the solution to (MFG,iii-v).

**Lemma 3.5.5.** *There exists  $C_b > 0$  such that for any  $\alpha \in (\mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip})^T$ ,*

$$m^*(\alpha) \in (\mathcal{P}_2^{C_b}(\mathbb{R}^d))^{T+1}, \quad \mu^*(\alpha) \in (\mathcal{P}_2^{C_b}(\mathbb{R}^{2d}))^T, \quad \text{and} \quad b^*(\alpha) \in \mathcal{B}_2^{C_b}.$$

*In addition the three mappings  $m^*, \mu^*$  and  $b^*$  are continuous.*

*Proof.* Let  $\alpha \in (\mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip})^T$ . All constants  $C$  in the proof are independent of  $\alpha$ . Let us first prove by induction that for any  $t \in \bar{\mathcal{T}}$ , there exists a constant  $C > 0$  independent of  $\alpha$  such that  $m^*(t, \cdot, \alpha) \in \mathcal{P}_2^C(\mathbb{R}^d)$  and such that,  $m^*(t, \cdot, \alpha)$  is continuous with respect to  $\alpha$ . The claim is clear for  $t = 0$ , since  $m^*(0, \cdot, \alpha) = \bar{m} \in \mathcal{P}_2^C(\mathbb{R}^d)$ , by Assumption 3.2.1. Now, let us assume that the claim holds true for some  $t \in \mathcal{T}$ . We recall that

$$m^*(t+1, \cdot, \alpha) = \nu(t) * [(id + \alpha_t) \# m^*(t, \cdot, \alpha)].$$

Since  $\nu(t) \in \mathcal{P}_2^C(\mathbb{R}^d)$  (by Assumption 3.2.1) and since  $\alpha_t \in \mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip}$ , we obtain with Lemma 3.4.1 and Lemma 3.4.2 that  $m^*(t+1, \cdot, \alpha) \in \mathcal{P}_2^C(\mathbb{R}^d)$  and that  $m^*(t+1, \cdot, \alpha)$  is a continuous function of  $\alpha$ , by composition.

It remains to justify the boundedness of  $\mu^*$  and  $b^*$ . We recall that for any  $t \in \mathcal{T}$ ,

$$\mu^*(t, \cdot, \alpha) = (id, \alpha_t) \# m^*(t, \cdot, \alpha).$$

We deduce from Lemma 3.4.2 that  $\mu^*(t, \cdot, \alpha) \in \mathcal{P}_2^C(\mathbb{R}^{2d})$  and that  $\mu^*(t, \cdot, \alpha)$  is a continuous function of  $\alpha$ , by composition. It immediately follows that  $b^*(\alpha) \in \mathcal{B}_2^C$  and that  $b^*$  is continuous.  $\square$

### 3.5.3 Existence of equilibrium

We are ready to prove the existence of a solution of system (MFG). The proof relies on the Schauder fixed point theorem, that we first recall.

**Theorem 3.5.6.** (Schauder) *Let  $C$  be a convex and compact set in a Banach space  $X$ , and let  $T: C \rightarrow C$  be a continuous mapping. Then  $T$  has a fixed point, i.e. there exists  $x \in C$  such that*

$$T(x) = x.$$

**Theorem 3.5.7.** *There exists  $(u, \alpha, m, \mu, b) \in (\mathcal{G}_2^{C_u})^T \times (\mathcal{G}_1^{C_\alpha} \cap 1\text{-Lip})^T \times (\mathcal{P}_2^{C_b}(\mathbb{R}^d))^{T+1} \times (\mathcal{P}_2^{C_b}(\mathbb{R}^{2d}))^T \times \mathcal{B}_2^{C_b}$  solution to system (MFG), where  $C_u, C_\alpha$  and  $C_b$  are the constants obtained in Lemma 3.5.2 and Lemma 3.5.5.*

*Proof.* By Lemma 3.5.4 and Lemma 3.5.5, the mapping

$$\mathcal{B}_2^{C_b} \ni b \mapsto b^* \circ \alpha^*(b) \in \mathcal{B}_2^{C_b}$$

is continuous for the distance  $d_1$ . Moreover,  $\mathcal{B}_2^{C_b}$  is compact for  $d_1$ , see [104, Lemma 25]. Therefore, by the Schauder fixed point theorem, there exists  $\bar{b} \in \mathcal{B}_2^{C_b}$  such that  $\bar{b} = b^* \circ \alpha^*(\bar{b})$ . Let us set  $\bar{u} = u^*(\bar{b})$ ,  $\bar{\alpha} = \alpha^*(\bar{b})$ ,  $\bar{m} = m^*(\bar{\alpha})$ , and  $\bar{\mu} = \mu^*(\bar{\alpha})$ . Then  $(\bar{u}, \bar{\alpha}, \bar{m}, \bar{\mu}, \bar{b})$  is solution to (MFG) and lies in the announced set.  $\square$

### 3.6 Connection with a finite player game

In this section we establish a connection between the coupled system (MFG) and a dynamic game with  $N$  players. More precisely, we fix a solution  $(\bar{u}, \bar{\alpha}, \bar{m}, \bar{\mu}, \bar{b})$  of system (MFG) and consider the situation where each of the  $N$  players adopts the feedback  $\bar{\alpha}$ . We show that this situation is an  $\varepsilon$ -Nash equilibrium for the  $N$ -player game and we quantify the rate of convergence of  $\varepsilon$  to 0 as  $N$  goes to infinity.

To show this, the following restriction on Assumption 3.2.4 (ii) will be required, in particular to prove Lemma 3.6.12.

**Assumption 3.6.1.** *There exists  $C > 0$  such that for any  $t \in \mathcal{T}$  and for any  $b_1$  and  $b_2$  in  $\mathcal{B}_2$ ,*

- (i)  $F(t, \cdot, b_1) \in \mathcal{Q}_1^C$ ,
- (ii)  $\|F(t, \cdot, b_1) - F(t, \cdot, b_2)\|_{\mathcal{G},1} \leq C d_1(b_1, b_2)$ .

We have already fixed a solution to system (MFG), now we also fix the number of players  $N$ ; all constants  $C$  appearing in the sequel are independent of  $N$ .

#### 3.6.1 Formulation of the game

Let  $\mathcal{N} := \{1, \dots, N\}$  and let  $i \in \mathcal{N}$ . For any vector  $(x^1, \dots, x^N)$  we denote

$$\begin{aligned} \mathbf{x} &= (x^1, \dots, x^N), \\ \mathbf{x}^{-i} &= (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^N). \end{aligned}$$

Consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $(X_0^i)_{i \in \mathcal{N}}$  be i.i.d. random variables with law  $\mathcal{L}(X_0^i) = \bar{m}$ . Let  $(Y_t^i)_{i \in \mathcal{N}, t \in \mathcal{T}}$  be independent random variables, independent of  $(X_0^i)_{i \in \mathcal{N}}$ , with law  $\mathcal{L}(Y_t^i) = \nu(t)$ . We denote  $\nu(t) := \prod_{i=1}^N \nu(t)$ . We define the filtration  $(\mathcal{F}_t)_{t \in \bar{\mathcal{T}}}$  as follows:  $\mathcal{F}_0 := \sigma(\mathbf{X}_0)$  is the sigma-algebra generated by  $\mathbf{X}_0$ ,  $\mathcal{F}_{t+1} := \sigma(\mathbf{X}_0, \mathbf{Y}_{[t]})$ . In this section we denote

$$\mathbf{L}_t^p(\Omega, \mathbb{R}^{d'}) := L^p(\Omega, \mathcal{F}_t, \mathbb{P}, \mathbb{R}^{d'}),$$

the space of  $\mathcal{F}_t$  measurable random variables with finite  $p$ -th order moment and value in  $\mathbb{R}^{d'}$ . When the dimension is  $d' = 1$ , we simplify the notation  $\mathbf{L}_t^p = L_t^p(\Omega, \mathbb{R})$ . For any  $t \in \mathcal{T}$ , we consider the control set

$$\mathcal{A}_t := L_t^2(\Omega, \mathbb{R}^d), \quad \mathcal{A} := \mathcal{A}_0 \times \dots \times \mathcal{A}_{T-1}, \quad \mathcal{A}^N := \prod_{i=1}^N \mathcal{A}.$$

For any  $t \in \mathcal{T}$  and for any constant  $C > 0$  we denote  $\mathcal{A}_t^C$  the set of controls  $A \in \mathcal{A}_t$  such that

$$\int_{\Omega} |A(\omega)|^2 d\mathbb{P}(\omega) \leq C$$

and we set  $\mathcal{A}^C := \mathcal{A}_0^C \times \dots \times \mathcal{A}_{T-1}^C$ . The control of player  $i \in \mathcal{N}$  is an adapted stochastic process  $A^i \in \mathcal{A}$ , whose associated trajectory  $(X_t^i[A^i])_{t \in \bar{\mathcal{T}}}$  is defined by the following state equation

$$X_{t+1}^i = X_t^i + A_t^i + Y_t^i.$$

**Remark 3.6.2.** *Let  $R > 0$ . There exists  $C > 0$  (depending on  $R$ ) such that for any  $i \in \mathcal{N}$  and for any  $A^i \in \mathcal{A}^R$ ,  $\mathbb{E} [|X_t^i[A^i]|^2] \leq C$  for any  $t \in \bar{\mathcal{T}}$ , since  $\mathcal{L}(X_0^i) \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\mathcal{L}(Y_t^i) \in \mathcal{P}_2(\mathbb{R}^d)$ .*

Given  $\mathbf{A} \in \mathcal{A}^N$ , we define the random empirical measure of the positions and the random empirical joint measure of the positions and actions of players by

$$m_{\mathbf{A}}^N(t) := \frac{1}{N} \sum_{i \in \mathcal{N}} \delta_{X_t^i[A^i]}, \quad \mu_{\mathbf{A}}^N(t) := \frac{1}{N} \sum_{i \in \mathcal{N}} \delta_{(X_t^i[A^i], A_t^i)},$$

where  $\delta$  denotes the Dirac measure. We set

$$b_{\mathbf{A}}^N := (\mu_{\mathbf{A}}^N(0), \dots, \mu_{\mathbf{A}}^N(T-1), m_{\mathbf{A}}^N(T)).$$

For any  $i \in \mathcal{N}$  and for any  $t \in \mathcal{T}$ , we define the *individual conditional risk measure*  $\rho_t^i: \mathbf{L}_{t+1}^1 \rightarrow \mathbf{L}_t^1$ ,

$$\rho_t^i(U_{t+1})(\mathbf{x}_0, \mathbf{y}_{[t-1]}) = \sup_{Z \in \mathcal{Z}_t} \int_{\Omega} U_{t+1}(\mathbf{x}_0, \mathbf{y}_{[t-1]}, \mathbf{Y}_t(\omega)) Z(Y_t^i(\omega)) d\mathbb{P}(\omega).$$

We define the set

$$\mathbf{Q}_{t+1}^i := \{Q \in \mathbf{L}_{t+1}^\infty, Q = Z(Y_t^i) \text{ a.s.}, Z \in \mathcal{Z}_t\}.$$

Then  $\rho_t^i$  can be expressed in the following form:

$$\rho_t^i(U_{t+1}) = \sup_{Q_{t+1} \in \mathbf{Q}_{t+1}^i} \mathbb{E}[U_{t+1} Q_{t+1} | \mathcal{F}_t]. \quad (3.43)$$

In addition we have that

$$\rho_t^i(U_{t+1})(\mathbf{x}_0, \mathbf{y}_{[t-1]}) = \sup_{\xi \in \mathcal{M}_t} \int_{\mathbb{R}^d} \int_{\mathbb{R}^{Nd}} U_{t+1}(\mathbf{x}_0, \mathbf{y}_{[t]}) d\nu^{-i}(t, \mathbf{y}_t^{-i}) d\xi(y_t^i),$$

where  $\nu^{-i}(t) := \prod_{j \in \mathcal{N} \setminus \{i\}} \nu^j(t)$ . Then  $(\rho_t^i)_{t \in \mathcal{T}}$  is a family of conditional risk mappings. We define the associated individual composite risk measure  $\rho^i: \mathbf{L}_T^1 \rightarrow \mathbb{R}$ ,

$$\rho^i(U) := \mathbb{E}[\rho_0^i \circ \dots \circ \rho_{T-1}^i(U)].$$

Here players are risk averse with respect to their individual noise only. For any  $\mathbf{A} \in \mathcal{A}^N$  the cost of the player  $i \in \mathcal{N}$  is given by

$$\mathcal{J}^{i,N}(A^i, \mathbf{A}^{-i}) := \rho^i \left( \sum_{t=0}^{T-1} \ell(t, X_t^i[A^i], A_t^i, b_{\mathbf{A}}^N) + F(T, X_T^i[A^i], b_{\mathbf{A}}^N) \right).$$

**Definition 3.6.3.** Let  $\varepsilon \geq 0$ . We say that an  $N$ -uplet  $\widehat{\mathbf{A}} \in \mathcal{A}^N$  is an  $\varepsilon$ -Nash equilibrium for the  $N$ -player game if for any  $i \in \mathcal{N}$ ,

$$\mathcal{J}^{i,N}(\widehat{A}^i, \widehat{\mathbf{A}}^{-i}) \leq \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \widehat{\mathbf{A}}^{-i}) + \varepsilon. \quad (3.44)$$

For  $\varepsilon = 0$ , we recover the usual definition of a Nash equilibrium.

### 3.6.2 An approximate Nash equilibrium

For any player  $i \in \mathcal{N}$ , we denote by  $(\bar{X}_t^i)_{t \in \bar{\mathcal{T}}}$  the solution to the closed-loop system

$$X_{t+1}^i = X_t^i + \bar{\alpha}_t(X_t^i) + Y_t^i.$$

We define the control  $\bar{A}^i \in \mathcal{A}$  by

$$\bar{A}_t^i = \bar{\alpha}_t(\bar{X}_t^i). \quad (3.45)$$

Since  $\bar{X}_t^i$  is adapted to  $\mathcal{F}_t$ , the control  $\bar{A}_t^i$  is also  $\mathcal{F}_t$ -measurable. Moreover,  $\bar{\alpha}_t$  is 1-Lipschitz and the random variables  $X_0$  and  $(Y_t)_{t \in \mathcal{T}}$  have a bounded second-order moment, thus  $\bar{A}^i \in \mathcal{A}$ . In addition, by Proposition 3.3.5,  $\bar{A}^i$  minimizes the following cost  $\mathcal{J}^i$ :

$$\mathcal{J}^i(A^i, \bar{b}) := \rho^i \left( \sum_{t=0}^{T-1} \ell(t, X_t^i[A^i], A_t^i, \bar{b}) + F(T, X_T^i[A^i], \bar{b}) \right). \quad (3.46)$$

Finally we set  $\bar{\mathbf{A}} = (\bar{A}^1, \dots, \bar{A}^N)$ . The following result states that  $\bar{\mathbf{A}}$  is an  $\varepsilon$ -Nash equilibrium.

**Theorem 3.6.4.** *Let  $\xi \in (0, 1/2)$ . There exists a constant  $C > 0$ , independent of  $N$ , such that the  $N$ -uplet  $\bar{\mathbf{A}}$  defined above is an  $\varepsilon$ -Nash equilibrium with*

$$\varepsilon = CN^{-\tau(d)/2}, \quad \tau(d) = \begin{cases} 1/2 - \xi & \text{if } d \in \{1, 2\}, \\ 1/d & \text{if } d \geq 3. \end{cases}$$

In addition we have that

$$|\mathcal{J}^{i,N}(\bar{\mathbf{A}}) - \mathcal{J}^i(\bar{A}^i, \bar{b})| \leq CN^{-\tau(d)/2}, \quad (3.47)$$

for any  $i \in \mathcal{N}$ .

The proof of the theorem can be found at the end of Subsection 3.6.3 (page 64), which contains technical intermediate lemmas. They rely on the following result.

**Theorem 3.6.5.** *(Fournier-Guillin) Let  $c > 0$ , let  $\xi \in (0, 1/2)$ , and let  $\mu \in \mathcal{P}_2^c(\mathbb{R}^d)$ . Consider  $N$  i.i.d. random variables  $(X_i)_{i \in \{1, \dots, N\}}$  in  $\mathbb{R}^d$  with law  $\mu$  and denote by  $\mu_N$  their empirical measure, defined by*

$$\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}. \quad (3.48)$$

There exists a constant  $C > 0$  depending only on  $c$ ,  $d$ , and  $\xi$  such that

$$\mathbb{E}[d_1(\mu, \mu_N)] \leq CN^{-\tau(d)}.$$

*Proof.* The theorem is a direct application of [58, Theorem 1] with  $q = \frac{2}{1+2\xi}$  if  $d \in \{1, 2\}$  and  $q = 2$  if  $d \geq 3$ .  $\square$

### 3.6.3 Proof of Theorem 3.6.4

We begin with four technical lemmas dealing with the regularity of the individual risk measures  $\rho^i$ .

**Lemma 3.6.6.** *For any player  $i \in \mathcal{N}$  the risk measure  $\rho^i$  is subadditive, that is*

$$\rho^i(U + V) \leq \rho^i(U) + \rho^i(V),$$

for any  $U$  and  $V$  in  $\mathbf{L}_T^1$ .

*Proof.* Let us define  $\pi_T^i(U) = U$  and  $\pi_t^i(U) = \rho_t^i \circ \rho_{t+1}^i \dots \circ \rho_{T-1}^i(U)$ , for any  $U \in \mathbf{L}_T^1$ . Note that  $\pi_t^i = \rho_t^i \circ \pi_{t+1}^i$ , for any  $t \in \mathcal{T}$ . We prove by backward induction that  $\pi_t^i$  is subadditive for any  $t \in \mathcal{T}$ . The claim is trivial for  $t = T$ . Let  $t \in \mathcal{T}$ . Assume that  $\pi_{t+1}^i$  is subadditive, let us prove that  $\pi_t^i$  is subadditive. First we observe that for any  $U$  and  $V$  in  $\mathbf{L}_T^1$ ,

$$\begin{aligned} \rho_t^i(U + V) &= \sup_{Q \in \mathcal{Q}_{t+1}^i} \mathbb{E}[(U + V)Q | \mathcal{F}_t] \\ &\leq \sup_{Q \in \mathcal{Q}_{t+1}^i} \mathbb{E}[UQ | \mathcal{F}_t] + \sup_{Q \in \mathcal{Q}_{t+1}^i} \mathbb{E}[VQ | \mathcal{F}_t] = \rho_t^i(U) + \rho_t^i(V), \quad \text{a.s.} \end{aligned}$$

It follows with the monotonicity of  $\rho_t^i$  that

$$\begin{aligned} \pi_t^i(U + V) &= \rho_t^i \circ \pi_{t+1}^i(U + V) \\ &\leq \rho_t^i(\pi_{t+1}^i(U) + \pi_{t+1}^i(V)) \\ &\leq \rho_t^i \circ \pi_{t+1}^i(U) + \rho_t^i \circ \pi_{t+1}^i(V) = \pi_t^i(U) + \pi_t^i(V), \quad \text{a.s.} \end{aligned}$$

Recalling that  $\rho^i(U) = \mathbb{E}[\rho_0^i \circ \dots \circ \rho_{T-1}^i(U)] = \mathbb{E}[\pi_0(U)]$ , we conclude that  $\rho^i$  is also subadditive.  $\square$

The following result is close to a triangle inequality for risk measures. The difference with the triangle inequality is due to the positive homogeneity of risk measures, while norms are absolutely homogeneous.

**Lemma 3.6.7.** *For any  $i \in \mathcal{N}$  and for any  $U$  and  $V$  in  $\mathbf{L}_T^1$ , we have*

$$|\rho^i(U + V) - \rho^i(U)| \leq \rho^i(|V|). \quad (3.49)$$

*Proof.* By the subadditivity and by the monotonicity of  $\rho^i$ , we have

$$\rho^i(U + V) - \rho^i(U) \leq (\rho^i(U) + \rho^i(V)) - \rho^i(U) = \rho^i(V) \leq \rho^i(|V|).$$

Similarly, we have

$$\rho^i(U) - \rho^i(U + V) = \rho^i(U + V - V) - \rho^i(U + V) \leq \rho^i(-V) \leq \rho^i(|V|).$$

Inequality (3.49) follows.  $\square$

**Lemma 3.6.8.** *There exists  $C > 0$  such that for any  $(i, t) \in \mathcal{N} \times \mathcal{T}$  and for any  $U \in \mathbf{L}_T^1$ ,*

$$\frac{1}{C} \mathbb{E}[|U|] \leq \rho^i(U) \leq C \mathbb{E}[|U|]. \quad (3.50)$$

*Proof.* All constants  $C$  in the proof are independent of  $U$ . Recall the definition of  $\pi_t^i$ , introduced in the proof of Lemma 3.6.6. We prove by backward induction that for any  $t \in \bar{\mathcal{T}}$ , there exists  $C > 0$  such that for any  $U \in \mathbf{L}_T^1$ ,

$$\frac{1}{C} \mathbb{E} [|U| | \mathcal{F}_t] \leq \pi_t^i(U) \leq C \mathbb{E} [|U| | \mathcal{F}_t], \quad \text{a.s.}$$

The claim is trivial for  $t = T$ . Let  $t \in \mathcal{T}$ . Assume that the claim holds true for  $t + 1$ . We first observe that for any  $U \in \mathbf{L}_{t+1}^1$ ,

$$\frac{1}{C} \mathbb{E} [|U| | \mathcal{F}_t] \leq \rho_t^i(U) \leq C \mathbb{E} [|U| | \mathcal{F}_t], \quad \text{a.s.}, \quad (3.51)$$

as a direct consequence of Assumption 3.2.2. It follows with the monotonicity of  $\rho_t^i$  that

$$\begin{aligned} \pi_t^i(U) &= \rho_t^i \circ \pi_{t+1}^i(U) \\ &\leq \rho_t^i (C \mathbb{E} [|U| | \mathcal{F}_{t+1}]) \\ &\leq C \mathbb{E} [C \mathbb{E} [|U| | \mathcal{F}_{t+1}] | \mathcal{F}_t] \leq C \mathbb{E} [|U| | \mathcal{F}_t], \quad \text{a.s.} \end{aligned}$$

Similarly we prove that  $\pi_t^i(|U|) \geq \frac{1}{C} \mathbb{E} [|U| | \mathcal{F}_t]$  a.s. Recalling that  $\rho^i(U) = \mathbb{E} [\pi_0(U)]$ , we finally obtain (3.50).  $\square$

The following lemma is an estimate of the second-order moment of suboptimal controls (for problem (3.44)).

**Lemma 3.6.9.** *There exists  $C > 0$  such that for any  $i \in \mathcal{N}$ , if  $\hat{A}^i$  satisfies*

$$\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) \leq \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) + 1, \quad (3.52)$$

then  $\hat{A}^i \in \mathcal{A}^C$ .

*Proof.* Let  $i \in \mathcal{N}$  and let  $\hat{A}^i$  satisfy (3.52). All constants  $C$  in the proof are independent of  $\hat{A}^i$ . We have

$$\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) \leq \mathcal{J}^{i,N}(0, \bar{\mathbf{A}}^{-i}) + 1 = \rho^i \left( \sum_{t=0}^T F \left( t, X_t^i[0], b_{(0, \bar{\mathbf{A}}^{-i})}^N \right) \right) + 1.$$

By Assumption 3.2.4 (i), Lemma 3.6.8, and Remark 3.6.2,

$$\rho^i \left( \sum_{t=0}^T F \left( t, X_t^i[0], b_{(0, \bar{\mathbf{A}}^{-i})}^N \right) \right) \leq C \mathbb{E} \left[ T + \sum_{t=0}^T |X_t^i[0]|^2 \right] \leq C.$$

Therefore,

$$\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) \leq C. \quad (3.53)$$

We need now to bound  $\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i})$  from below. We obtain by using successively Lemma 3.6.8, Assumptions 3.2.4 (i) and (iv), and Young's inequality that

$$\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) \geq \frac{1}{C} \mathbb{E} \left[ \sum_{t=0}^{T-1} \left( \frac{1}{2} |\hat{A}_t^i|^2 - C |\hat{A}_t^i| \right) \right] - C \geq \frac{1}{C} \mathbb{E} \left[ \sum_{t=0}^{T-1} |\hat{A}_t^i|^2 \right] - C. \quad (3.54)$$

We deduce then from (3.53) and (3.54) that  $\mathbb{E} \left[ \sum_{t=0}^{T-1} |\hat{A}_t^i|^2 \right] \leq C$ , which concludes the proof.  $\square$

In the following we fix a constant  $c > 0$  such that the result of Lemma 3.6.9 holds and such that  $\bar{A}^i \in \mathcal{A}^c$  for any  $i \in \mathcal{N}$ . Let  $b$  and  $b'$  in  $\mathcal{B}_2$ , for any  $(t, t', x) \in \mathcal{T} \times \bar{\mathcal{T}} \times \mathbb{R}^d$  we define

$$\Delta P(t, b, b') := P(t, b) - P(t, b'), \quad \Delta F(t', x, b, b') := F(t', x, b) - F(t', x, b').$$

For any  $(x, A) \in \mathbb{R}^{Td} \times \mathcal{A}$  we define

$$\Delta \ell(x, A, b, b') := \sum_{t=0}^{T-1} \langle A_t, \Delta P(t, b, b') \rangle + \sum_{t=0}^T \Delta F(t, x_t, b, b').$$

**Remark 3.6.10.** For any  $t \in \mathcal{T}$  and for any  $b$  and  $b'$  in  $\mathcal{B}_2$ , we have

$$\|\Delta F(t, \cdot, b, b')\|_{\mathcal{G},1} \leq 2Cd_1(b, b')^{1/2}.$$

Indeed if  $d_1(b, b') \geq 1$ , Assumption 3.6.1 (i) yields

$$\|\Delta F(t, \cdot, b, b')\|_{\mathcal{G},1} \leq 2 \sup_{b \in \mathcal{B}_2} \|F(t, \cdot, b)\|_{\mathcal{G},1} \leq 2C.$$

If  $d_1(b, b') \leq 1$ , by Assumption 3.6.1 (ii) we have

$$\|\Delta F(t, \cdot, b, b')\|_{\mathcal{G},1} \leq Cd_1(b, b') \leq Cd_1(b, b')^{1/2}.$$

In the following lemma we study the convergence of the empirical belief to the reference belief  $\bar{b} \in \mathcal{B}_2$ .

**Lemma 3.6.11.** There exists  $C > 0$  such that for any  $i \in \mathcal{N}$  and for any  $A^i \in \mathcal{A}^c$ ,

$$\mathbb{E} \left[ d_1 \left( b_{(A^i, \bar{A}^{-i})}^N, \bar{b} \right) \right] \leq CN^{-\tau(d)}. \quad (3.55)$$

*Proof.* Let  $i \in \mathcal{N}$  and let  $A^i \in \mathcal{A}^c$ . For any  $t \in \mathcal{T}$ , we have by the triangle inequality

$$d_1 \left( \mu_{(A^i, \bar{A}^{-i})}^N(t), \bar{\mu}(t) \right) \leq d_1 \left( \mu_{(A^i, \bar{A}^{-i})}^N(t), \mu_{\bar{A}}^N(t) \right) + d_1 \left( \mu_{\bar{A}}^N(t), \bar{\mu}(t) \right). \quad (3.56)$$

Let us consider the first term of the right-hand side. By definition of the distance  $d_1$ ,

$$d_1 \left( \mu_{(A^i, \bar{A}^{-i})}^N(t), \mu_{\bar{A}}^N(t) \right) \leq \frac{1}{N} (|X_t^i[A^i] - \bar{X}_t^i| + |A_t^i - \bar{A}_t^i|), \quad \text{a.s.} \quad (3.57)$$

Since the controls  $\bar{A}^i$  and  $A^i$  belong to  $\mathcal{A}^c$ , the first-order moment of  $X_t^i[A^i]$  and  $\bar{X}_t^i$  are finite as a consequence of Remark 3.6.2, thus

$$\mathbb{E} [|X_t^i[A^i] - \bar{X}_t^i| + |A_t^i - \bar{A}_t^i|] \leq C. \quad (3.58)$$

Therefore, by (3.57) and (3.58), we have

$$d_1 \left( \mu_{(A^i, \bar{A}^{-i})}^N(t), \mu_{\bar{A}}^N(t) \right) \leq \frac{C}{N}. \quad (3.59)$$

Let us consider now the second-term of the right-hand side of (3.56). We recall that  $\bar{\mu}(t) = (id, \bar{\alpha}_t) \# \bar{m}(t)$ . Since  $\bar{A}_t^j = \bar{\alpha}_t(\bar{X}_t^j)$ , we also have  $\mu_{\bar{A}}^N(t) = (id, \bar{\alpha}_t) \# \bar{m}_{\bar{A}}^N(t)$ . We deduce from the Lipschitz continuity of  $(id, \bar{\alpha}_t)$  and from Lemma 3.4.2 that

$$\begin{aligned} d_1 \left( \mu_{\bar{A}}^N(t), \bar{\mu}(t) \right) &= d_1 \left( (id, \bar{\alpha}_t) \# \bar{m}_{\bar{A}}^N(t), (id, \bar{\alpha}_t) \# \bar{m}(t) \right) \\ &\leq Cd_1 \left( \bar{m}_{\bar{A}}^N(t), \bar{m}(t) \right). \end{aligned} \quad (3.60)$$



The random variables  $\bar{X}_t^j$  are independent and  $\mathcal{L}(\bar{X}_t^j) \sim \bar{m}(t)$ . Therefore, Theorem 3.6.5 applies and yields

$$\mathbb{E} \left[ d_1 \left( \bar{m}_{\mathbf{A}}^N(t), \bar{m}(t) \right) \right] \leq CN^{-\tau(d)}. \quad (3.61)$$

Combining (3.56), (3.59), (3.60), and (3.61), we obtain

$$\mathbb{E} \left[ d_1 \left( \mu_{(A^i, \bar{\mathbf{A}}^{-i})}^N(t), \bar{\mu}(t) \right) \right] \leq CN^{-\tau(d)}.$$

It is then easy to verify that

$$\mathbb{E} \left[ d_1 \left( m_{(A^i, \bar{\mathbf{A}}^{-i})}^N(T), \bar{m}(T) \right) \right] \leq CN^{-\tau(d)}.$$

Estimate (3.55) follows immediately.  $\square$

**Lemma 3.6.12.** *There exists  $C > 0$  such that for any  $i \in \mathcal{N}$ , for any  $t \in \bar{\mathcal{T}}$ , and for any  $A^i \in \mathcal{A}^c$ , we have*

$$\mathbb{E} \left[ \left| \Delta F \left( t, X_t^i[A^i], b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right| \right] \leq CN^{-\tau(d)/2}.$$

*Proof.* Let  $N \in \mathbb{N}^*$  and let  $t \in \bar{\mathcal{T}}$ . By Remark 3.6.10, we have

$$\mathbb{E} \left[ \left| \Delta F \left( t, X_t^i[A^i], b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right| \right] \leq C \mathbb{E} \left[ (1 + |X_t^i[A^i]|) d_1 \left( b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right)^{1/2} \right]. \quad (3.62)$$

Since  $A^i \in \mathcal{A}^c$ , by Remark 3.6.2 we have that  $\mathbb{E} [|X_t^i[A^i]|^2] \leq C$ . We obtain with the Cauchy-Schwarz inequality and Lemma 3.6.11 that

$$\begin{aligned} \mathbb{E} \left[ (1 + |X_t^i[A^i]|) d_1 \left( b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right)^{1/2} \right] &\leq \mathbb{E} \left[ (1 + |X_t^i[A^i]|)^2 \right]^{1/2} \mathbb{E} \left[ d_1 \left( b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right]^{1/2} \\ &\leq CN^{-\tau(d)/2}. \end{aligned} \quad (3.63)$$

Combining (3.62) and (3.63), we obtain the announced inequality.  $\square$

**Lemma 3.6.13.** *There exists  $C > 0$  such that for any  $i \in \mathcal{N}$ , for any  $t \in \mathcal{T}$  and for any  $A^i \in \mathcal{A}^c$ , we have*

$$\mathbb{E} \left[ \left| \langle A_t^i, \Delta P \left( t, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \rangle \right| \right] \leq CN^{-\tau(d)/2}. \quad (3.64)$$

*Proof.* Let  $i \in \mathcal{N}$ , let  $t \in \mathcal{T}$ , and let  $A^i \in \mathcal{A}^c$ . By the Cauchy-Schwarz inequality, we have

$$\mathbb{E} \left[ \left| \langle A_t^i, \Delta P \left( t, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \rangle \right| \right] \leq C \left( \mathbb{E} \left[ \left| \Delta P \left( t, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right|^2 \right] \right)^{1/2}. \quad (3.65)$$

We obtain with Assumptions 3.2.4 (iii-iv) and Lemma 3.6.11 that

$$\begin{aligned} \mathbb{E} \left[ \left| \Delta P \left( t, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right|^2 \right] &\leq 2C \mathbb{E} \left[ \left| \Delta P \left( t, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right| \right] \\ &\leq C \mathbb{E} \left[ d_1 \left( b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right] \leq CN^{-\tau(d)}. \end{aligned} \quad (3.66)$$

Combining (3.65) and (3.66), we deduce (3.64).  $\square$

We finally prove the main result of the section.

*Proof of Theorem 3.6.4.* Let  $i \in \mathcal{N}$ . We first show that for any  $A^i \in \mathcal{A}^c$ , the inequality

$$|\mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) - \mathcal{J}^i(A^i, \bar{b})| \leq CN^{-\tau(d)/2} \quad (3.67)$$

holds for some constant  $C > 0$  independent of  $A^i$ . This will imply (3.47). For any  $A^i \in \mathcal{A}^c$ , we can write  $\mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) = \rho^i(Y)$  and  $\mathcal{J}^i(A^i, \bar{b}) = \rho^i(Z)$ , where

$$\begin{aligned} Y &:= \sum_{t=0}^{T-1} \ell \left( t, X_t^i[A^i], A_t^i, b_{(A^i, \bar{\mathbf{A}}^{-i})}^N \right) + F \left( T, X_T^i[A^i], b_{(A^i, \bar{\mathbf{A}}^{-i})}^N \right), \\ Z &:= \sum_{t=0}^{T-1} \ell \left( t, X_t^i[A^i], A_t^i, \bar{b} \right) + F \left( T, X_T^i[A^i], \bar{b} \right). \end{aligned}$$

Applying Lemma 3.6.7 with  $U = Z$  and  $V = Y - Z$  we have

$$|\mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) - \mathcal{J}^i(A^i, \bar{b})| = |\rho^i(Y) - \rho^i(Z)| \leq \rho^i(|Y - Z|).$$

In addition, Lemma 3.50 yields

$$\rho^i(|Y - Z|) \leq C\mathbb{E}[|Y - Z|] = C\mathbb{E} \left[ \left| \Delta \ell \left( X^i[A^i], b_{(A^i, \bar{\mathbf{A}}^{-i})}^N, \bar{b} \right) \right| \right].$$

We finally obtain (3.67) with Lemma 3.6.12 and Lemma 3.6.13.

Let us fix now  $\hat{A}^i \in \mathcal{A}$  such that

$$\mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) \leq \left( \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) \right) + \min \left\{ 1, N^{-\tau(d)/2} \right\}. \quad (3.68)$$

By Lemma 3.6.9, we have  $\hat{A}^i \in \mathcal{A}^c$ . Thus inequality (3.67) yields

$$\begin{aligned} \mathcal{J}^i(\hat{A}^i, \bar{b}) &\leq \mathcal{J}^{i,N}(\hat{A}^i, \bar{\mathbf{A}}^{-i}) + CN^{-\tau(d)/2} \\ &\leq \left( \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) \right) + CN^{-\tau(d)/2}. \end{aligned} \quad (3.69)$$

We apply again inequality (3.67) to  $A^i = \bar{A}^i$ . Using also the optimality of  $\bar{A}^i$  (with respect to  $\mathcal{J}^i$ ), we obtain

$$\mathcal{J}^{i,N}(\bar{\mathbf{A}}) - CN^{-\tau(d)/2} \leq \mathcal{J}^i(\bar{A}^i, \bar{b}) \leq \mathcal{J}^i(\hat{A}^i, \bar{b}). \quad (3.70)$$

Finally, combining (3.69) and (3.70) we have

$$\mathcal{J}^{i,N}(\bar{\mathbf{A}}) \leq \left( \inf_{A^i \in \mathcal{A}} \mathcal{J}^{i,N}(A^i, \bar{\mathbf{A}}^{-i}) \right) + CN^{-\tau(d)/2},$$

which shows that  $\bar{\mathbf{A}}$  is an  $\varepsilon$ -Nash equilibrium with  $\varepsilon = CN^{-\tau(d)/2}$ .  $\square$

## 3.7 Conclusion

This paper has studied a mean field game model with risk averse agents, and provided a framework under which an equilibrium holds, for a large class of composite risk measures and congestion terms. The specific structure of the integral cost of the agents has been exploited in order to rewrite the dynamic programming equations in a functional form (using the Moreau envelope and the proximal operator). In that way, the coupled system could be formulated as an equivalent fixed point equation, yielding the existence of a

solution. Regularity properties have been obtained for risk averse agents. This has allowed to show that an optimal feedback control (for the mean field game) results in an  $\varepsilon$ -Nash equilibrium for a related dynamic game with  $N$  players. Future work could focus on the uniqueness of the Nash equilibrium with contraction arguments and smallness assumptions on the coupling terms. In this work, risk averse (with respect to their own noise) agents have been considered; investigating a mean field game model with common noise and risk averse agents would be of particular interest. Finally, we could investigate variants of our model involving agents driven by nonlinear dynamical systems, nonconvex data functions, or exponential utility cost functions. In such a setting we cannot expect anymore the value function to be convex and thus, a feedback policy cannot be defined in a unique manner. A different notion of equilibrium must then be employed. An appropriate one may rely on the distribution of the controls of the agents at each time, conditioned to their position, as for example in [113], where an existence result is obtained with Kakutani's theorem.



## Chapter 4

# Discrete potential mean field games

### 4.1 Introduction<sup>1</sup>

The class of mean field game (MFG) problems was introduced by J.-M. Lasry and P.-L. Lions in [88, 89, 90] and M. Huang, R. Malhamé, and P. Caines in [80] to study interactions among a large population of agents. The agents of the game optimize their own dynamical system with respect to a criterion; the criterion is parameterized by some endogenous coupling terms. These coupling terms are linked to the collective behavior of all agents and thus induce an interaction between them. It is assumed that an isolated agent has no impact on the coupling terms and that all agents are identical. At a mathematical level, MFGs typically take the form of a system of coupled equations: a dynamic programming equation (characterizing the optimal behavior of the agents), a Kolmogorov equation (describing the distribution of the agents), and coupling equations.

In this work we study a class of discrete time and finite state space mean field games with potential structure. The dynamical system of each agent is a Markov chain, with controlled probability transitions. Our results are motivated by discrete time and discrete space models as such but can be applied to discretized MFGs. Few publications deal with fully discrete models; in a seminal work, D. Gomes, J. Mohr, and R. R. Souza [65] have studied the existence of a Nash equilibrium via a fixed point approach and investigated the long-term behavior of the game.

Potential (also called variational) MFGs are coupled systems which can be interpreted as first-order conditions of two control problems in duality whose state equations are respectively a Kolmogorov equation and a dynamic programming equation. The primal problem (involving the Kolmogorov equation) can be interpreted as a stochastic optimal control problem with cost and constraints on the law of the state and the control. Its numerical resolution is thus of interest beyond the context of MFGs.

In our model, the agents interact with each other via two coupling terms: a congestion variable  $\gamma$  and a price variable  $P$ . The congestion  $\gamma$  is linked to the distribution of the agents via the subdifferential of a proper convex and l.s.c. potential  $F$ . The price  $P$  is linked to the joint law of states and controls of the agents via the subdifferential of a proper convex and l.s.c. potential  $\phi$ . A specificity of our discrete model is that the potentials  $F$  and  $\phi$  can take the value  $+\infty$  and thus induce constraints on the distribution of the agents, referred to as hard constraints. In the continuous case, four classes of variational MFGs can be identified. Our model is general enough to be seen as

---

<sup>1</sup>The corresponding article [22] has been submitted to Mathematical Programming.

the discrete counterpart of these four cases. Case 1: MFGs with monotone congestion terms ( $F$  is differentiable,  $\phi = 0$ ). The first variational formulation was given in [89] and has been widely studied in following works [17, 34, 38, 93, 106]. Case 2: MFGs with density constraints ( $F$  has a bounded domain,  $\phi = 0$ ). These models are of particular interest for describing crowd motions. The coupling variable  $\gamma$  has there an incentive role. The reader can refer to [38, 93, 114, 115]. Case 3: MFGs with Cournot interactions ( $F = 0$ ,  $\phi$  is differentiable). In this situation, each agent optimally chooses a quantity to be sold at each time step of the game. Interactions with the other players occur through the gradient of  $\phi$  which maps the mean strategy (the market demand) to a market price. See for example [21, 67, 68, 70, 72]. Case 4: MFGs with price formation ( $F = 0$ ,  $\phi$  has a bounded domain). These models incorporate a hard constraint on the demand. The price variable is the associated Lagrange multiplier and has a incentive role. We refer to [66].

The first part of the article is devoted to the theoretical analysis of the MFG system. We first introduce a potential problem, shown to be equivalent to a convex problem involving the Kolmogorov equation via a change of variable, similar to the one widely employed in the continuous setting (e.g. in [14]). Under a suitable qualification condition, we establish a duality result between this problem and an optimal control problem involving the dynamic programming equation. We show the existence of solutions to these problems and finally we show the existence of a solution to the MFG system. A uniqueness result is proved (when  $F$  and  $\phi$  are differentiable).

The second part of the article is devoted to the numerical resolution of the MFG system. We focus on two families of methods: primal-dual methods and augmented Lagrangian methods. These two classes exploit the duality structure discussed above and can deal with hard constraints. They have already been applied to continuous MFGs, see for example the survey article [6]. Primal-dual methods have been applied to stationary MFGs with hard congestion terms in [28] and to time-dependent MFGs in [27]. Augmented Lagrangian methods have been applied to MFGs in [15] and to MFGs with hard congestion terms in [17]. Other methods exploiting the potential structure have been investigated in the literature, they are out of the scope of the current article. Let us mention the fictitious play method [35, 55, 77] and the Sinkhorn algorithm [16]. Let us emphasize that the above references all deal with interaction terms depending on the distribution of the states of the agents; very few publications are concerned by interactions through the controls (see [5]). The present work is the first to address methods for ‘‘Cournot’’ mean field games.

Let us comment further on the families of methods under investigation and our contribution. The primal-dual algorithms that we have implemented were introduced by A. Chambolle and T. Pock [45]. A novelty of our work is to show that the extension of primal-dual methods of [46], involving nonlinear proximity operators (based on Bregman divergences), can also be used to solve MFGs. The augmented Lagrangian method that we have implemented is applied to the dual problem (involving the dynamic programming equation), as originally proposed in [14] for optimal transportation problems. As in [14], we have actually implemented a variant of the augmented Lagrangian method, called alternating direction method of multipliers (ADMM). The method was introduced by R. Glowinski and A. Marroco [64] and studied by D. Gabay and B. Mercier [61]. It relies on a successive minimization of the augmented Lagrangian function. One of the main limitations of ADMM is that when the number of involved variables is greater or equal to three, as it is the case for our problem, convergence is not granted. A novelty of our work is to consider a variant of ADMM, the alternating direction method with Gaussian

back substitution (ADM-G), introduced in [78]. At each iteration of this method, the ADMM step is followed by a Gaussian back substitution step. Convergence is ensured. The practical implementation of the additional step turns out to be inexpensive in our framework.

The last contribution of this work is to propose and solve numerically two hard constraints problems: a congestion mean field game problem and a ‘‘Cournot’’ mean field game. Following our analysis we define a notion of residuals allowing us to compare the empirical convergence of each method in a common setting.

The article is organized as follows. In section 4.2 we provide the main notations, the mean field game system under study and the underlying individual player problem. In section 4.3 we formulate a potential problem and perform the announced change of variable. In section 4.4 we form a dual problem and we establish a duality result. In section 4.5 we provide our main results: existence and uniqueness of a solution to the mean field game. In section 4.6 we provide a detailed implementation of the primal-dual proximal algorithms, ADMM and ADM-G, and we give theoretical convergence results when possible. In section 4.7 we present numerical results for two concrete problems. We provide outputs obtained for each method: errors, value function, equilibrium measure, mean displacement, congestion, demand and price.

## 4.2 Discrete mean field games

### 4.2.1 Notation

**Sets.** Let  $T \in \mathbb{N}^*$  denote the duration of the game. We set  $\mathcal{T} = \{0, \dots, T - 1\}$  and  $\bar{\mathcal{T}} = \{0, \dots, T\}$ . Let  $S = \{0, \dots, n - 1\}$  denote the state space. We set

$$\Delta(S) = \left\{ \pi: S \rightarrow [0, 1] \mid \sum_{x \in S} \pi(x) = 1 \right\},$$

$$\Delta = \left\{ \pi: \mathcal{T} \times S \times S \rightarrow [0, 1] \mid \pi(t, x, \cdot) \in \Delta(S), \forall (t, x) \in \mathcal{T} \times S \right\}.$$

For any finite set  $A$ , we denote by  $\mathbb{R}(A)$  the finite-dimensional vector space of mappings from  $A$  to  $\mathbb{R}$ . All along the article, we make use of the following spaces:

$$\begin{aligned} \mathcal{R} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T} \times S^2), & \mathcal{U} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T}), \\ \mathcal{C} &= \mathcal{R} \times \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T}), & \mathcal{K} &= \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathcal{U}. \end{aligned}$$

**Convex analysis.** For any function  $g: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ , we denote

$$\text{dom}(g) = \{x \in X \mid g(x) < +\infty\}.$$

The subdifferential of  $g$  is defined by

$$\partial g(x) = \left\{ x^* \in \mathbb{R}^d \mid g(x') \geq g(x) + \langle x^*, x' - x \rangle, \forall x' \in \mathbb{R}^d \right\}.$$

By convention,  $\partial g(x) = \emptyset$  if  $g(x) = +\infty$ . Note also that  $x^* \in \partial g(x)$  if and only if  $g(x) + g^*(x^*) = \langle x, x^* \rangle$ , where  $g^*$  is the Fenchel transform of  $g$ , defined by

$$g^*(x^*) = \sup_{x \in \mathbb{R}^d} \langle x, x^* \rangle - g(x).$$

Note that the subdifferential and Fenchel transforms of  $\ell$ ,  $F$ , and  $\phi$  (introduced in the next paragraph) are considered for fixed values of the time and space variables.

We denote by  $\chi$  the indicator function of  $\{0\}$  (without specifying the underlying vector space). For any subset  $C \subseteq \mathbb{R}^d$ , we denote by  $\chi_C$  the indicator function of  $C$ . For any  $x \in C$ , we denote by  $N_C(x)$  the normal cone to  $C$  at  $x$ ,

$$N_C(x) = \left\{ x^* \in \mathbb{R}^d \mid \langle x^*, x' - x \rangle \leq 0, \forall x' \in C \right\}.$$

We set  $N_C(x) = \emptyset$  if  $x \notin C$ .

**Nemytskii operators.** Given two mappings  $g: \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$  and  $u: \mathcal{X} \rightarrow \mathcal{Y}$ , we call Nemytskii operator the mapping  $\mathbf{g}[u]: \mathcal{X} \rightarrow \mathcal{Z}$  defined by

$$\mathbf{g}[u](x) = g(x, u(x)).$$

We will mainly use this notation in order to avoid the repetition of time and space variables, for example, we will write  $\ell[\pi](t, x)$  instead of  $\ell(t, x, \pi(t, x))$ .

All along the article, we will transpose some notions associated with  $g$  to the Nemytskii operator  $\mathbf{g}[u]$ . When  $\mathcal{Y} = \mathbb{R}^d$  and  $\mathcal{Z} = \mathbb{R} \cup \{+\infty\}$ , we define the domain of  $g$  by

$$\mathbf{dom}(g) = \left\{ u: \mathcal{X} \rightarrow \mathbb{R}^d \mid u(x) \in \mathbf{dom}(g(x, \cdot)), \forall x \in \mathcal{X} \right\}.$$

We define  $\mathbf{g}^*[v]: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  by  $\mathbf{g}^*[v](x) = g^*(x, v(x))$ , where  $g^*$  is the Fenchel transform of  $g$  with respect to the second variable.

### 4.2.2 Coupled system

**Data and assumption.** We fix an initial distribution  $m_0 \in \Delta(S)$  and four maps: a running cost  $\ell$ , a potential price function  $\phi$ , a potential congestion cost  $F$ , and a displacement cost  $\alpha$ ,

$$\begin{aligned} \ell: \mathcal{T} \times S \times \mathbb{R}(S) &\rightarrow \mathbb{R} \cup \{+\infty\}, & \phi: \mathcal{T} \times \mathbb{R} &\rightarrow \mathbb{R} \cup \{+\infty\}, \\ F: \bar{\mathcal{T}} \times \mathbb{R}(S) &\rightarrow \mathbb{R} \cup \{+\infty\}, & \alpha: \mathcal{T} \times S^2 &\rightarrow \mathbb{R}. \end{aligned}$$

The following convexity assumption is in force all along the article. Note that we will later make use of an additional qualification assumption (Assumption 4.4.1).

**Assumption 4.2.1** (Convexity). *For any  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ , the maps  $\ell(t, x, \cdot)$ ,  $F(s, \cdot)$ , and  $\phi(t, \cdot)$  are proper, convex and lower semi-continuous. In addition  $\mathbf{dom}(\ell(t, x, \cdot)) \subseteq \Delta(S)$ .*

**Coupled system.** The unknowns of the MFG system introduced below are denoted  $((m, \pi), (u, \gamma, P)) \in \mathcal{R} \times \mathcal{K}$ . They can be described as follows:

- $\gamma$  and  $P$  are the coupling terms of the MFG:  $\gamma(t, x)$  is a congestion term incurred by agents located at  $x \in S$  at time  $t \in \bar{\mathcal{T}}$  and  $P(t)$  is a price variable
- $\pi(t, x, y)$  denotes the probability transition from  $x \in S$  to  $y \in S$ , for agents located at  $x$  at time  $t$
- $m(t, x)$  denotes the proportion of agents located at  $x \in S$  at time  $t \in \bar{\mathcal{T}}$
- $u(t, x)$  is the value function of the agents.



For any  $(\gamma, P) \in \mathcal{U}$ , we define the individual cost  $c: \mathcal{T} \times S \times S \times \Delta(S) \rightarrow \mathbb{R}$ ,

$$c_{\gamma, P}(t, x, y, \rho) = \ell(t, x, \rho) + \gamma(t, x) + \alpha(t, x, y)P(t).$$

Given  $(m, \pi) \in \mathcal{R}$ , we denote

$$\mathbf{Q}[m, \pi](t) = \sum_{(x, y) \in S^2} m(t, x) \pi(t, x, y) \alpha(t, x, y).$$

We aim at finding a quintuplet  $(m, \pi, u, \gamma, P)$  such that for any  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ ,

$$\left\{ \begin{array}{l} \text{(i)} \quad \begin{cases} u(t, x) = \inf_{\rho \in \Delta(S)} \sum_{y \in S} \rho(y) (c_{\gamma, P}(t, x, y, \rho) + u(t+1, y)), \\ u(T, x) = \gamma(T, x), \end{cases} \\ \text{(ii)} \quad \pi(t, x, \cdot) \in \arg \min_{\rho \in \Delta(S)} \sum_{y \in S} \rho(y) (c_{\gamma, P}(t, x, y, \rho) + u(t+1, y)), \\ \text{(iii)} \quad \begin{cases} m(t+1, x) = \sum_{y \in S} m(t, y) \pi(t, y, x), \\ m(0, x) = m_0(x), \end{cases} \\ \text{(iv)} \quad \gamma(s, \cdot) \in \partial F(s, m(s, \cdot)), \\ \text{(v)} \quad P(t) \in \partial \phi(t, \mathbf{Q}[m, \pi](t)). \end{array} \right. \quad (\text{MFG})$$

### Heuristic interpretation.

- The dynamical system of each agent is a Markov chain  $(X_s^\pi)_{s \in \bar{\mathcal{T}}}$  controlled by  $\pi \in \Delta$ , with initial distribution  $m_0$ : for any  $(t, x, y) \in \mathcal{T} \times S^2$ ,

$$\mathbb{P}(X_{t+1}^\pi = y | X_t^\pi = x) = \pi(t, x, y), \quad \mathbb{P}(X_0^\pi = x) = m_0(x).$$

Given the coupling terms  $(\gamma, P) \in \mathcal{U}$ , the individual control problem is

$$\inf_{\pi \in \Delta} J_{\gamma, P}(\pi) := \mathbb{E} \left( \sum_{t \in \mathcal{T}} c_{\gamma, P}(t, X_t^\pi, X_{t+1}^\pi, \pi(t, X_t^\pi)) + \gamma(T, X_T^\pi) \right). \quad (4.1)$$

The equations (MFG,i-ii) are the associated dynamic programming equations: given  $(\gamma, P) \in \mathcal{U}$ , if  $u$  and  $\pi$  satisfy these equations, then  $\pi$  is a solution to (4.1). The reader can refer to [20, Chapter 7] for a detailed presentation of the dynamic programming approach for the optimal control of Markov chains.

- Given  $\pi \in \Delta$ , denote by  $m^\pi$  the probability distribution of  $X^\pi$ , that is,  $m^\pi(t, x) = \mathbb{P}(X_t^\pi = x)$ . Then  $m^\pi$  is obtained by solving the Kolmogorov equation (MFG,iii). In the limit when the number of agents tends to  $\infty$ , the distribution  $m^\pi$  coincides with the empirical distribution of the agents.
- Finally, the equations (MFG,iv-v) link the coupling terms  $\gamma$  and  $P$  to the distribution of the agents  $m$  and their control  $\pi$ .

In summary: given a solution  $(m, \pi, u, \gamma, P) \in \mathcal{R} \times \mathcal{K}$  to (MFG), the triplet  $(\pi, \gamma, P)$  is a solution to the mean field game

$$\pi \in \arg \min_{\rho \in \Delta} J_{\gamma, P}(\rho), \quad \gamma \in \partial \mathbf{F}[m^\pi], \quad P \in \partial \phi[\mathbf{Q}[m^\pi, \pi]].$$

**Remark 4.2.2.** • *At any time  $t \in \mathcal{T}$ , it is possible to encode constraints on the transitions of the agents located at  $x \in S$  by defining  $\ell$  in such a way that  $\text{dom}(\ell(t, x, \cdot))$  is strictly included into  $\Delta(S)$ . An example will be considered in Section 4.7.*

- *If  $F$  and  $\phi$  are differentiable, then their subdifferentials are singletons and thus the coupling terms  $\gamma$  and  $P$  are uniquely determined by  $m$  and  $\pi$  through the equations (MFG,iv-v).*
- *The equations (MFG,iv-v) imply that  $m \in \text{dom}(\mathbf{F})$  and  $\mathbf{Q}[m, \pi] \in \text{dom}(\phi)$ . Thus they encode hard constraints on  $m$  and  $\pi$  if the coupling functions  $F$  or  $\phi$  take the value  $+\infty$ . For example, they can be chosen in the form  $G: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $G = g + \chi_K$ , where  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  is convex and differentiable and where  $K$  is a closed and convex subset of  $\mathbb{R}^d$ . Then by [11, Corollary 16.38],*

$$\partial G(x) = \nabla g(x) + N_K(x), \quad \forall x \in \mathbb{R}^d.$$

### 4.2.3 Further notation

We introduce now two linear operators,  $\mathbf{A}$  and  $\mathbf{S}$ . They will allow to bring out the connection between the coupled system and the potential problem. The operator  $\mathbf{A}: \mathbb{R}(\mathcal{T} \times S^2) \rightarrow \mathbb{R}(\mathcal{T})$  and its adjoint  $\mathbf{A}^*: \mathbb{R}(\mathcal{T}) \rightarrow \mathbb{R}(\mathcal{T} \times S^2)$  are given by

$$\mathbf{A}[w](t) = \sum_{(x,y) \in S^2} w(t, x, y) \alpha(t, x, y), \quad \mathbf{A}^*[P](t, x, y) = \alpha(t, x, y) P(t).$$

The operator  $\mathbf{S}: \mathbb{R}(\mathcal{T} \times S^2) \rightarrow \mathbb{R}(\bar{\mathcal{T}} \times S)$  and its adjoint  $\mathbf{S}^*: \mathbb{R}(\bar{\mathcal{T}} \times S) \rightarrow \mathbb{R}(\mathcal{T} \times S^2)$  are given by

$$\mathbf{S}[w](s, x) = \begin{cases} \sum_{y \in S} w(s-1, y, x) & \text{if } s > 0, \\ 0 & \text{if } s = 0, \end{cases} \quad \mathbf{S}^*[u](t, x, y) = u(t+1, y).$$

We can now reformulate the dynamic programming equations of the coupled system (MFG,i-ii) as follows:

$$\begin{cases} \text{(i)} & \begin{cases} u(t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) = \gamma(t, x), \\ u(T, x) = \gamma(T, x), \end{cases} \\ \text{(ii)} & (\ell[\pi] + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u])(t, x) = -\langle \pi(t, x), (\mathbf{A}^*P + \mathbf{S}^*u)(t, x) \rangle. \end{cases}$$

## 4.3 Potential problem and convex formulation

### 4.3.1 Perspective functions

Given  $h: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  a proper l.s.c. and convex function with bounded domain, we define the perspective function  $\tilde{h}: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  by

$$\tilde{h}(\theta, x) = \begin{cases} \theta h(x/\theta) & \text{if } \theta > 0, \\ 0 & \text{if } (\theta, x) = (0, 0), \\ +\infty & \text{otherwise.} \end{cases}$$

**Lemma 4.3.1.** *The perspective function  $\tilde{h}$  is proper, convex, l.s.c. and its domain is given by  $\text{dom}(\tilde{h}) = \{(\theta, x) \in \mathbb{R}_+ \times \mathbb{R}^d \mid x \in \theta \text{dom}(h)\}$ . For any  $(\theta^*, x^*) \in \mathbb{R} \times \mathbb{R}^d$ , we have*

$$\tilde{h}^*(\theta^*, x^*) = \chi_Q(\theta^*, x^*), \quad (4.2)$$

where  $Q := \{(\theta^*, x^*) \in \mathbb{R} \times \mathbb{R}^d, h^*(x^*) + \theta^* \leq 0\}$ .

*Proof.* The proof is a direct application of [25, Lemmas 1.157, 1.158] when  $h$  has a bounded domain. In this case the recession function of  $h$  is the indicator function of zero.  $\square$

**Lemma 4.3.2.** *Let  $(\theta, x), (\theta^*, x^*) \in \mathbb{R} \times \mathbb{R}^d$ . Then  $(\theta^*, x^*) \in \partial\tilde{h}(\theta, x)$  if and only if*

$$\begin{aligned} \text{either: } & h^*(x^*) + \theta^* \leq 0 \quad \text{and} \quad (\theta, x) = (0, 0) \\ \text{or: } & h^*(x^*) + \theta^* = 0, \quad h(x/\theta) + h^*(x^*) - \langle x/\theta, x^* \rangle = 0, \quad \text{and} \quad \theta > 0. \end{aligned}$$

*Proof.* Direct application of [48, Proposition 2.3].  $\square$

### 4.3.2 Potential problem

We define the following criterion

$$\mathcal{J}(m, \pi) = \sum_{(t,x) \in \mathcal{T} \times S} m(t, x) \ell[\pi](t, x) + \sum_{t \in \mathcal{T}} \phi[\mathbf{Q}[m, \pi]](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m](s)$$

and the following potential problem (recall that  $m^\pi$  is the solution to the Kolmogorov equation (MFG,iii), given  $\pi \in \Delta$ ):

$$\inf_{(m, \pi) \in \mathcal{R}} \mathcal{J}(m, \pi), \quad \text{subject to: } m = m^\pi. \quad (P)$$

The link between the mean field game system (MFG) and the potential problem (P) will be exhibited in Section 4.5. Notice that Problem (P) is not convex. Yet we can define a closely related convex problem, whose link with (P) is established in Lemma 4.3.3.

We denote by  $\tilde{\ell}: \mathcal{T} \times S \times \mathbb{R} \times \mathbb{R}(S) \rightarrow \mathbb{R} \cup \{+\infty\}$  the perspective function of  $\ell$  with respect to the third variable. By Lemma 4.3.1 the function  $\tilde{\ell}(t, x, \cdot, \cdot)$  is proper convex and l.s.c. for any  $(t, x) \in \mathcal{T} \times S$ . We define

$$\tilde{\mathcal{J}}(m, w) = \sum_{(t,x) \in \mathcal{T} \times S} \tilde{\ell}[m, w](t, x) + \sum_{t \in \mathcal{T}} \phi[\mathbf{A}w](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m](s).$$

In the above definition,  $\tilde{\ell}$  is the Nemytskii operator of  $\tilde{\ell}$ , that is, for any  $(t, x) \in \mathcal{T} \times S$ ,

$$\tilde{\ell}[m, w](t, x) = \begin{cases} m(t, x) \ell\left(t, x, \frac{w(t, x, \cdot)}{m(t, x)}\right), & \text{if } m(t, x) > 0, \\ 0, & \text{if } m(t, x) = 0 \text{ and } w(t, x, \cdot) = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

We consider now the following convex problem:

$$\inf_{(m, w) \in \mathcal{R}} \tilde{\mathcal{J}}(m, w), \quad \text{subject to: } \mathbf{S}w - m + \bar{m}_0 = 0, \quad (\tilde{P})$$

where  $\bar{m}_0 \in \mathbb{R}(\bar{\mathcal{T}} \times S)$  is defined by

$$\bar{m}_0(s, x) = \begin{cases} m_0(x) & \text{if } s = 0, \\ 0 & \text{otherwise.} \end{cases}$$

**Lemma 4.3.3.** *Let  $\text{val}(P)$  and  $\text{val}(\tilde{P})$  respectively denote the values of problems  $(P)$  and  $(\tilde{P})$ . Then  $\text{val}(P) = \text{val}(\tilde{P})$ . In addition, if Problem  $(P)$  is feasible, then both problems  $(P)$  and  $(\tilde{P})$  have a non-empty bounded set of solutions.*

*Proof. Step 1:*  $\text{val}(P) \geq \text{val}(\tilde{P})$ . Let  $(m, \pi) \in \text{dom}(\mathcal{J})$  be such that  $m = m^\pi$ . Let

$$w(t, x, \cdot) := m(t, x)\pi(t, x, \cdot), \quad (4.3)$$

for any  $(t, x) \in \mathcal{T} \times S$ . Then  $(m, w)$  is feasible for problem  $(\tilde{P})$  and

$$m(t, x)\ell(t, x, \pi(t, x, \cdot)) = \tilde{\ell}(t, x, m(t, x), w(t, x, \cdot)), \quad (4.4)$$

for any  $(t, x) \in \mathcal{T} \times S$ . Indeed by definition of  $\tilde{\ell}(t, x, \cdot, \cdot)$ , if  $m(t, x) > 0$  then (4.4) holds and if  $m(t, x) = 0$  then  $w(t, x, \cdot) = 0$  and (4.4) still holds. It follows that  $\mathcal{J}(m, \pi) = \tilde{\mathcal{J}}(m, w)$  and consequently,  $\text{val}(P) \geq \text{val}(\tilde{P})$ .

*Step 2:*  $\text{val}(P) \leq \text{val}(\tilde{P})$ . Let  $(m, w) \in \text{dom}(\tilde{\mathcal{J}})$  be such that  $\mathbf{S}w - m = \bar{m}_0$  and let  $\pi$  be such that

$$\begin{cases} \pi(t, x, \cdot) = w(t, x, \cdot)/m(t, x) & \text{if } m(t, x) > 0, \\ \pi(t, x, \cdot) \in \text{dom}(\ell(t, x, \cdot)) & \text{otherwise,} \end{cases} \quad (4.5)$$

for all  $(t, x) \in \mathcal{T} \times S$ . Then (4.4) is satisfied and  $(m, \pi)$  is feasible for  $(P)$ . Thus  $\mathcal{J}(m, \pi) = \tilde{\mathcal{J}}(m, w)$ , and consequently,  $\text{val}(P) \leq \text{val}(\tilde{P})$ .

*Step 3: non-empty and bounded sets of solutions.* Since  $\mathcal{J}(m^\pi, \pi)$  is l.s.c. with non-empty bounded domain, it reaches its minimum on its domain. Then the set of solutions to  $(P)$  is non-empty and bounded. Now let  $(m, \pi)$  be a solution to  $(P)$  and let  $w$  be given by (4.3). We have that

$$\tilde{\mathcal{J}}(m, w) = \mathcal{J}(m, \pi) = \text{val}(P) = \text{val}(\tilde{P}),$$

thus we deduce that the set of solutions to  $(\tilde{P})$  is non-empty. It remains to show that the set of solutions to  $(\tilde{P})$  is bounded. Let  $(m, w)$  be a solution to  $(\tilde{P})$ . The Kolmogorov equation implies that  $0 \leq m(t, x) \leq 1$ , for any  $(t, x) \in \bar{\mathcal{T}} \times S$ . By Lemma 4.3.1, we have  $w(t, x, \cdot) \in m(t, x)\Delta(S)$ , which implies that  $0 \leq w(t, x, y) \leq 1$ .  $\square$

Note that the above proof shows how to deduce a solution to  $(\tilde{P})$  out of a solution to  $(P)$  and vice-versa, thanks to relations (4.3) and (4.5).

## 4.4 Duality

We show in this section that Problem  $(\tilde{P})$  is the dual of an optimization problem, denoted  $(D)$ , itself equivalent to an optimal control problem of the dynamic programming equation, problem  $(\tilde{D})$ . For this purpose, we introduce a new assumption (Assumption 4.4.1), which is assumed to be satisfied all along the rest of the article.

### 4.4.1 Duality result

The dual problem is given by

$$\begin{aligned} & \sup_{(u, \gamma, P) \in \mathcal{K}} \mathcal{D}(u, \gamma, P) := \langle m_0, u(0, \cdot) \rangle - \sum_{t \in \mathcal{T}} \phi^*[P](t) - \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}^*[\gamma](s), \\ \text{subject to: } & \begin{cases} u(t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) \leq \gamma(t, x) & (t, x) \in \mathcal{T} \times S, \\ u(T, x) = \gamma(T, x), & x \in S. \end{cases} \end{aligned} \quad (D)$$

Note that the above kind of dynamic programming equation involves inequalities (and not equalities as in (MFG,i)).

We introduce now a qualification condition, which will allow to prove the main duality result of the section. For any  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \mathcal{K}$  and  $\pi \in \mathbf{dom}(\ell)$  we define  $m_1[\varepsilon, \pi]$  the solution to the following perturbed Kolmogorov equation

$$m_1(t+1, x) = \sum_{y \in S} m_1(t, y) \pi(t, y, x) - \varepsilon_1(t+1, x), \quad m_1(0) - \varepsilon_1(0) = \bar{m}_0. \quad (4.6)$$

We also define, for any  $(t, x, y) \in \mathcal{T} \times S \times S$ ,

$$\begin{aligned} w[\varepsilon, \pi](t, x, y) &= m_1[\varepsilon, \pi](t, x) \pi(t, x, y) \\ m_2[\varepsilon, \pi](t, x) &= m_1[\varepsilon, \pi](t, x) + \varepsilon_2(t, x) \\ D[\varepsilon, \pi](t) &= \sum_{(x, y) \in S^2} w[\varepsilon, \pi](t, x, y) \alpha(t, x, y) + \varepsilon_3(t). \end{aligned} \quad (4.7)$$

**Assumption 4.4.1** (Qualification). *There exists  $\alpha > 0$  such that for any  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$  in  $\mathcal{K}$  with  $\|\varepsilon\| \leq \alpha$ , there exists  $\pi \in \mathbf{dom}(\ell)$  such that*

$$m_1[\varepsilon, \pi] \geq 0, \quad m_2[\varepsilon, \pi] \in \mathbf{dom}(\mathbf{F}), \quad D[\varepsilon, \pi] \in \mathbf{dom}(\phi). \quad (4.8)$$

Note that the qualification assumption implies the feasibility of Problems ( $\tilde{P}$ ) and ( $P$ ).

**Remark 4.4.2.** *Assume that  $\text{int}(\mathbf{dom}(\mathbf{F}))$  and  $\text{int}(\mathbf{dom}(\phi))$  are non-empty sets. Then in this case, Assumption 4.4.1 is satisfied if there exists  $\pi \in \mathbf{dom}(\ell)$  such that*

$$m_1[0, \pi] = m_2[0, \pi] \in \text{int}(\mathbf{dom}(\mathbf{F}) \cap \mathbb{R}_+(\bar{\mathcal{T}} \times S)), \quad D[0, \pi] \in \text{int}(\mathbf{dom}(\phi)).$$

**Theorem 4.4.3.** *Let Assumption 4.4.1 hold true. Then the dual problem ( $D$ ) has a bounded set of solutions and  $\text{val}(D) = \text{val}(\tilde{P})$ .*

*Proof.* The primal problem ( $\tilde{P}$ ) can be formulated as follows:

$$\inf_{(m_1, w, m_2, D) \in \mathcal{C}} \mathcal{F}(m_1, w, m_2, D) + \mathcal{G}(\mathcal{A}(m_1, w, m_2, D)), \quad (\mathfrak{P})$$

where the maps  $\mathcal{F}: \mathcal{C} \rightarrow \mathbb{R} \cup \{+\infty\}$  and  $\mathcal{G}: \mathcal{K} \rightarrow \mathbb{R} \cup \{+\infty\}$  and the operator  $\mathcal{A}: \mathcal{C} \rightarrow \mathcal{K}$  are defined by

$$\begin{aligned} \mathcal{F}(m_1, w, m_2, D) &= \sum_{(t, x) \in \mathcal{T} \times S} \tilde{\ell}[m_1, w](t, x) + \sum_{t \in \mathcal{T}} \phi[D](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}[m_2](s), \\ \mathcal{G}(y_1, y_2, y_3) &= \chi(y_1 + \bar{m}_0) + \chi(y_2) + \chi(y_3), \end{aligned} \quad (4.9)$$

$$\mathcal{A}(m_1, w, m_2, D) = (\mathbf{S}w - m_1, m_1 - m_2, \mathbf{A}w - D).$$

We next prove that the qualification condition

$$0 \in \text{int}(\text{dom}(\mathcal{G}) - \mathcal{A} \text{dom}(\mathcal{F}))$$

is satisfied. This is equivalent to show the existence of  $\alpha > 0$  such that for any  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \mathcal{K}$ , with  $\|\varepsilon\| \leq \alpha$ , there exists  $(m_1, w, m_2, D) \in \text{dom}(\mathcal{F})$  satisfying

$$(\mathbf{S}w - m_1 + \varepsilon_1, m_1 - m_2 + \varepsilon_2, \mathbf{A}w - D + \varepsilon_3) \in \text{dom}(\mathcal{G}) = \{\bar{m}_0\} \times \{0\} \times \{0\}.$$

This is a direct consequence of Assumption 4.4.1. Therefore, we can apply the Fenchel-Rockafellar theorem (see [109, Theorem 31.2]) to problem  $(\mathfrak{P})$ . It follows that the following dual problem has the same value as  $(\mathfrak{P})$  and possesses a solution:

$$\inf_{(u, \gamma, P) \in \mathcal{K}} \mathcal{F}^*(-\mathcal{A}^*(u, \gamma, P)) + \mathcal{G}^*(u, \gamma, P). \quad (\mathfrak{D})$$

It remains to calculate  $\mathcal{F}^*$ ,  $\mathcal{G}^*$ , and  $\mathcal{A}^*$ . For any  $(s, x) \in \bar{\mathcal{T}} \times S$ , we define

$$Q_{s,x} = \begin{cases} \{(a, b) \in \mathbb{R} \times \mathbb{R}(S), & \ell^*(s, x, b) + a \leq 0\}, & \text{if } s < T, \\ \{a \in \mathbb{R}, & a = 0\}, & \text{if } s = T. \end{cases}$$

We then define

$$Q = \prod_{(s,x) \in \bar{\mathcal{T}} \times S} Q_{s,x}. \quad (4.10)$$

For any  $(y_1, y_2, y_3, y_4) \in \mathcal{C}$  we have by Lemma 4.3.1 that

$$\mathcal{F}^*(y_1, y_2, y_3, y_4) = \chi_Q(y_1, y_2) + \sum_{t \in \bar{\mathcal{T}}} \phi^*[y_4](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}^*[y_3](s).$$

The adjoint operator  $\mathcal{A}^*: \mathcal{K} \rightarrow \mathcal{C}$  is given by

$$\mathcal{A}^*(u, \gamma, P) = (\gamma - u, \mathbf{A}^*P + \mathbf{S}^*u, -\gamma, -P).$$

It follows that

$$\mathcal{F}^*(-\mathcal{A}^*(u, \gamma, P)) = \chi_Q(u - \gamma, -\mathbf{A}^*P - \mathbf{S}^*u) + \sum_{t \in \bar{\mathcal{T}}} \phi^*[P](t) + \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}^*[\gamma](s).$$

Moreover,  $\mathcal{G}^*(u, \gamma, P) = -\langle u(0, \cdot), m_0 \rangle$ . It follows that  $(D)$  and  $(\mathfrak{D})$  are equivalent, which concludes the proof of the theorem.  $\square$

#### 4.4.2 A new dual problem

We introduce in this section a new optimization problem, equivalent to  $(D)$ . We define the mapping  $\mathbf{U}: \mathcal{U} \rightarrow \mathbb{R}(\bar{\mathcal{T}} \times S)$  which associates with  $(\gamma, P) \in \mathcal{U}$  the solution  $u \in \mathbb{R}(\bar{\mathcal{T}} \times S)$  to the dynamic programming equation

$$\begin{cases} u(t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) = \gamma(t, x) & (t, x) \in \mathcal{T} \times S, \\ u(T, x) = \gamma(T, x), & x \in S. \end{cases} \quad (4.11)$$

We define the following problem

$$\begin{aligned} \max_{(\gamma, P) \in \mathcal{U}} \tilde{\mathcal{D}}(\gamma, P) &:= \mathcal{D}(\mathbf{U}[\gamma, P], \gamma, P) \\ &= \langle \tilde{m}_0, \mathbf{U}[\gamma, P] \rangle - \sum_{t \in \bar{\mathcal{T}}} \phi^*[P](t) - \sum_{s \in \bar{\mathcal{T}}} \mathbf{F}^*[\gamma](s). \end{aligned} \quad (\tilde{D})$$

**Lemma 4.4.4.** *Problems  $(D)$  and  $(\tilde{D})$  have the same value. Moreover, for any solution  $(u, \gamma, P)$  to  $(D)$ ,  $(\gamma, P)$  is a solution to  $(\tilde{D})$ ; conversely, for any solution  $(\gamma, P)$  to  $(\tilde{D})$  (there exists at least one),  $(\mathbf{U}[\gamma, P], \gamma, P)$  is a solution to  $(D)$ .*

*Proof.* Let  $(\gamma, P) \in \mathcal{U}$ . Then  $(u := \mathbf{U}[\gamma, P], \gamma, P)$  is feasible for problem (D) and by definition,  $\mathcal{D}(u, \gamma, P) = \tilde{\mathcal{D}}(\gamma, P)$ . Therefore,  $\text{val}(\mathbf{D}) \geq \text{val}(\tilde{\mathbf{D}})$ .

Conversely, let  $(u, \gamma, P)$  be feasible for (D). Let  $\hat{u} = \mathbf{U}[\gamma, P]$ . Now we claim that  $\hat{u}(t, x) \geq u(t, x)$ , for any  $(t, x) \in \bar{\mathcal{T}} \times S$  (this is nothing but a comparison principle for our dynamic programming equation). The proof of the claim relies on a monotonicity property of  $\ell^*$ . Given  $b$  and  $b' \in \mathbb{R}(S)$ , we say that  $b \leq b'$  if  $b(x) \leq b'(x)$ , for all  $x \in S$ . Since  $\ell(t, x, \cdot)$  has its domain included in  $\Delta(S)$ , we have

$$b \leq b' \implies \ell^*(t, x, b) \leq \ell^*(t, x, b').$$

Using the above property, it is easy to prove the claim by backward induction. It follows that  $\tilde{\mathcal{D}}(\gamma, P) = \mathcal{D}(\hat{u}, \gamma, P) \geq \mathcal{D}(u, \gamma, P)$  and finally,  $\text{val}(\tilde{\mathbf{D}}) \geq \text{val}(\mathbf{D})$ . Thus the two problems have the same value.

The other claims of the lemma are then easy to verify.  $\square$

**Lemma 4.4.5.** *For any  $(t, x) \in \mathcal{T} \times S$ , the map  $(\gamma, P) \in \mathcal{U} \mapsto \mathbf{U}[\gamma, P](t, x)$  is concave.*

*Proof.* Let  $(t, x) \in \mathcal{T} \times S$ . Given  $\pi \in \Delta$ , consider the Markov chain  $(X_s^\pi)_{s=t, \dots, T}$  defined by

$$\mathbb{P}(X_{s+1}^\pi = y | X_s^\pi = x) = \pi(s, x, y), \quad \forall s = t, \dots, T-1, \quad X_t^\pi = x.$$

By the dynamic programming principle, we have

$$\mathcal{U}[\gamma, P](t, x) = \inf_{\pi \in \Delta} \mathbb{E} \left( \sum_{s=t}^T c_{\gamma, P}(t, X_s^\pi, X_{s+1}^\pi, \pi(s, X_s^\pi)) + \gamma(T, X_T^\pi) \right).$$

The criterion to be minimized in the above equality is affine with respect to  $(\gamma, P)$ , thus it is concave. The infimum of a family of concave functions is again concave, therefore,  $\mathcal{U}[\gamma, P](t, x)$  is concave with respect to  $(\gamma, P)$ .  $\square$

As a consequence of the above Lemma, the criterion  $\tilde{\mathcal{D}}$  is concave.

## 4.5 Connection between the MFG system and potential problems

The connection between the MFG system and the potential problems can be established with the help of seven conditions, which we introduce first. We say that  $(m_1, w, m_2, D) \in \mathcal{C}$  and  $(u, \gamma, P) \in \mathcal{K}$  satisfy the condition (C1) if for any  $(t, x) \in \mathcal{T} \times S$ ,

$$\text{either: } \begin{cases} u(t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) \leq \gamma(t, x), \\ (m_1(t, x), w(t, x)) = (0, 0), \end{cases}$$

$$\text{or: } \begin{cases} u(t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) = \gamma(t, x), \\ \ell[\pi](t, x) + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u](t, x) + \langle \pi(t, x), (\mathbf{A}^*P + \mathbf{S}^*u)(t, x) \rangle = 0, \\ m_1(t, x) > 0, \end{cases}$$

where  $\pi(t, x) = w(t, x)/m_1(t, x)$ . We say that the conditions (C2-C7) are satisfied if

$$\begin{aligned} \text{(C2)} \quad u(T) &= \gamma(T), & \text{(C5)} \quad m_1 &= \mathbf{S}w + \bar{m}_0, \\ \text{(C3)} \quad \gamma &\in \partial \mathbf{F}[m_2], & \text{(C6)} \quad m_1 &= m_2, \\ \text{(C4)} \quad P &\in \partial \phi[D], & \text{(C7)} \quad D &= \mathbf{A}w. \end{aligned}$$

We show in the next lemma that the conditions (C1-C7) are necessary and sufficient optimality conditions for  $(\mathfrak{P})$  and  $(\mathfrak{D})$ .

**Lemma 4.5.1.** *We have that  $(m_1, w, m_2, D) \in \mathcal{C}$  and  $(u, \gamma, P) \in \mathcal{K}$  are respectively solutions of  $(\mathfrak{P})$  and  $(\mathfrak{D})$  if and only if the conditions (C1-C7) hold.*

*Proof.* Let  $(m_1, w, m_2, D) \in \mathcal{C}$  and  $(u, \gamma, P) \in \mathcal{K}$ . We define the two quantities  $a$  and  $b$  as follows:

$$a = \mathcal{F}(m_1, w, m_2, D) + \mathcal{F}^*(-\mathcal{A}^*(u, \gamma, P)) + \langle (m_1, w, m_2, D), \mathcal{A}^*(u, \gamma, P) \rangle,$$

$$b = \mathcal{G}(\mathcal{A}(m_1, w, m_2, D)) + \mathcal{G}^*(u, \gamma, P) - \langle \mathcal{A}(m_1, w, m_2, D), (u, \gamma, P) \rangle.$$

By Theorem 4.4.3,  $(m_1, w, m_2, D) \in \mathcal{C}$  and  $(u, \gamma, P) \in \mathcal{K}$  are respectively solutions of  $(\mathfrak{P})$  and  $(\mathfrak{D})$  if and only if  $a + b = 0$ . Then we have the following decomposition

$$a = \sum_{(s,x) \in \mathcal{T} \times S} a_1(t, x) + \sum_{x \in S} a_2(x) + \sum_{s \in \bar{\mathcal{T}}} a_3(s) + \sum_{t \in \mathcal{T}} a_4(t),$$

$$b = \sum_{t \in \mathcal{T}} b_1(t) + \sum_{s \in \bar{\mathcal{T}}} b_2(s) + b_3(s),$$

where

$$a_1(t, x) := \tilde{\ell}[m_1, w](t, x) + \chi_{Q_{t,x}}((\gamma - u)(t, x), (-\mathbf{A}^*P - \mathbf{S}^*u)(t, x)) \\ + \langle m_1(t, x), (u - \gamma)(t, x) \rangle + \langle w(t, x), (\mathbf{A}^*P + \mathbf{S}^*u)(t, x) \rangle,$$

$$a_2(x) := \chi_{Q_{T,x}}((\gamma - u)(T, x)) + \langle m_1(T, x), (u - \gamma)(T, x) \rangle,$$

$$a_3(s) := \mathbf{F}[m_2](s) + \mathbf{F}^*[\gamma](s) - \langle m_2(s), \gamma(s) \rangle,$$

$$a_4(t) := \phi[D](t) + \phi^*[P](t) - \langle D(t), P(t) \rangle,$$

$$b_1(t) := \chi((\mathbf{A}w - D)(t)) - \langle P(t), (\mathbf{A}w - D)(t) \rangle,$$

$$b_2(s) := \chi((\mathbf{S}w - m_1 + \bar{m}_0)(s)) - \langle u(s), (\mathbf{S}w - m_1 + \bar{m}_0)(s) \rangle,$$

$$b_3(s) := \chi((m_1 - m_2)(s)) - \langle \gamma(s), (m_1 - m_2)(s) \rangle,$$

for any  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ . By the Fenchel-Young inequality,

$$a_1(s, x) \geq 0, \quad a_2(x) \geq 0, \quad a_3(s) \geq 0, \quad a_4(t) \geq 0, \\ b_1(t) \geq 0, \quad b_2(s) \geq 0, \quad b_3(s) \geq 0.$$

Then  $a + b = 0$  if and only if

$$a_1(s, x) = 0, \quad a_2(x) = 0, \quad a_3(s) = 0, \quad a_4(t) = 0, \\ b_1(t) = 0, \quad b_2(s) = 0, \quad b_3(s) = 0. \tag{4.12}$$

By Lemma 4.3.2 we have that (C1) holds if and only if  $a_1(s, x) = 0$  and it is obvious that (C2-C7) holds if and only if  $a_2(x) = a_3(s) = a_4(t) = b_1(t) = b_2(s) = b_3(s) = 0$ . Then the conditions (C1-C7) hold if and only if (4.12) holds, which concludes the proof.  $\square$



**Proposition 4.5.2.** *Let  $(m_1, \pi, u, \gamma, P) \in \mathcal{R} \times \mathcal{K}$  be a solution to (MFG) and let*

$$w(t, x, \cdot) = m_1(t, x)\pi(t, x, \cdot), \quad m_2 = m_1, \quad D = \mathbf{A}w,$$

for any  $(t, x) \in \mathcal{T} \times S$ . Then  $(m_1, w, m_2, D)$  and  $(u, \gamma, P)$  are respectively solutions to  $(\mathfrak{P})$  and  $(\mathfrak{D})$ . Moreover,  $(m_1, w)$  is solution to  $(\tilde{P})$ ,  $(m_1, \pi)$  is solution to  $(P)$ , and  $(\gamma, P)$  is solution to  $(\tilde{D})$ .

*Proof.* The conditions (C1-C7) are obviously satisfied. It immediately follows from Lemma 4.5.1 that  $(m_1, w, m_2, D)$  and  $(u, \gamma, P)$  are optimal for  $(\mathfrak{P})$  and  $(\mathfrak{D})$ . The optimality of  $(m_1, w)$  and  $(m_1, \pi)$  is then deduced from the proof of Lemma 4.3.3. The optimality of  $(\gamma, P)$  is a consequence of Lemma 4.4.4.  $\square$

For any  $(m, w) \in \mathcal{R}$ ,  $(u, \gamma, P) \in \mathcal{K}$  we define the set  $\pi[m, w, u, \gamma, P]$  of controls  $\pi \in \Delta$  satisfying

$$\pi(t, x, \cdot) = w(t, x, \cdot)/m(t, x)$$

if  $m(t, x) > 0$  and

$$\pi(t, x, \cdot) \in \arg \min_{\rho \in \Delta(S)} \ell(t, x, \rho) + \sum_{y \in S} \rho(y)(P(t)\alpha(t, x, y) + u(t+1, y))$$

if  $m(t, x) = 0$ , for any  $(t, x) \in \mathcal{T} \times S$ . Note that for any  $\pi \in \pi[m, w, u, \gamma, P]$ , we have  $w(t, x, \cdot) = m(t, x)\pi(t, x, \cdot)$ , for any  $(t, x) \in \mathcal{T} \times S$ . We have now the following converse property to Proposition 4.5.2.

**Proposition 4.5.3.** *Let  $(m_1, w, m_2, D)$  and  $(u, \gamma, P)$  be respectively solutions to  $(\mathfrak{P})$  and  $(\mathfrak{D})$ . Let  $\hat{u} = \mathbf{U}[\gamma, P]$  and let  $\pi \in \pi[m, w, \hat{u}, \gamma, P]$ . Then  $(m, \hat{\pi}, \hat{u}, \gamma, P)$  is a solution to (MFG).*

*Proof.* By Lemma 4.4.4,  $(\hat{u}, \gamma, P)$  is a solution to  $(D)$ . The pairs  $(m_1, w, m_2, D)$  and  $(\hat{u}, \gamma, P)$  are solutions to  $(\mathfrak{P})$  and  $(\mathfrak{D})$ , respectively, therefore they satisfy conditions (C1-C7), by Lemma 4.5.1. Equations (MFG,iii-v) are then obviously satisfied. By definition,  $\hat{u}$  satisfies (MFG,i). Finally, (MFG,ii) is satisfied, by condition (C1) and by definition of the set  $\pi[m, w, u, \gamma, P]$ . It follows that  $(m, \pi, \hat{u}, \gamma, P) \in \mathcal{R} \times \mathcal{K}$  is solution to (MFG).  $\square$

Since the existence of solutions to  $(\mathfrak{P})$  and  $(\mathfrak{D})$  has been established in Lemmas 4.3.3 and 4.4.4, we have the following corollary.

**Corollary 4.5.4.** *There exists a solution to (MFG).*

We finish this section with a uniqueness result.

**Proposition 4.5.5.** *Let  $(m, \pi, u, \gamma, P)$  and  $(m', \pi', u', \gamma', P')$  be two solutions to the coupled system (MFG). Assume that  $F$  and  $\phi$  are differentiable with respect to their second variable. Then  $(u, \gamma, P) = (u', \gamma', P')$ . If moreover, for any  $(t, x) \in \mathcal{T} \times S$ ,  $\ell(t, x, \cdot)$  is strictly convex, then  $(m, \pi) = (m', \pi')$  and thus (MFG) has a unique solution.*

*Proof.* It follows from Proposition 4.5.2 that  $(m, w := m\pi, m, D := \mathbf{A}w)$  is a solution to  $(\mathfrak{P})$  and that  $(u, \gamma, P)$  and  $(u', \gamma', P')$  are solutions to  $(\mathfrak{D})$ . Thus by Lemma 4.5.1, the conditions (C3) and (C4) are satisfied, both for  $(m, w, m, D)$  and  $(u, \gamma, P)$  and for  $(m, w, m, D)$  and  $(u', \gamma', P')$ , which implies that  $\gamma = \nabla \mathbf{F}[m] = \gamma'$  and  $P = \nabla \phi[D] = P'$ . It further follows that  $u = \mathbf{U}[\gamma, P] = \mathbf{U}[\gamma', P'] = u'$ .

If moreover  $\ell(t, x, \cdot)$  is strictly convex for any  $(t, x) \in \mathcal{T} \times S$  then the minimal argument in (MFG,ii) is unique, which implies that  $\pi = \pi'$  and finally that  $m = m^\pi = m^{\pi'} = m'$ .  $\square$

## 4.6 Numerical methods

In this section we investigate the numerical resolution of the problems  $(\mathfrak{P})$  and  $(\mathfrak{D})$ . We investigate different methods: primal-dual proximal algorithms, ADMM and ADM-G. For all methods, it is assumed that the prox operators (defined below) of  $\ell(t, x, \cdot)$ ,  $F(t, \cdot)$  and  $\phi(t, \cdot)$  are known. We explain in the Appendix 4.8 how to calculate the prox of  $\ell$  (and the nonlinear proximator based on the entropy function) in the special case where  $\ell$  is linear on its domain. We explain in Section 4.6.4 how to recover a solution to  $(\text{MFG})$ .

### 4.6.1 Notations

Let  $X_1$  be a subset of  $\mathbb{R}^d$ , let  $\bar{X}_1$  denote its closure. Let  $f: \bar{X}_1 \rightarrow \mathbb{R}$ . Assume that the following assumption holds true.

**Assumption 4.6.1.** *The set  $\bar{X}_1$  is convex and the map  $f$  is continuous and 1-strongly convex on  $\bar{X}_1$ . There exists an open subset  $X_2$  containing  $X_1$  such that  $f$  can be extended to a differentiable function on  $X_2$ .*

We define then the Bregman distance  $d_f: X_1 \times X_1 \rightarrow \mathbb{R}$  by

$$d_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle.$$

If  $f$  is the Euclidean distance  $\frac{1}{2}|\cdot|^2$ , then  $d_f(x, y) = \frac{1}{2}|x - y|^2$ .

Given a l.s.c., convex and proper function  $g: \mathbb{R}^d \rightarrow \mathbb{R}$ , we define its proximal operator  $\text{prox}_g: \mathbb{R}^d \rightarrow \mathbb{R}^d$  as follows:

$$\text{prox}_g(x) = \arg \min_{y \in \mathbb{R}^d} \frac{1}{2}|x - y|^2 + g(y).$$

For any non-empty, convex and closed  $K \subseteq \mathbb{R}^d$ , we define the projection operator  $\text{proj}_K$  of  $x \in \mathbb{R}^d$  on  $K$  by

$$\text{proj}_K(x) = \text{prox}_{\chi_K}(x).$$

Finally, we denote  $\bar{\alpha}(t) = \sum_{(x,y) \in \mathcal{S} \times \mathcal{S}} \alpha(t, x, y)^2$  for any  $t \in \mathcal{T}$ .

### 4.6.2 Primal-dual proximal algorithms

In this subsection we present the primal-dual algorithms proposed by A. Chambolle and T. Pock in [45] and [46]. For the sake of simplicity, we denote by  $x$  the primal variable  $(m_1, w, m_2, D)$  and by  $y$  the dual variable  $(u, \gamma, P)$ . The primal-dual algorithms rely on the following saddle-point problem

$$\min_{x \in \mathcal{C}} \max_{y \in \mathcal{K}} \mathcal{L}(x, y) := \mathcal{F}(x) - \mathcal{G}^*(y) + \langle \mathcal{A}x, y \rangle, \quad (4.13)$$

which is equivalent to problem  $(\mathfrak{P})$  (defined in the proof of Theorem 4.4.3). Let  $\mathcal{C}_1$  and  $\mathcal{K}_1$  be two subsets of  $\mathcal{C}$  and  $\mathcal{K}$ , respectively. Let  $f: \bar{\mathcal{C}}_1 \rightarrow \mathbb{R}$  and let  $g: \bar{\mathcal{K}}_1 \rightarrow \mathbb{R}$  satisfy Assumption 4.6.1.

For any  $\tau, \sigma > 0$  and for any  $(x', y') \in \mathcal{C} \times \mathcal{K}$  we define:

Iteration  $(\hat{x}, \hat{y}) = \mathcal{S}_{\tau, \sigma}[d_f, d_g](x', y')$ ,

$$\begin{cases} \text{(i)} & \hat{x} = \arg \min_{x \in \mathcal{C}_1} \mathcal{F}(x) + \langle x, \mathcal{A}^*y' \rangle + \frac{1}{\tau}d_f(x, x'), \\ \text{(ii)} & \tilde{x} = 2\hat{x} - x', \\ \text{(iii)} & \hat{y} = \arg \min_{y \in \mathcal{K}_1} \mathcal{G}^*(y) - \langle \mathcal{A}\tilde{x}, y \rangle + \frac{1}{\sigma}d_g(y, y'). \end{cases} \quad (4.14)$$

Then we define the following algorithm.

---

**Algorithm 2** Chambolle-Pock
 

---

Choose  $\sigma, \tau > 0$  and  $(x^0, y^0) \in \mathcal{C} \times \mathcal{K}$   
**for**  $0 \leq k < N$  **do**  
   Compute  $(x^{k+1}, y^{k+1}) = \mathcal{S}_{\tau, \sigma}[d_f, d_g](x^k, y^k)$ .  
**end for**  
**return**  $(x^N, y^N)$ .

---

**Theorem 4.6.2.** *Let  $\tau, \sigma > 0$  be such that  $\tau\sigma\|\mathcal{A}\|^2 < 1$ , where  $\|\mathcal{A}\|$  denotes the operator norm of  $\mathcal{A}$  (for the Euclidean norm). Assume that  $\text{dom}(\mathcal{F}) \subseteq \bar{\mathcal{C}}_1$  and  $\text{dom}(\mathcal{G}) \subseteq \bar{\mathcal{K}}_1$ . Assume that the iteration (4.14) is well-defined, that is, the minimal arguments in (i) and (iii) exist. Let  $(x^k, y^k)_{k \in \mathbb{N}}$  denote the sequence generated by the algorithm. For any  $k \in \mathbb{N}$  we set*

$$\bar{x}^k = \frac{1}{k} \sum_{n=0}^k x^n, \quad \text{and} \quad \bar{y}^k = \frac{1}{k} \sum_{n=0}^k y^n. \quad (4.15)$$

Let  $(x, y) \in \mathcal{C} \times \mathcal{K}$ . Then the following holds:

1. The sequence  $(\bar{x}^k)_{k \in \mathbb{N}}$  converges to a solution of  $(\tilde{P})$  and the sequence  $(\bar{y}^k)_{k \in \mathbb{N}}$  converges to a solution of  $(D)$ . In addition the saddle-point gap is such that

$$\mathcal{L}(\bar{x}^k, y) - \mathcal{L}(x, \bar{y}^k) \leq \frac{1}{k} \left( d_f(x, \bar{x}^k)/\tau + d_g(y, \bar{y}^k)/\sigma - \langle \mathcal{A}(x - x^0), (y - y^0) \rangle \right). \quad (4.16)$$

2. If  $f$  and  $g$  are the Euclidean distance  $\frac{1}{2}|\cdot|^2$ , then the sequence  $(x^k)_{k \in \mathbb{N}}$  converges to a solution of  $(\tilde{P})$  and the sequence  $(y^k)_{k \in \mathbb{N}}$  converges to a solution of  $(D)$ .

*Proof.* Point 1 holds as a direct application of [46, Theorem 1]. Point 2 holds as a direct application of [45, Theorem 1], applied with  $\theta = 1$ .  $\square$

**Remark 4.6.3.** *Fix  $(x, y)$ , solution to (4.13). Let  $(\hat{x}, \hat{y}) \in \mathcal{C} \times \mathcal{K}$ . Then we have that  $0 \leq \delta(\hat{x}) := \mathcal{L}(\hat{x}, y) - \mathcal{L}(x, y)$  and  $0 \leq \delta'(\hat{y}) := \mathcal{L}(x, \hat{y}) - \mathcal{L}(x, y)$ , with equality if  $\hat{x}$  (resp.  $\hat{y}$ ) is a primal (resp. dual) solution. These measures of optimality (for the saddle-point problem) trivially satisfy*

$$0 \leq \delta(\hat{x}) + \delta'(\hat{y}) = \mathcal{L}(\hat{x}^k, y) - \mathcal{L}(x, \hat{y}^k), \quad (4.17)$$

for which an upper-bound is provided by (4.16).

**Lemma 4.6.4.** *Let  $a = \max_{t \in \mathcal{T}} \bar{\alpha}(t)$ . Then  $\|\mathcal{A}\| \leq \sqrt{\max\{n + a, 4\}}$ , where  $n$  is the cardinal of the set  $S$ .*

*Proof.* For any  $(m_1, w, m_2, D) \in \mathcal{C}$ , we have

$$\begin{aligned} |\mathcal{A}(m_1, w, m_2, D)|^2 &\leq |\mathcal{S}w - m_1|^2 + |m_1 - m_2|^2 + |\mathcal{A}w - D|^2 \\ &\leq (\|\mathcal{A}\| + \|\mathcal{S}\|)|w|^2 + 4|m_1|^2 + 2|m_2|^2 + 2|D|^2. \end{aligned}$$

We have  $\|\mathcal{A}\| \leq a$  and  $\|\mathcal{S}\| \leq n$ , which concludes the proof.  $\square$

### Euclidean distance

Now we explicit the update rule (4.14) in the case where  $f$  and  $g$  are both equal to the Euclidean distance  $\frac{1}{2}|\cdot|^2$  and defined on  $\mathcal{C}$  and  $\mathcal{K}$  respectively. In this situation (i,4.14) and (iii,4.14) can be expressed via proximal operators:

$$\begin{cases} \text{(i)} & \hat{x} = \text{prox}_{\tau\mathcal{F}}(x' - \tau\mathcal{A}^*y'), \\ \text{(iii)} & \hat{y} = \text{prox}_{\sigma\mathcal{G}^*}(y' + \tau\mathcal{A}\tilde{x}). \end{cases} \quad (4.18)$$

Now we detail the computation of the proximal steps in the above algorithm.

**Primal step.** For any  $x = (x_1, x_2, x_3, x_4) \in \mathcal{C}$ , we have by Moreau's identity

$$\text{prox}_{\tau\mathcal{F}}(x) = x - \tau \text{prox}_{\mathcal{F}^*/\tau}(x/\tau).$$

As a consequence of (4.11), the proximal operator of  $\mathcal{F}^*$  is given by

$$\text{prox}_{\mathcal{F}^*}(x) = \arg \min_{x' \in \mathcal{C}} \frac{1}{2}|x - x'|^2 + \chi_Q(x'_1, x'_2) + \sum_{s \in \bar{\mathcal{T}}} F^*(s, x'_3(s)) + \sum_{t \in \bar{\mathcal{T}}} \phi^*(t, x'_4(t)).$$

Then (i,4.18) is given by

$$\begin{aligned} (\hat{m}_1, \hat{w}) &= (m'_1 - \tau(\gamma' - u'), w' - \tau(\mathbf{A}^*P' + \mathbf{S}^*u')) \\ &\quad - \tau \text{proj}_Q(m'_1/\tau - \gamma' + u'), w'/\tau - \mathbf{A}^*P' - \mathbf{S}^*u'), \end{aligned} \quad (4.19)$$

and for any  $(t, s) \in \mathcal{T} \times \bar{\mathcal{T}}$ ,

$$\begin{aligned} \hat{m}_2(s) &= m'_2(s) + \tau\gamma'(s) - \tau \text{prox}_{F^*(s)/\tau}(m'_2(s)/\tau + \gamma'(s)), \\ \hat{D}(t) &= D'(t) + \tau P'(t) - \tau \text{prox}_{\phi^*(t)/\tau}(D'(t)/\tau + P'(t)). \end{aligned} \quad (4.20)$$

**Dual step.** It follows from (4.9) that  $\text{prox}_{\sigma\mathcal{G}^*}(y_1, y_2, y_3) = (y_1 + \sigma\bar{m}_0, y_2, y_3)$ . Then (iii,4.18) is given by

$$\hat{u} = u' + \sigma(\mathbf{S}\tilde{w} - \tilde{m}_1 + \bar{m}_0), \quad \hat{\gamma} = \gamma' + \sigma(\tilde{m}_1 - \tilde{m}_2), \quad \hat{P} = P' + \sigma(\mathbf{A}\tilde{w} - \tilde{D}).$$

### Kullback-Leibler divergence

In this section we slightly modify the Euclidean framework above. Instead of considering a Euclidean distance  $d_f$  in (i,4.14), we consider an entropy based Bregman distance called Kullback-Leibler divergence. Let us define

$$\begin{aligned} \mathcal{C}_1 &= \left\{ (m_1, w, m_2, D) \in \mathcal{C} \mid 0 < m_1(t, x) \leq 1, 0 < w(t, x, y) \leq 1 \right\}, \\ \mathcal{C}_2 &= \left\{ (m_1, w, m_2, D) \in \mathcal{C} \mid 0 < m_1(t, x) < 2, 0 < w(t, x, y) < 2 \right\}. \end{aligned}$$

For any  $(m_1, w, m_2, D) \in \bar{\mathcal{C}}_1$ , we define

$$\begin{aligned} f(m_1, w, m_2, D) &= \sum_{(s,x) \in \bar{\mathcal{T}} \times S} m_1(s, x) \ln(m_1(s, x)) \\ &\quad + \sum_{(t,x,y) \in \mathcal{T} \times S^2} w(t, x, y) \ln(w(t, x, y)) + \frac{1}{2}|(m_2, D)|^2. \end{aligned} \quad (4.21)$$

Then for any  $(m_1, w, m_2, D)$  and  $(m'_1, w', m'_2, D') \in \mathcal{C}_1$ , we have

$$d_f((m_1, w, m_2, D), (m'_1, w', m'_2, D')) = d_{KL}((m_1, w), (m'_1, w')) + \frac{1}{2}|(m_2, D) - (m'_2, D')|^2,$$

where

$$d_{KL}((m_1, w), (m'_1, w')) = \sum_{(s,x) \in \bar{\mathcal{T}} \times S} m_1(s,x)(\ln(m_1(s,x)/m'_1(s,x)) - 1) + \sum_{(t,x,y) \in \mathcal{T} \times S^2} w(t,x,y)(\ln(w(t,x,y)/w'(t,x,y)) - 1). \quad (4.22)$$

As can be easily verified, the map  $f$  is 1-strongly convex on  $\bar{\mathcal{C}}_1$ . The domain of  $\mathcal{F}$  is not contained in  $\bar{\mathcal{C}}_1$  in general (as required by Theorem 4.6.2), however  $f$  is not 1-strongly convex on  $\mathcal{C}$ . This is a minor issue, since any solution to (4.13) lies in  $\bar{\mathcal{C}}_1$ , thus we can replace  $\mathcal{F}$  by  $\mathcal{F} + \chi_{\bar{\mathcal{C}}_1}$  without modifying the solution set to the problem.

Compared to the Subsection 4.6.2, the computations of (4.20) still hold. The projection step (4.19) is now replaced by

$$(\hat{m}_1, \hat{w}) = \arg \min_{(m_1, w) \in \mathcal{R}} \sum_{(t,x) \in \mathcal{T} \times S} \tilde{\ell}[m_1, w](t, x) + \langle m_1, \gamma' - u' \rangle + \langle w, \mathbf{A}^* P' + \mathbf{S}^* u' \rangle + \frac{1}{\tau} d_{KL}((m_1, w), (m'_1, w')) + \sum_{(t,x) \in \bar{\mathcal{T}} \times S} \chi_{\mathbb{R}^-}(m_1(t, x) - 1). \quad (4.23)$$

Note that it is not necessary to explicit the constraint  $w(t, x, y) \leq 1$  in the above problem; it is satisfied as a consequence of Assumption 4.2.1 and Lemma 4.3.1.

### 4.6.3 ADMM and ADM-G

We now present ADMM and ADM-G. Introducing the variables

$$(a, b) = (u - \gamma, -\mathbf{A}^* P - \mathbf{S}^* u) \quad (4.24)$$

and recalling the definition of  $Q$  and  $\mathcal{F}^*$  (see the proof of Theorem 4.4.3), the problem (D) can be written as follows:

$$\begin{aligned} & \sup_{(u, \gamma, P) \in \mathcal{K}, (a, b) \in Q} \mathcal{D}(u, \gamma, P) \\ \text{s.t.} & \begin{cases} u(s, x) - \gamma(s, x) = a(s, x) & (s, x) \in \bar{\mathcal{T}} \times S, \\ -\alpha(t, x, y)P(t) - u(t+1, y) = b(t, x, y) & (t, x, y) \in \mathcal{T} \times S^2. \end{cases} \end{aligned} \quad (4.25)$$

**Remark 4.6.5.** Let  $D_t$  and  $D_x$  be finite difference operators defined for any  $(t, x, y) \in \mathcal{T} \times S \times S$  by

$$D_t[u](t, x) = \begin{cases} u(t+1, x) - u(t, x) & \text{if } t < T, \\ -u(T, x) & \text{if } t = T, \end{cases}$$

$$D_x[u](t, x, y) = u(t+1, x) - u(t+1, y).$$

Since  $\text{dom}(\ell(t, x, \cdot)) \subseteq \Delta(S)$ , for any  $(u, b) \in \mathcal{R}$  we have that

$$\ell^*[b + \mathbf{S}^* u](t, x) = \ell^*[b + D_x u](t, x) - u(t+1, x),$$

for any  $(t, x) \in \mathcal{T} \times S$ . Then we have that  $(a, b) \in Q$  if and only if  $(\tilde{a}, \tilde{b}) \in Q$ , where

$$\tilde{a}(t, x) = a(t, x) - u(t+1, x), \quad \tilde{b}(t, x, y) = b(t, x, y) + u(t+1, x).$$

Thus the problem (4.25) can be alternatively written

$$\sup_{(u, \gamma, P) \in \mathcal{K}, (\tilde{a}, \tilde{b}) \in Q} \mathcal{D}(u, \gamma, P), \quad \text{subject to: } \begin{cases} D_t u = -\gamma - \tilde{a}, \\ D_x u = A^* P + \tilde{b}. \end{cases}$$

This problem is close to the problem studied in [14, Section 4] in the context of optimal transport theory.

Let  $r > 0$ . The Lagrangian and augmented Lagrangian associated with problem (4.25) are defined by

$$\begin{aligned} \mathcal{L} &= \mathcal{D}(u, \gamma, P) - \chi_Q(a, b) + \langle m, u - \gamma - a \rangle + \langle w, -A^* P - S^* u - b \rangle \\ \mathcal{L}_r &= \mathcal{L}(u, \gamma, P, a, b, m, w) + \frac{r}{2} |(u - \gamma - \bar{a}, -A^* P - S^* u - b)|^2, \end{aligned} \quad (4.26)$$

when evaluated at  $(u, \gamma, P, a, b, m, w)$ . Note that their definition is different from the one introduced in (4.13). We define an ADMM step which consists in the updates of  $u$ ,  $(\gamma, P)$  and  $(a, b)$  via three successive minimization steps and in the update of  $(m, w)$  via a gradient ascent step of the augmented Lagrangian:

Iteration  $(\hat{u}, \hat{\gamma}, \hat{P}, \hat{a}, \hat{b}, \hat{m}, \hat{w}) = L_r(u, \gamma, P, a, b, m, w)$ ,

$$\begin{cases} \text{(i)} & \hat{u} \in \arg \min_{u \in \mathbb{R}(\bar{\mathcal{T}} \times S)} \mathcal{L}_r(u, \gamma, P, a, b, m, w), \\ \text{(ii)} & (\hat{\gamma}, \hat{P}) \in \arg \min_{(\gamma, P) \in \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}(\mathcal{T})} \mathcal{L}_r(\hat{u}, \gamma, P, a, b, m, w), \\ \text{(iii)} & (\hat{a}, \hat{b}) \in \arg \min_{(a, b) \in Q} \mathcal{L}_r(\hat{u}, \hat{\gamma}, \hat{P}, a, b, m, w), \\ \text{(iv)} & (\hat{m}, \hat{w}) = (m, w) + r(\hat{u} - \hat{\gamma} - \hat{a}, -A^*[\hat{P}] - S^*[\hat{u}] - \hat{b}). \end{cases} \quad (4.27)$$

## ADMM

The ADMM method is given by Algorithm 3.

---

### Algorithm 3 ADMM

---

Choose  $r > 0$ ,  $(\bar{m}^0, \bar{w}^0) \in \mathcal{R}$ ,  $(u^0, \gamma^0, P^0) \in \mathcal{K}$ ,  $(a^0, b^0) \in Q$ .

Let  $v^0 = (u, \gamma, P, a, b, m, w)$ .

**for**  $0 \leq k < N$  **do**

ADMM step:  $v^{k+1} = L_r(v^k)$ ,

**end for**

**return**  $v^N$ .

---

Unlike in [14] this algorithm does not reduce to ALG2, thus we have no theoretical guarantee about the convergence. But as we will see in subsection 4.6.3, convergence results are available for ADM-G. The relation (4.27,i) is given by

$$u^{k+1} = -(m^k - \bar{m}_0 - S w^k)/r + \gamma^k + a^k - S[A^* P^{k+1} + b^k].$$

The relation (4.27,ii) can be written under a proximal form,

$$\gamma^{k+1}(s) = \text{prox}_{\mathbf{F}^*(s)/r} \left( m^k(s)/r + u^{k+1}(s) - \bar{a}^k(s) \right),$$

$$P^{k+1}(t) = \text{prox}_{\phi^*(t)/(r\bar{\alpha}(t))} \left( \mathbf{A}[w^k/r - \mathbf{S}^*u^k - b^k](t)/\bar{\alpha}(t) \right),$$

for any  $(t, s) \in \mathcal{T} \times \bar{\mathcal{T}}$ . The relation (4.27, iii) can be written as a projection step

$$(a^{k+1}, b^{k+1}) = \text{proj}_Q \left( (m^k/r + u^{k+1} - \gamma^{k+1}, w^k/r - \mathbf{A}^*P^{k+1} - \mathbf{S}^*u^{k+1}) \right). \quad (4.28)$$

### ADM-G

We explicit now the implementation of the ADM-G algorithm introduced in [78]. To fit their framework we define

$$A_1 = \begin{pmatrix} id \\ -\mathbf{S}^* \end{pmatrix}, \quad A_2 = \begin{pmatrix} -id & 0 \\ 0 & -\mathbf{A}^* \end{pmatrix}, \quad A_3 = \begin{pmatrix} -id & 0 \\ 0 & -id \end{pmatrix},$$

with appropriate dimensions, so that the constraint of problem (4.25) writes  $A_1u + A_2(\gamma, P) + A_3(a, b) = 0$ . We define

$$M = \begin{pmatrix} rA_2^*A_2 & 0 & 0 \\ rA_3^*A_2 & rA_3^*A_3 & 0 \\ 0 & 0 & id/r \end{pmatrix}, \quad H = \begin{pmatrix} rA_2^*A_2 & 0 & 0 \\ 0 & rA_3^*A_3 & 0 \\ 0 & 0 & id/r \end{pmatrix}.$$

Then we have

$$(M^*H^{-1})^{-1} = \begin{pmatrix} id & -(A_2^*A_2)^{-1}A_2A_3 & 0 \\ 0 & id & 0 \\ 0 & 0 & id \end{pmatrix}.$$

---

#### Algorithm 4 ADM-G

---

Choose  $r > 0$  and  $\xi \in (0, 1)$ . Let  $(m^0, w^0) \in \mathcal{R}$ ,  $(u^0, \gamma^0, P^0) \in \mathcal{K}$ ,  $(a^0, b^0) \in Q$ .

Let  $v^0 = (\gamma, P, a, b, m, w)$ .

**for**  $0 \leq k < N$  **do**

ADMM step:  $(\tilde{u}^{k+1}, \tilde{v}^{k+1}) = L_r(u^k, v^k)$ ,

Substitution step:  $\begin{cases} v^{k+1} = v^k + \xi(M^*H^{-1})^{-1}(\tilde{v}^k - v^k), \\ u^{k+1} = \tilde{u}^{k+1}, \end{cases}$

**end for**

**return**  $(u^N, v^N)$ .

---

**Theorem 4.6.6.** *Let  $(u^k, \gamma^k, P^k, a^k, b^k, m_1^k, w^k)_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 4, and let  $m_2^k = m_1^k$ ,  $D^k = \mathbf{A}w^k$ , for any  $k \in \mathbb{N}$ . Then the sequence  $(m_1^k, w^k, m_2^k, D^k)_{k \in \mathbb{N}}$  converges to a solution of  $(\tilde{P})$  and the sequence  $(u^k, \gamma^k, P^k)_{k \in \mathbb{N}}$  converges to a solution of  $(D)$ .*

*Proof.* By [78, Theorem 4.7] we have that  $(u^k, \gamma^k, P^k, a^k, b^k, m_1^k, w^k)_{k \in \mathbb{N}}$  converges to a saddle-point of the Lagrangian (4.26). Thus by definition of  $(m_2^k, D^k)$ , the sequence  $(m_1^k, w^k, m_2^k, D^k)_{k \in \mathbb{N}}$  converges to a solution of  $(\mathfrak{P})$  and the sequence  $(u^k, \gamma^k, P^k)_{k \in \mathbb{N}}$  converges to a solution of  $(\mathfrak{D})$ .  $\square$

**Remark 4.6.7.** *In our case the first equality of the Gaussian back substitution step in Algorithm (4) can be written*

$$\begin{aligned} v^{k+1} &= v^k + \xi(M^*H^{-1})^{-1}(\tilde{v}^k - v^k) \\ &= (\tilde{\gamma}^k - \xi(\tilde{a}^k - a^k), \tilde{P}^k - \xi(\mathbf{A}\mathbf{A}^*)^{-1}\mathbf{A}^*(\tilde{b}^k - b^k), \tilde{a}^k, \tilde{b}^k, \tilde{m}^k, \tilde{w}^k). \end{aligned}$$

The Gaussian back substitution step is thus given by

$$\gamma^{k+1} = \tilde{\gamma}^k - \xi(\tilde{a}^k - a^k), \quad (4.29)$$

$$P^{k+1} = \tilde{P}^k - \xi(\mathbf{A}\mathbf{A}^*)^{-1}\mathbf{A}^*(\tilde{b}^k - b^k), \quad (4.30)$$

$$(u^{k+1}, a^{k+1}, b^{k+1}, m^{k+1}, w^{k+1}) = (\tilde{u}^k, \tilde{a}^k, \tilde{b}^k, \tilde{m}^k, \tilde{w}^k),$$

where  $(\mathbf{A}\mathbf{A}^*)^{-1}P(t) = P(t)/\bar{\alpha}(t)$  for any  $t \in \mathcal{T}$ . Then the differences between ADM-G and ADMM can be summarized by the two corrections (4.29) and (4.30).

#### 4.6.4 Residuals

Let  $(m_1^k, w^k, m_2^k, D^k)_{k \in \mathbb{N}}$  and  $(u^k, \gamma^k, P^k)_{k \in \mathbb{N}}$  denote the two sequences generated by a numerical method. Let us consider

$$\hat{u}^k = \mathbf{U}[\gamma^k, P^k] \quad \text{and} \quad \pi^k \in \boldsymbol{\pi}[m_1^k, w^k, \hat{u}^k, \gamma^k, P^k]. \quad (4.31)$$

It was shown in Proposition (4.5.3) that if for some  $k \in \mathbb{N}$ ,  $(m_1^k, w^k, m_2^k, D^k)$  and  $(u^k, \gamma^k, P^k)$  are solutions to  $(\mathfrak{P})$  and  $(\mathfrak{D})$ , then  $(m_1^k, \pi^k, \hat{u}^k, \gamma^k, P^k)$  is a solution to  $(\text{MFG})$ .

Therefore, we look the sequence  $(m_1^k, \pi^k, \hat{u}^k, \gamma^k, P^k)_{k \in \mathbb{N}}$  as a sequence of approximate solutions to  $(\text{MFG})$ . Note that  $(\text{MFG}, \text{i})$  is exactly satisfied, by construction. We consider the residuals  $(\varepsilon_m, \varepsilon_\pi, \varepsilon_\gamma, \varepsilon_P) \in \mathcal{R} \times \mathcal{U}$  defined as follows, in order to measure the satisfaction of the remaining relations in the coupled system:

$$\left\{ \begin{array}{l} \varepsilon_\pi(t, x) = (\ell[\pi] + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*\hat{u}])(t, x) - \langle \pi(t, x), (\mathbf{A}^*P + \mathbf{S}^*\hat{u})(t, x) \rangle, \\ \varepsilon_m(s, x) = m^\pi(s, x) - m(s, x), \\ \varepsilon_\gamma(s) = m(s) - \text{proj}_{\partial \mathbf{F}^*[\gamma](s)}(m(s)), \\ \varepsilon_P(t) = Q[m, \pi](t) - \text{proj}_{\partial \phi^*[P](t)}(Q[m, \pi](t)\Delta_x), \end{array} \right.$$

for all  $(t, s, x) \in \mathcal{T} \times \bar{\mathcal{T}} \times S$ . If the residuals are null, then  $(m_1^k, \pi^k, \hat{u}^k, \gamma^k, P^k)$  is a solution to  $(\text{MFG})$ . The errors are then defined as the norms of  $\varepsilon_\pi$ ,  $\varepsilon_m$ ,  $\varepsilon_\gamma$ , and  $\varepsilon_P$ .

## 4.7 Numerical Results

In this section we provide two problems that we solve with the algorithms presented in the previous section. We set  $n = T = 50$  and we define two scaling coefficients  $\Delta_x = 1/n$



and  $\Delta_x = 1/T$ . We solve two instances of the following scaled system:

$$\left\{ \begin{array}{l} \text{(i)} \quad \begin{cases} u(t, x)/\Delta_t + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u/\Delta_t](t, x) = \gamma(t, x), \\ u(T, x) = \gamma(T, x), \end{cases} \\ \text{(ii)} \quad (\ell[\pi] + \ell^*[-\mathbf{A}^*P - \mathbf{S}^*u/\Delta_t])(t, x) = -\langle \pi(t, x), (\mathbf{A}^*P + \mathbf{S}^*u/\Delta_t)(t, x) \rangle, \\ \text{(iii)} \quad \begin{cases} m(t+1, x) = \sum_{y \in S} m(t, y)\pi(t, y, x), \\ m(0, x) = m_0(x)/\Delta_x, \end{cases} \\ \text{(iv)} \quad \gamma \in \partial \mathbf{F}[m], \\ \text{(v)} \quad P \in \partial \phi[\mathbf{Q}[m, \pi]\Delta_x]. \end{array} \right. \quad (\text{MFG}_\Delta)$$

One can show that this system is connected to two optimization problems of very similar nature as Problems  $(\mathfrak{P})$  and  $(\mathfrak{D})$ , which can be solved as described previously. For both examples,  $\ell$  is defined by

$$\ell(t, x, \rho) = \sum_{y \in S} \rho(y)\beta(t, x, y) + \chi_{\Delta(S_x)}(\rho), \quad \beta(t, x, y) = \left( (y-x)\frac{\Delta_x}{\Delta_t} \right)^2 / 4, \quad (4.32)$$

where  $\beta$  is a displacement cost from state  $x$  to state  $y$  and  $S_x = \{x, x-1, x+1\} \cap \{0, \dots, n-1\}$  is an admissible transition set. In Appendix 4.8 the reader can find detailed computations of the Euclidean projection (4.19) (Subsection 4.8.1) and the computation of (4.23) (Subsection 4.8.2) for this particular choice of running cost  $\ell$ . The notion of residuals that we use in the following is adapted from Section 4.6.4 to the scaled system  $(\text{MFG}_\Delta)$ . In all subsequent graphs, the state space is represented by  $\{0, \Delta_x, \dots, 1\}$  and the set of time steps by  $\{0, \Delta_t, \dots, 1\}$ .

### 4.7.1 Example 1

In our first example, we take  $\phi = 0$  and  $\alpha = 0$ . We consider a potential  $\mathbf{F}$  of the form  $\mathbf{F}[m] = \mathbf{F}_1[m] + \mathbf{F}_2[m]$ , where

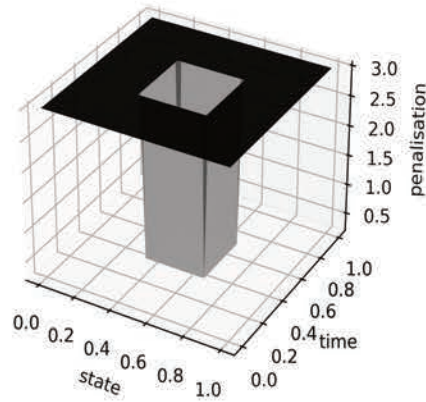
$$\mathbf{F}_1[m](s) = |m(s)|^2/2, \quad \mathbf{F}_2[m](s) = \chi_{[0, \eta(s)]}(m(s)), \quad (4.33)$$

and where  $\eta \in \mathbb{R}_+(\bar{\mathcal{T}} \times S)$  is given by

$$\eta(s, x) := \begin{cases} 0.5 & \text{if } T/3 \leq s \leq 2T/3 \text{ and } n/3 \leq x \leq 2n/3, \\ 3 & \text{else,} \end{cases}$$

for any  $(s, x) \in \bar{\mathcal{T}} \times \mathbb{R}(S)$ . We refer to  $\mathbf{F}_1$  as the soft congestion term and to  $\mathbf{F}_2$  as the hard congestion term. We call narrow region the set of points  $(s, x)$  for which  $\eta(s, x) = 0.5$  and we call spacious region the set of points for which  $\eta(s, x) = 3$ . In this situation the state of an agent represents its physical location on the interval  $[0, 1]$ . Each agent aims at minimizing her displacement cost and avoids congestion as time evolves from time  $t = 0$  to  $t = 1$ . The congestion term is linked to  $\eta$  by the following relation (see Remark 4.2.2):

$$\gamma \in \partial \mathbf{F}[m] = \nabla \mathbf{F}_1[m] + \partial \mathbf{F}_2[m] = m + N_{[0, \eta]}(m).$$

Figure 4.1: Hard constraint  $\eta$ 

As shown on the graphs below, we have two regimes at equilibrium: in the spacious regions  $\gamma$  plays the role of a classical congestion term and  $\gamma = \nabla F_1[m]$ . In the narrow region the constraint is binding,  $\gamma$  is such that the constraint  $m \in [0, \eta]$  is satisfied at the equilibrium and is maximal for the dual problem.

We give a representation of the solution to the mean field game system in Figure 4.2. Since it is hard to give a graphical representation of  $\pi$ , we give instead a graph of the mean displacement  $v$ , defined by

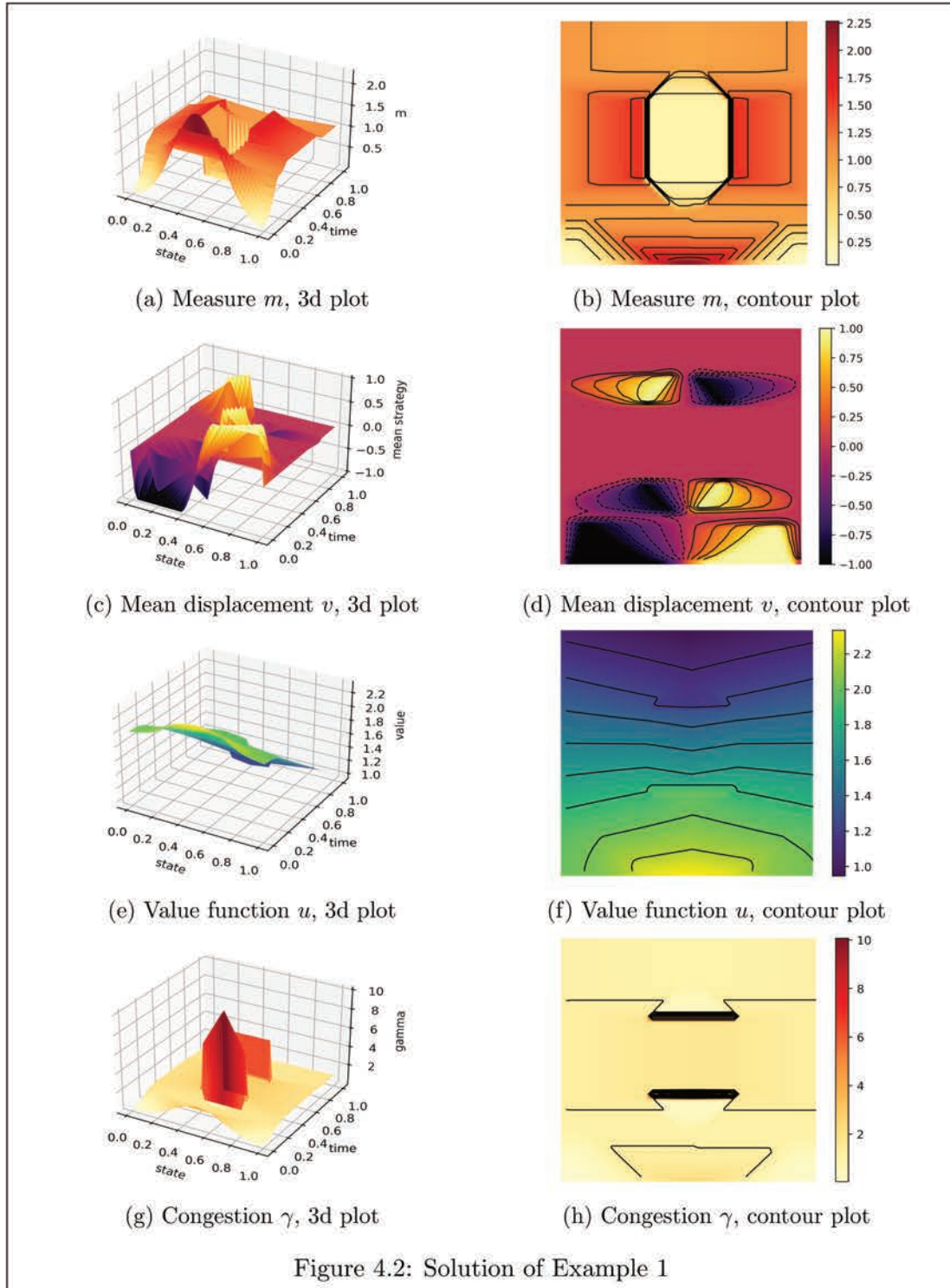
$$v(t, x) = \sum_{y \in S} \pi(t, x, y)(y - x), \quad \forall (t, x) \in \mathcal{T} \times S.$$

For each variable, a 3D representation of the graph and a 2D representation of the contour plots are provided. For the contour plots, the horizontal axis corresponds to the state space and the vertical axis to the time steps (to be read from the bottom to the top).

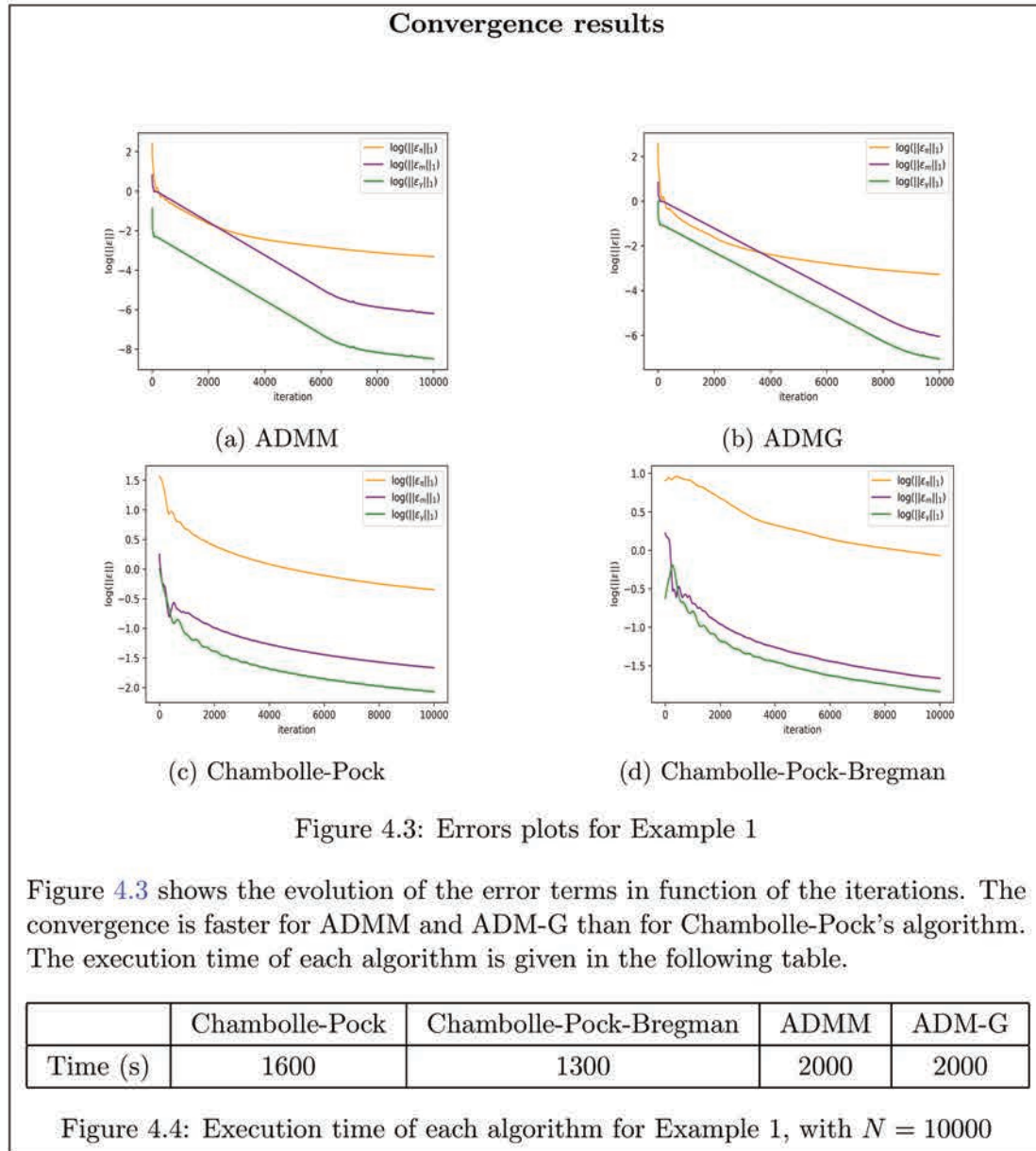
Let us comment the results. We start with the interpretation of the measure  $m$  and the mean displacement  $v$ . At the beginning of the game, the distribution of players is given by the initial condition  $m(0) = \bar{m}_0$ . Then the players spread since they are in the spacious region to avoid congestion.

Thus the mean displacement is negative on the left (black region) and positive on the right (yellow region), around  $t = 0$ . The distribution becomes uniform after some time. In a second phase, the agents move again towards the border of the state space, anticipating the narrow region. They start their displacement before entering into the narrow region due to their limited speed and displacement cost. Then we are in a stationary regime (purple region), the mean displacement is null and the mass does not vary until the end of the narrow region. At the end of the narrow region, the agents spread again along the state axis and the distribution  $m$  becomes uniform.

We now interpret the value  $u$  and the congestion  $\gamma$ . The value function has to be interpreted backward in time. At the end of the game, the terminal condition imposes that the value is equal to the congestion. Since the congestion is positive and accumulates backward in the value function (which can be seen in the dynamic programming equation), the value function increases backward in time. At the end and at the beginning of the narrow region we observe irregularities in the value function due to the irregularities of the congestion term  $\gamma$ . But the impact on the value function is limited due to the trade-off between the variables  $u$  and  $\gamma$  in the dual problem. At the beginning



of the game the value function is higher at the middle of the space because of the initial distribution of players that are accumulated at this point. The congestion term  $\gamma$  is high enough at the beginning of the narrow region to ensure that the constraint on the distribution of players is satisfied at this point. Then  $\gamma$  is high enough at the end of the narrow region to ensure that the constraint on the distribution of players is satisfied for all time indices  $T/3 \leq s \leq 2T/3$ . At the exception of these two moments,  $\gamma$  plays the role of a classical congestion term.



### 4.7.2 Example 2

Here we assume that  $F = 0$ . In this situation the state of an individual agent represents a level of stock. We set  $\alpha(t, x, y) = y - x$ ; it represents the quantity bought in order to "move" from  $x$  to  $y$ . Therefore the variable  $D$  (used in the primal problem) is the average

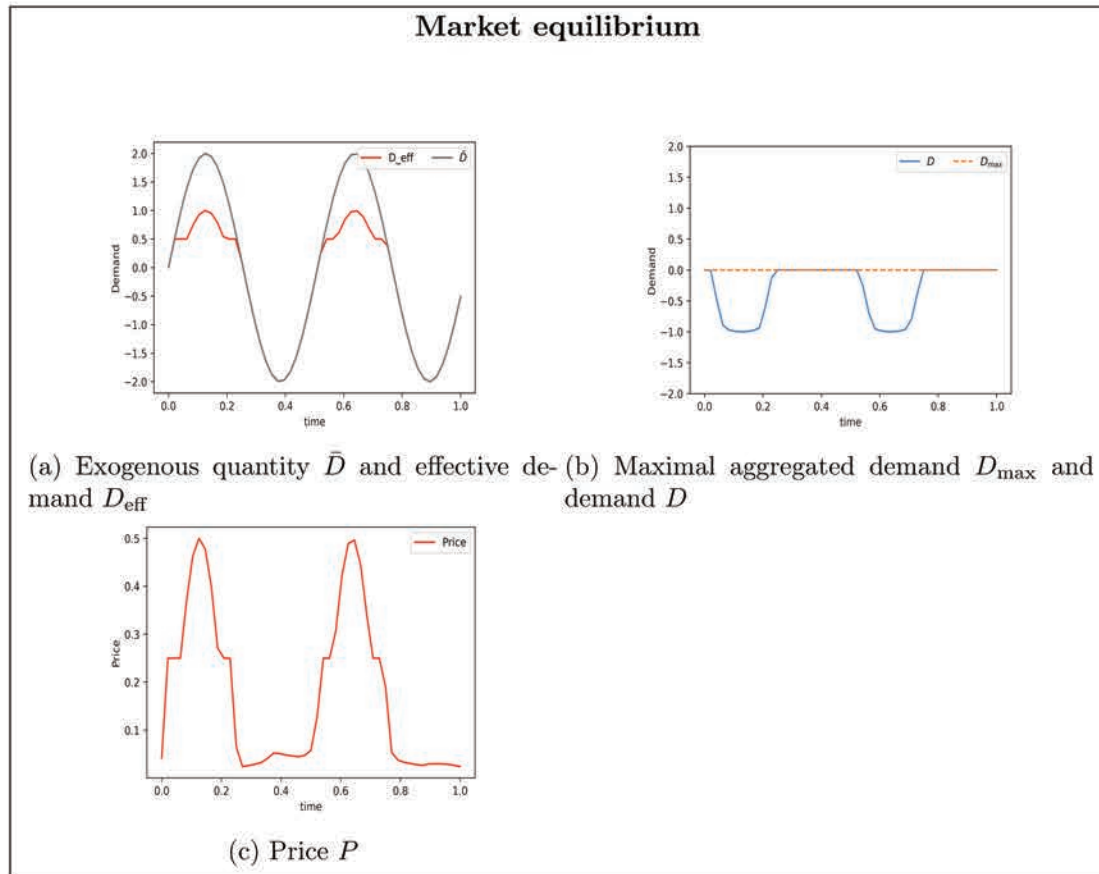
quantity which is bought; it has to be understood as a demand, since at equilibrium,

$$D(t) = \mathcal{Q}[m, \pi](t) = \sum_{(x,y) \in S^2} m(t,x)\pi(t,x,y)\alpha(t,x,y).$$

We define the potential  $\phi[D] = \phi_1[D] + \phi_2[D]$ , where

$$\phi_1[D] = \frac{1}{4}(D + \bar{D})^2, \quad \phi_2[D] = \chi_{(-\infty, D_{\max}]}(D).$$

The potential  $\phi$  is the sum of a convex and differential term  $\phi_1$  with full domain and a convex non-differentiable term  $\phi_2$ . The quantity  $\bar{D}$  is a given exogenous quantity which represent a net demand (positive or negative) to be satisfied by the agents. In this example  $\bar{D}(t) = 2 \sin(4\pi t/(T-1))$  for any  $t \in \mathcal{T}$  and  $D_{\max} = 0$ .

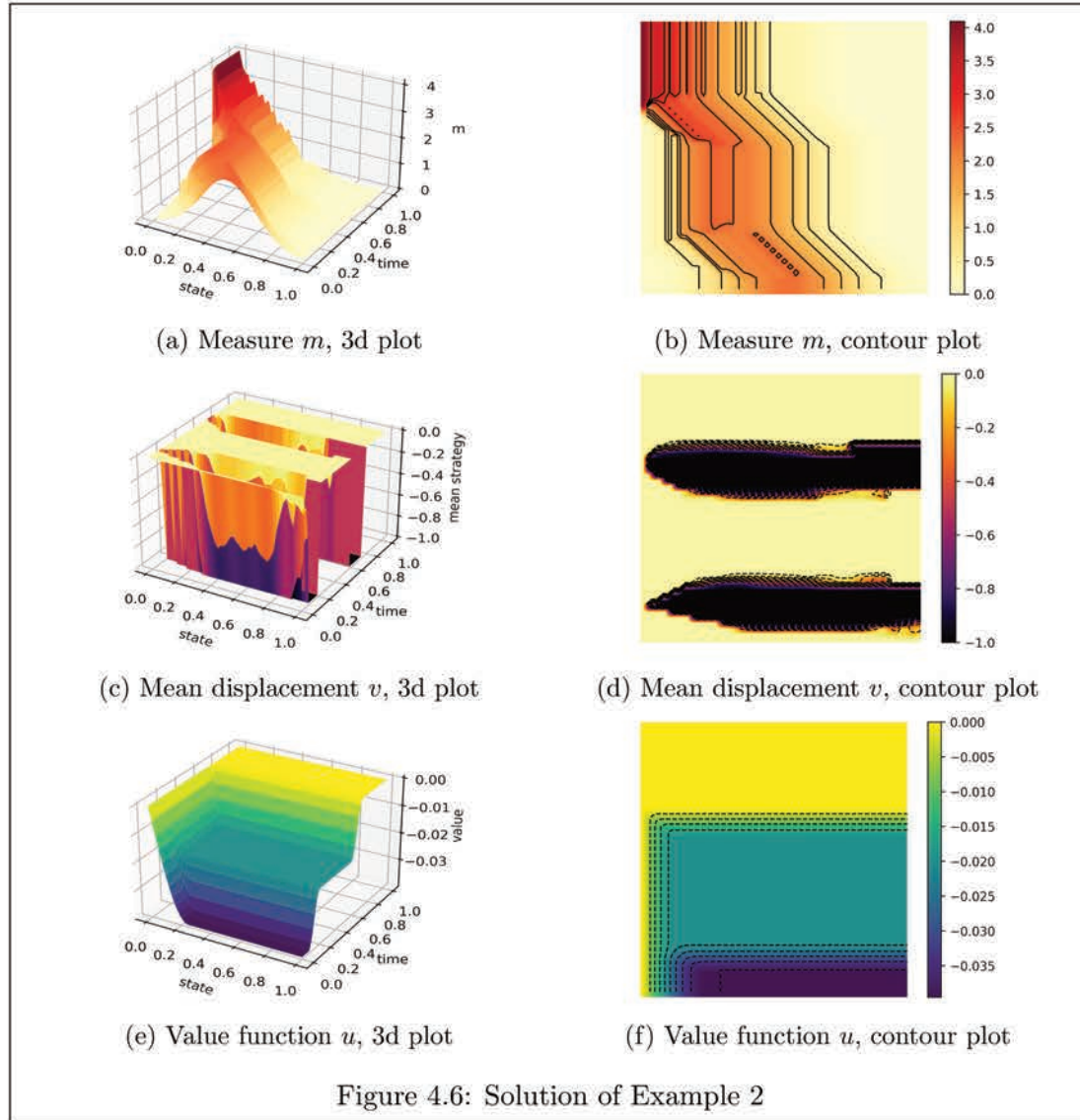


In this situation each agent faces a price and chooses to increase or deplete her stock. The price mechanism is given by

$$P(t) \in \partial\phi[D](t) = \nabla\phi_1[D] + \partial\phi_2[D] = \frac{1}{2}D_{\text{eff}}(t) + N_{(-\infty, D_{\max}]}(D(t))$$

where  $D_{\text{eff}} := D + \bar{D}$  is called the effective demand and follows two regimes.

When the constraint on the demand  $D$  is not binding, we are in a soft regime, the price plays the role of a classical price term and is given by  $P(t) = \frac{1}{2}D_{\text{eff}}(t)$ . The quantity  $\bar{D}$  is an exogenous quantity which can be positive or negative. If the quantity  $\bar{D}(t) > 0$ , the exogenous quantity is interpreted as being a demand and the agents have an incentive to deplete their stock to satisfy this demand. If  $\bar{D}(t) < 0$ , the exogenous



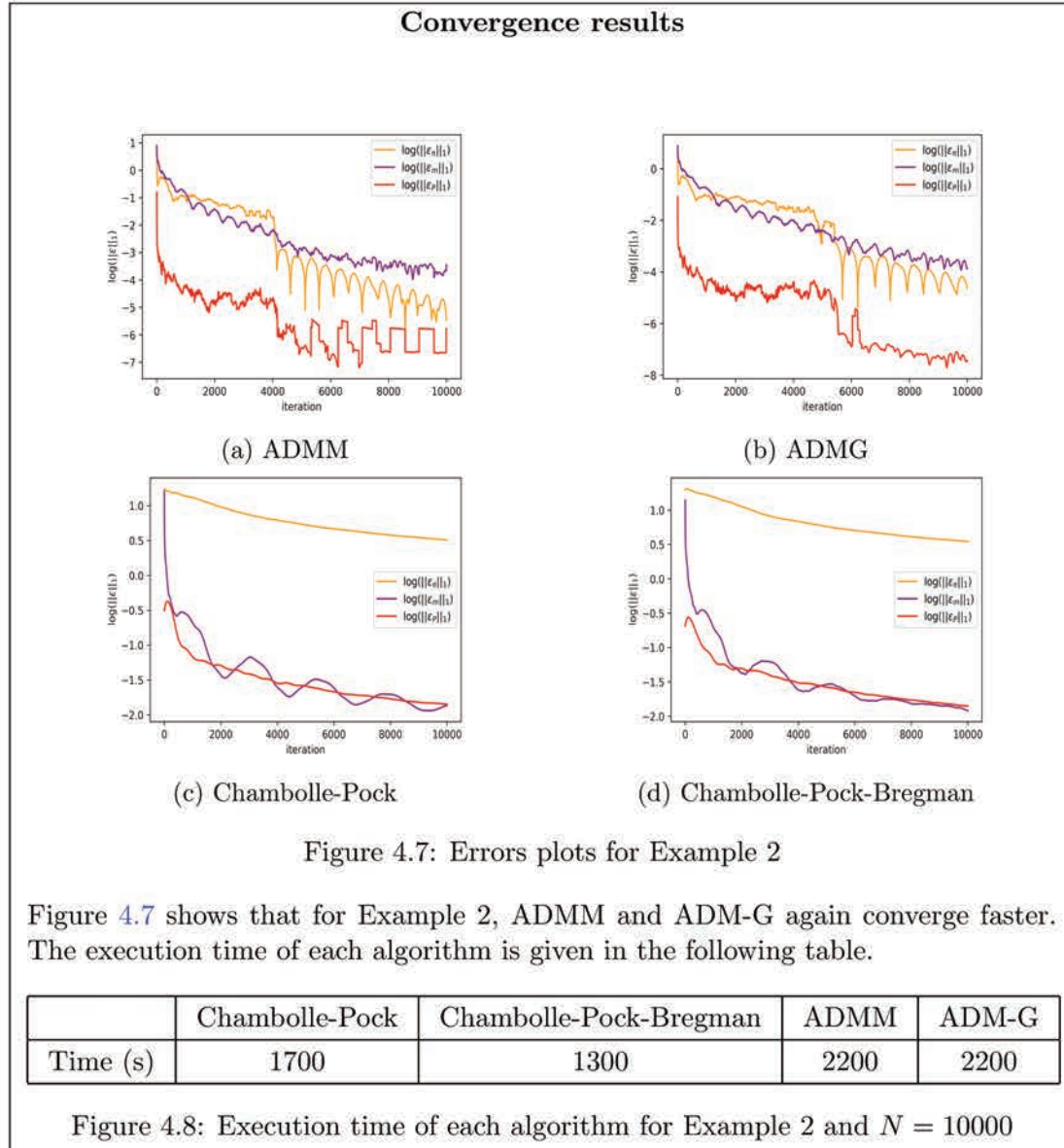
quantity is interpreted as being a supply. In the absence of a hard constraint, the agents would have interest to increase their stock to absorb this supply. When the constraint on the demand is binding, we are in a hard regime and the price plays the role of an adjustment variable so that the constraint  $D(t) \leq D_{max}$  is satisfied and is maximal for the dual problem.

In the case where it is not profitable to buy or sell, we have that  $D(t) = 0$  and thus  $D_{eff}(t) = \bar{D}(t)$ . This situation occurs when the quantity  $\bar{D}(t) \leq 0$ , since the hard constraint prevents the agents from buying on the market. On the graph this corresponds to the case where the red and the black curves coincide.

When  $\bar{D}(t) \geq 0$  we observe that the red curve is lower than the black curve meaning that a certain amount (given by the blue curve on the following graph) of the demand has been satisfied by the agents. Three effects prevent the agents from fully satisfying the demand: their level of stock, their trading cost and their depletion speed limitation.

At the optimum we observe that the demand  $D$  is indeed below the threshold  $D_{max}$ , meaning that the constraint is satisfied. We now comment the measure  $m$ , the mean displacement  $v$  and the value function  $u$ . For a given initial distribution of the level of stock, we observe that the measure  $m$  is shifted to the left with time. This means that

the agents deplete their stocks with time. This is consistent with the mean displacement  $v$  where we observe two regimes: either the agents choose to sell as much as possible or the agents choose not to sell on average. The value function  $u$  can be interpreted backward. At the end of the game the value is null due to the terminal condition. Then the higher the level of stock, the lower the value function that is to say the value function is increasing in time and decreasing in space. This comes from the definition of  $\alpha$  and the constraint  $D \leq 0$ , which implies that the price is positive.



## 4.8 Appendix

We detail here the calculation of the projection on  $Q$  and the non-linear proximity operator in (4.23), for a running cost of the form

$$\ell(t, x, \rho) = \sum_{y \in S} \rho(y) \beta(t, x, y) + \chi_{\Delta(S)}(\rho).$$

The adaptation to the case where  $\ell$  is defined by (4.32) is straightforward.

### 4.8.1 Projection on $Q$

We detail the computation of  $\text{proj}_Q$ , as it appears in (4.19) and (4.28). First notice that the projection is decoupled in space and time, then for any  $(t, x) \in \mathcal{T} \times S$  and  $(\bar{a}, \bar{b}) \in \mathbb{R} \times \mathbb{R}(S)$ , we need to compute

$$\text{proj}_{Q_{t,x}}(\bar{a}, \bar{b}) = \arg \min_{(a,b) \in Q_{t,x}} (a - \bar{a})^2/2 + \sum_{y \in S} (b(y) - \bar{b}(y))^2/2,$$

where  $Q_{t,x} = \{(a, b) \in \mathbb{R} \times \mathbb{R}(S), a + b(y) - \beta(y) \leq 0\}$ . The corresponding problem is

$$\min_{a \in \mathbb{R}} \left( (a - \bar{a})^2/2 + \min_{\substack{b \in \mathbb{R}(S) \\ b(y) \leq \beta(y) - a, \forall y \in S}} \sum_{y \in S} (b(y) - \bar{b}(y))^2/2 \right). \quad (4.34)$$

For any  $a \in \mathbb{R}$ , the solution of the inner minimization problem is given by

$$b^*(a, y) := \min\{\bar{b}(y), \beta(y) - a\}, \quad \forall y \in S.$$

Then replacing into (4.34), the minimization problem is now given by

$$\min_{a \in \mathbb{R}} g(a), \quad g(a) := (a - \bar{a})^2/2 + \sum_{y \in S} \max(0, a - \tilde{\beta}(y))^2/2,$$

where  $\tilde{\beta}(y) := \beta(y) - \bar{b}(y)$ . It is now relatively easy to minimize  $g$ . Let us sort the sequence  $(\tilde{\beta}(y))_{y \in S}$ , that is, let us consider  $(y_i)_{i \in \{0, \dots, n-1\}}$  such that  $\tilde{\beta}(y_0) \leq \dots \leq \tilde{\beta}(y_{n-1})$ . It is obvious that the function  $g$  is strictly convex and polynomial of degree 2 on each of the intervals  $(-\infty, \tilde{\beta}(y_0))$ ,  $(\tilde{\beta}(y_0), \tilde{\beta}(y_1))$ , ..., and  $(\tilde{\beta}(y_{n-1}), +\infty)$ . One can identify on which of these intervals a stationary point of  $g$  exists, by evaluating  $\partial g(\tilde{\beta}(y_i))$ , for all  $i = 0, \dots, n-1$ . Then one can obtain an analytic expression of the (unique) stationary point  $a^*$ , which minimizes  $g$ . Finally, we have  $\text{proj}_{Q_{t,x}}(\bar{a}, \bar{b}) = (a^*, b^*(a^*, \cdot))$ .

### 4.8.2 Entropic proximity operator

Here we detail the computation of the solution to (4.23). For notational purpose we set  $c_1 = \tau(-u' + \gamma')$  and  $c_2 = \tau(\beta + \mathbf{A}^*P' + \mathbf{S}^*u')$ . By definition of the running cost  $\ell$ , we have that

$$\sum_{(t,x) \in \mathcal{T} \times S} \tilde{\ell}[m_1, w](t, x) = \langle w, \beta \rangle + \chi_{\text{dom}(\tilde{\ell})}(m_1, w).$$

Problem (4.23) writes

$$\begin{aligned} & \min_{(m_1, w) \in \mathcal{R}} \langle m_1, c_1 \rangle + \langle w, c_2 \rangle + \frac{1}{\tau} d_{KL}((m_1, w), (m'_1, w')) \\ & \text{subject to: } \begin{cases} m_1(t, x) \leq 1 \\ m_1(t, x) - \sum_{y \in S} w(t, x, y) = 0. \end{cases} \end{aligned}$$

To find the solution, we define the following Lagrangian with associated multipliers  $(\lambda_1, \lambda_2) \in \mathbb{R}(\mathcal{T} \times S) \times \mathbb{R}_+(\bar{\mathcal{T}} \times S)$ :

$$\begin{aligned} \mathcal{L}(m_1, w, \lambda_1, \lambda_2) &= \langle m_1, c_1 \rangle + \langle w, c_2 \rangle + d_{KL}((m_1, w), (m'_1, w')) \\ &+ \sum_{(t,x) \in \mathcal{T} \times S} \lambda_1(t, x) \left( m_1(t, x) - \sum_{y \in S} w(t, x, y) \right) + \sum_{(s,x) \in \bar{\mathcal{T}} \times S} \lambda_2(s, x) (m_1(s, x) - 1). \end{aligned}$$



For any  $(t, s, x, y) \in \mathcal{T} \times \bar{\mathcal{T}} \times S \times S$ , a saddle point of the Lagrangian is given by the following first order conditions,

$$\begin{cases} \hat{m}_1(T, x) &= m'_1(T, x) \exp(-\lambda_2(T, x) - c_1(T, x)), \\ \hat{m}_1(t, x) &= m'_1(t, x) \exp(-\lambda_1(t, x) - \lambda_2(t, x) - c_1(t, x)), \\ \hat{w}(t, x, y) &= w'(t, x, y) \exp(\lambda_1(t, x) - c_2(t, x, y)), \\ \hat{m}_1(t, x) &= \sum_{y' \in S} \hat{w}(t, x, y'), \\ 0 &= \min \{ \lambda_2(s, x), \hat{m}_1(s, x) - 1 \}. \end{cases} \quad (4.35)$$

**Case 1:**  $\lambda_2(s, x) > 0$ . At time  $s = T$  we have that  $\hat{m}_1(s, x) = 1$ . For any  $s < T$  we have that  $\hat{m}_1(s, x) = 1$  and  $\sum_{y \in S} \hat{w}(s, x, y) = 1$  and by a direct computation we have that

$$\begin{cases} \hat{m}_1(s, x) &= 1, \\ \hat{w}(s, x, y) &= w'(s, x, y) \exp(-c_2(s, x, y)) C(s, x), \\ \lambda_1(s, x) &= \ln(C(s, x)), \\ \lambda_2(s, x) &= \ln(m'_1(s, x)/C(s, x)) - c_1(s, x), \end{cases} \quad (4.36)$$

where  $C(s, x) = \sum_{y \in S} w'(s, x, y) \exp(-c_2(s, x, y))$ .

**Case 2:**  $\lambda_2(s, x) = 0$ . At time  $s = T$  we have that  $\hat{m}_1(s, x) = m'_1(s, x) \exp(-c_1(s, x))$ . For any  $s < T$  we have by a direct computation

$$\begin{cases} \hat{m}_1(s, x) &= m'_1(s, x) C(s, x)^{-1} \exp(-c_1(s, x)), \\ \hat{w}(s, x, y) &= w'(s, x, y) C(s, x) \exp(-c_2(s, x, y)), \\ \lambda_1(s, x) &= \ln(C(s, x)), \\ \lambda_2(s, x) &= 0, \end{cases} \quad (4.37)$$

where  $C(s, x) = \left( m'_1(s, x) \exp(-c_1(s, x)) / \sum_{y \in S} w'(s, x, y) \exp(-c_2(s, x, y)) \right)^{1/2}$ .

In order to identify which of the two cases arises, one can compute a solution with formula (4.36) and check a posteriori that  $\lambda_2(s, x) > 0$ . If this is not the case, we deduce that the solution to (4.35) is given by (4.37).



## Chapter 5

# Generalized conditional gradient and learning in potential mean field games

### 5.1 Introduction

Mean field games were introduced by J.-M. Lasry and P.-L. Lions in [88, 89, 90] and M. Huang, R. Malhamé, and P. Caines in [80], to study interactions among a large population of players. Mean field games have found various applications such as epidemic control [52, 54], electricity management [7, 49], finance and banking [37, 40, 41, 56, 84], social network [12], economics [2, 74], crowd motion [85]. In these models, the nature of the interactions can be of two kinds. Interactions through the density  $m$  of players, which appear typically in epidemic or crowd motion models, will be modeled in the following by a congestion function denoted  $f$ . Interactions through the controls  $v$ , which rather appear in economics, finance or energy management models, will be modeled by a price function denoted  $\phi$ .

**Framework** In this article, we study the generalized conditional gradient algorithm to solve potential mean field game problems. We consider the continuous and finite time framework formulated in [21], consisting of a Hamilton-Jacobi-Bellman equation, a Fokker-Planck equation, and other coupling equations. We show that the generalized conditional gradient method can be interpreted as a learning procedure called fictitious play. This perspective allows us to:

1. borrow and apply classical tools from the conditional gradient theory and derive, under suitable assumptions, convergence rates for the potential cost, the different variables generated by the fictitious play algorithm, and the exploitability;
2. show that the notion of exploitability from game theory is equivalent to the notion of primal-dual gap defined (as defined in Section 5.5).

**Potential mean field games** We say that a mean field game has a convex potential formulation if the congestion and price mappings  $f$  and  $\phi$  derive from convex potentials  $F$  and  $\Phi$ . In the mean field game literature, potential (or variational) mean field games were first considered in [89]. This class of games has been widely investigated, we refer the reader to [17, 34, 38, 93, 106] for congestion interactions and [21, 67, 68, 71, 70, 72] for price interactions. A key interest of potential mean field games is that the mean field

game system stands as sufficient first order conditions for the potential control problem. This is of particular interest for numerical resolution: in such a case one expects classical optimization algorithms to be applicable.

**Algorithms** The numerical resolution of mean field games has been widely studied, see [6] for a survey. Primal-dual methods [22, 27, 28] fully use the primal-dual structure of the potential problem. The augmented Lagrangian algorithm [15, 17, 22] is a primal method based on successive minimization of the primal variable and gradient ascent step of dual variables. Other methods have been investigated such as the Sinkhorn algorithm [16].

Let us emphasize that most of the above references deal with interaction terms depending on the distribution of the states of the agents; few publications are concerned with interactions through the controls (see [5, 22]).

**Generalized conditional gradient** The generalized conditional gradient algorithm is a variant of the conditional gradient algorithm, also called Frank-Wolfe algorithm, first developed in [59]. The conditional gradient method is designed to minimize a convex objective function on a convex and compact set. The idea is to linearize the objective function at each iteration  $k \in \mathbb{N}$ , at a given point  $\bar{x}_k$ , and to find a minimizer  $x_k$  of this linearized problem. Then a new point  $\bar{x}_{k+1} = (1 - \delta_k)\bar{x}_k + \delta_k x_k$  is computed for some step size  $\delta_k \in [0, 1]$ . As we will see later, the step size  $\delta_k$  can be interpreted as a learning rate for games. A classical choice of step size is given by  $\delta_k = 2/(k + 2)$  (see [53, 81]) which yields the convergence of the objective function in  $O(1/k)$ . For a recent description of the conditional gradient algorithm, we refer to [82, Chapter 1]. In our study we consider the generalized conditional gradient algorithm (first studied in [26]), which is based on a semi-linearization of the objective function instead of a full linearization. An interesting feature of this method is that most of the existing convergence results obtained for the conditional gradient remain true for the generalized conditional gradient method. We refer to [107] for a study. We mention that the previous references deal with finite dimensional problems but these algorithms have been also investigated in infinite dimensional setting, see [26, 105, 121] respectively for studies in Hilbert, measures and Banach spaces.

**Learning and exploitability** Since most models in social science or engineering rely on Nash equilibria, one can wonder whether such equilibria can be reached if all agents follow their personal interests. Learning is thus a central question in game theory [60]. Fictitious play is a best response iterative method for solving games, introduced in [29, 108]. The idea is the following: at each step of the algorithm, for a given belief on the strategy of the others, find the best response of the players; then learn by averaging all the best responses found from the beginning of the learning procedure. An application of the fictitious play to potential games can be found in [94]. The fictitious play has been investigated in [35, 55, 77, 103]. The convergence results for learning methods can be of various forms. In potential games, one can study the convergence of the potential cost along a sequence generated by the fictitious play algorithm. In general, one can consider the exploitability of the game at each iteration and try to show its convergence to zero. Given a player and a belief on the others behaviors, the exploitability is the expected relative reward that the player can get by choosing a best response. This notion has recently received a growing attention [50, 51, 62, 101, 102, 103]. The convergence of the exploitability has been addressed in [103] in the context of continuous time learning and

discrete mean field games, and a convergence rate is provided.

**Link between the generalized conditional gradient and fictitious play** A key message of this article is that, in the context of continuous potential mean field games, the generalized conditional gradient algorithm can be interpreted as a fictitious play method. It relies on the following fact: at each step of the method, the problem to be solved (arising from a semi-linearization of the potential problem) coincides with the individual control problem of the agents, for a given belief of the coupling terms. The update formula  $\bar{x}_{k+1} = (1 - \delta_k)\bar{x}_k + \delta_k x_k$  corresponds to the learning step in the fictitious play algorithm, where the agents update their belief by averaging the past and the new distributions of states and controls.

This interpretation has already been highlighted in a very recent work [62], for a class of potential mean field games with some discrete structure. To the best of our knowledge, no other contribution in the literature has investigated the conditional gradient method for mean field games and has pointed out this interpretation. A minor difference between our framework and the one of [62] is the linearity of the running cost of the agents, so that they can apply the classical conditional gradient algorithm (and do not need to rely on semi-linearizations of the potential cost). In our PDE setting, we must employ the standard change of variable “à la Benamou-Brenier” and the perspective function of the running cost to get a convex potential problem. It turns out that in order to get an interpretation of the method as a learning method, the contribution of the perspective function (in the potential cost) must not be linearized, whence the use of the generalized conditional gradient algorithm.

**Contributions** Our contributions concern the well-posedness of the generalized conditional gradient algorithm and its convergence to the solution of the problem. The well-posedness is established with the help of suitable regularity estimates for the Hamilton-Jacobi-Bellman equation and the Fokker-Planck equation.

Similarly to [62], we use the standard convergence results of the conditional gradient method to prove that the potential cost converges at a rate  $O(1/k)$  and the exploitability at a rate  $O(1/\sqrt{k})$ , when  $\delta_k = 2/(k+2)$ .

In comparison with [62], the main novelty of our work (besides the different analytical framework) is the proof of convergence of all variables of the game: the coupling terms (price and congestion), the distribution of the agents, and their value function, at a rate  $O(1/\sqrt{k})$ . A key tool for the proof of convergence is a kind of quadratic growth property satisfied by the potential cost, which itself follows from the (assumed) strong convexity of the running cost of the agents.

Let us mention that we also provide convergence rates for the case  $\delta_k = 1/(k+1)$  which is more standard in the fictitious play algorithm:  $O(\ln(k)/k)$  for the potential cost,  $O(\sqrt{\ln(k)/k})$  for the exploitability and the different variables of the game.

**Plan of the paper** In Section 5.2 we provide our framework, the mean field game system we are interested in, and give our main assumptions. In Section 5.3 we study a stochastic individual control problem. We derive the Hamilton-Jacobi-Bellman equation associated with the value function of the control problem, and provide some regularity results. We link this problem with a partial differential equation (PDE) control problem of a Fokker-Planck equation and show existence of a (regular) optimal policy. In Section 5.4 we explicit the potential problem under study. We derive uniqueness results for the potential and the individual control problem. In Section 5.5 we recall the generalized

conditional gradient algorithm and apply it to our context. We show that the algorithm is well-defined. We define the exploitability and show the equality with the primal-dual gap. At the end of the section we exhibit the link with the fictitious play learning method. Finally, in Section 5.6, we provide our convergence results.

## 5.2 Data and main assumptions

### 5.2.1 Notations

We fix  $T > 0$  the duration of the game and  $d, k \in \mathbb{N}^*$  two dimensional coefficients.

**Sets** We set  $Q = \mathbb{T}^d \times [0, T]$ . Given a metric space  $X$ , we denote by  $X^*$  its dual. For any  $\alpha \in (0, 1)$ , we denote by  $\mathcal{C}^\alpha(Q)$  the set of Hölder continuous mappings on  $Q$  of exponent  $\alpha$  and by  $\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$  the set of continuous mappings  $u$  with Hölder continuous derivatives  $\partial_t u$ ,  $\nabla u$  and  $D_{xx}^2 u$  on  $Q$  of exponent  $\alpha$ . We also denote by  $\mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$  the set of all  $v \in \mathcal{C}^\alpha(Q; \mathbb{R}^d)$  with  $D_x v \in \mathcal{C}^\alpha(Q, \mathbb{R}^{d \times d})$ .

Sobolev spaces are denoted by  $W^{n,q}(Q)$ , the order of derivation  $n$  being possibly non-integral (following the definition in [87, section II.2]). We set

$$W^{2,1,q}(Q) = W^{1,q}(Q) \cap L^q(0, T; W^{2,q}(\mathbb{T}^d)), \quad W^{1,0,q}(Q) = L^q(0, T; W^{1,q}(\mathbb{T}^d)).$$

We define

$$\mathcal{D}_1(\mathbb{T}^d) = \left\{ m \in L^\infty(\mathbb{T}^d), m \geq 0, \int_{\mathbb{T}^d} m(x) dx = 1 \right\}.$$

We fix a real number  $p$  such that  $p > d + 2$ .

**Nemytskii notations** For any mappings  $g: Q \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $u: Q \rightarrow \mathbb{R}^d$ , we define  $\mathbf{g}[u]: Q \rightarrow \mathbb{R}^d$ ,

$$\mathbf{g}[u](x, t) = g(x, t, u(x, t))$$

called Nemytskii operator. This notation will mainly be used for the Hamiltonian  $H$ . Note that  $\mathbf{H}_p$  will denote the Nemytskii operator associated with the partial derivative of  $H$  with respect to  $p$  (a similar notation will be used for the other partial derivatives).

**Data of the problem** We fix an initial distribution and a terminal cost

$$m_0 \in \mathcal{D}_1(\mathbb{T}^d), \quad g: \mathbb{T}^d \rightarrow \mathbb{R},$$

and four maps: a running cost  $L$ , a congestion cost  $f$ , a vector of price  $\phi$  and an aggregation term  $a$ ,

$$\begin{aligned} L: & \quad Q \times \mathbb{R}^d \rightarrow \mathbb{R}, \\ f: & \quad Q \times \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}, \\ \phi: & \quad [0, T] \times \mathbb{R}^k \rightarrow \mathbb{R}^d, \\ a: & \quad Q \rightarrow \mathbb{R}^{k \times d}. \end{aligned}$$

We assume that  $L$  is strongly convex, more precisely, we assume that there exists a constant  $C_0 > 0$  such that for any  $v, v' \in \mathbb{R}^d$  and for any  $(x, t) \in Q$ , we have

$$\langle L_v(x, t, v) - L_v(x, t, v'), v - v' \rangle \geq \frac{1}{C_0} |v - v'|. \quad (\text{A1})$$

For any  $(x, t, p) \in Q \times \mathbb{R}^d$ , we define the Hamiltonian  $H$ ,

$$H(x, t, p) = \sup_{v \in \mathbb{R}^d} -\langle p, v \rangle - L(x, t, v).$$

The strong convexity assumption on  $L$  ensures that  $H$  takes finite values and is continuously differentiable (more regularity properties on  $H$  are collected in Appendix 5.7.1). We define the perspective function  $\tilde{L}: Q \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\tilde{L}(x, t, m, w) = \begin{cases} mL(x, t, \frac{w}{m}), & \text{if } m > 0, \\ 0, & \text{if } m = 0 \text{ and } w = 0, \\ +\infty, & \text{otherwise.} \end{cases} \quad (5.1)$$

Note that  $\tilde{L}$  is convex and lower semi-continuous with respect to  $(m, w)$ . We define  $A: L^1(Q; \mathbb{R}^d) \rightarrow L^1(0, T; \mathbb{R}^k)$  and  $A^*: L^1(0, T; \mathbb{R}^k) \rightarrow L^1(Q; \mathbb{R}^d)$  as follows,

$$A[w](t) = \int_{\mathbb{T}^d} a(x, t)w(x, t)dx, \quad A^*[P](x, t) = a^*(x, t)P(t),$$

for any  $(x, t) \in Q$ .

### 5.2.2 Coupled system and assumptions

The mean field game system under study is the following,

$$\left\{ \begin{array}{ll} \text{(i)} & \begin{cases} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^*P] = \gamma, & (x, t) \in Q, \\ u(x, T) = g(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(ii)} & v = -\mathbf{H}_p[\nabla u + A^*P], \quad (x, t) \in Q, \\ \text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, & (x, t) \in Q, \\ m(0, x) = m_0(x), & x \in \mathbb{T}^d, \end{cases} \\ \text{(iv)} & \gamma(x, t) = f(x, t, m(t)), \quad (x, t) \in Q, \\ \text{(v)} & P(t) = \phi[A[vm]](t), \quad t \in [0, T], \end{array} \right. \quad (\text{MFG})$$

where the unknown is  $(m, v, u, \gamma, P)$  with  $m(x, t) \in \mathbb{R}$ ,  $v(x, t) \in \mathbb{R}^d$ ,  $u(x, t) \in \mathbb{R}$ ,  $\gamma(x, t) \in \mathbb{R}$ , and  $P(t) \in \mathbb{R}^k$ , for any  $(x, t) \in Q$ . The equation (MFG,i) is a Hamilton-Jacobi-Bellman equation and describes the evolution of the value function as time goes backward. Equation (MFG,ii) defines the optimal control  $v$ , which is given by the gradient  $H_p$  of the Hamiltonian. Equation (MFG,iii) is a Fokker-Planck equation, describing the evolution of the state distribution of the agents. Equation (MFG,iv) defines the congestion  $\gamma$  and equation (MFG,v) the price  $P$ .

**Regularity assumptions** We assume that  $L_v$  is differentiable with respect to  $x$  and  $v$  and that  $a$  is differentiable with respect to  $x$ . All along the article, we make use of the following assumptions.

**Growth assumptions** There exists  $C_0 > 0$  such that for all  $(x, t) \in Q$ ,  $y \in \mathbb{T}^d$ ,  $v \in \mathbb{R}^d$ ,  $z \in \mathbb{R}^k$ , and  $m \in \mathcal{D}_1(\mathbb{T}^d)$ ,

$$L(x, t, v) \leq C_0|v|^2 + C_0, \quad (\text{A2})$$

$$|L(x, t, v) - L(y, t, v)| \leq C_0|x - y|(1 + |v|^2), \quad (\text{A3})$$

$$|\phi(t, z)| \leq C_0, \quad (\text{A4})$$

$$|f(x, t, m)| \leq C_0. \quad (\text{A5})$$

**Hölder and Lipschitz continuity assumptions** For all  $R > 0$ , there exists  $\alpha_0 \in (0, 1)$  such that

$$\left\{ \begin{array}{l} L \in \mathcal{C}^{\alpha_0}(B_R), \\ L_v \in \mathcal{C}^{\alpha_0}(B_R, \mathbb{R}^d), \\ L_{vx} \in \mathcal{C}^{\alpha_0}(B_R, \mathbb{R}^{d \times d}), \\ L_{vv} \in \mathcal{C}^{\alpha_0}(B_R, \mathbb{R}^{d \times d}), \end{array} \right. \quad \left\{ \begin{array}{l} \phi \in \mathcal{C}^{\alpha_0}(B'_R, \mathbb{R}^d), \\ a \in \mathcal{C}^{\alpha_0}(Q, \mathbb{R}^{k \times d}), \\ D_x a \in \mathcal{C}^{\alpha_0}(Q, \mathbb{R}^{k \times d \times d}), \end{array} \right. \quad (\text{A6})$$

where  $B_R = Q \times B(\mathbb{R}^d, R)$  and  $B'_R = [0, T] \times B(\mathbb{R}^k, R)$ . There exists  $\alpha_0 \in (0, 1)$  and  $C_0 > 0$  such that

$$|f(x_2, t_2, m_2) - f(x_1, t_1, m_1)| \leq C_0 \left( |x_2 - x_1| + |t_2 - t_1|^{\alpha_0} + \|m_2 - m_1\|_{L^2(\mathbb{T}^d)} \right), \quad (\text{A7})$$

for all  $(x_1, t_1)$  and  $(x_2, t_2) \in Q$  and for all  $m_1$  and  $m_2 \in \mathcal{D}_1(\mathbb{T}^d)$ . We further assume that  $\phi$  is Lipschitz continuous with respect to its second variable,

$$|\phi(t, z_2) - \phi(t, z_1)| \leq C_0|z_2 - z_1|, \quad (\text{A8})$$

for all  $(x, t) \in Q$ , for all  $z_1$  and  $z_2 \in \mathbb{R}^k$ .

**Remark 5.2.1.** Note that compared to the framework of [21] the Assumptions (A4) and (A7) are strengthened. Indeed, we require here more regularity: on  $f$  with respect to its third variable; on  $\phi$  with respect to its second variable.

**Boundary conditions and convention on constants** We assume that there exists  $\varepsilon_0 > 0$  such that  $m_0(x) \geq \varepsilon_0$  for any  $x \in \mathbb{T}^d$ . There exists  $\alpha_0 \in (0, 1)$  such that

$$m_0 \in \mathcal{C}^{2+\alpha_0}(\mathbb{T}^d), \quad g \in \mathcal{C}^{2+\alpha_0}(\mathbb{T}^d). \quad (\text{A9})$$

All along the article, we make use of two generic constants  $C > 0$  and  $\alpha \in (0, 1)$ . The value of  $C$  may increase from an inequality to the next one; the value of  $\alpha$  may decrease. The constants depend on the data of the problem introduced above.

### 5.2.3 Potentials

**Congestion** We assume that  $f$  is monotone, that is to say,

$$\int_{\mathbb{T}^d} (f(x, t, m_2) - f(x, t, m_1))(m_2(x) - m_1(x))dx \geq 0,$$

for any  $m_1$  and  $m_2 \in \mathcal{D}_1(\mathbb{T}^d)$  and for any  $t \in [0, T]$ . We assume that  $f$  has a primitive, that is, we assume the existence of a map  $F: [0, T] \times \mathcal{D}_1(\mathbb{T}^d)$  such that

$$F(t, m_2) - F(t, m_1) = \int_0^1 \int_{\mathbb{T}^d} f(x, t, sm_2 + (1-s)m_1)(m_2(x) - m_1(x))dx ds. \quad (5.2)$$



The monotonicity assumption implies that

$$F(t, m_2) \geq F(t, m_1) + \int_{\mathbb{T}^d} f(x, t, m_1)(m_2(x) - m_1(x))dx.$$

Since this inequality holds for any  $m_1 \in \mathcal{D}_1(\mathbb{T}^d)$ ,  $F$  is convex with respect to its second variable as the supremum of affine functions.

**Price** We assume that  $\phi$  has a convex potential  $\Phi$ , that is to say there exists a measurable mapping  $\Phi : [0, T] \times \mathbb{R}^k \rightarrow \mathbb{R}$ , convex with respect to its second variable and such that  $\phi(t, z) = \nabla_z \Phi(t, z)$  for any  $(t, z) \in [0, T] \times \mathbb{R}^k$ .

### 5.3 Estimates for the individual control problem

In this section we establish regularity results on the variables  $u$ ,  $v$ , and  $m$ , when obtained by solving the equations (MFG,i-iii), for fixed congestion and price. We investigate the stochastic optimal control problem associated with the HJB equation (MFG,i). In the section we fix  $\beta \in (0, 1)$  and we consider

$$\mathcal{U}^\beta = \mathcal{C}^{1,\beta}(Q) \times \mathcal{C}^\beta(0, T; \mathbb{R}^k). \quad (5.3)$$

We also fix a pair  $(\gamma, P) \in \mathcal{U}^\beta$  and a constant  $R > 0$  such that

$$\|\gamma\|_{L^\infty(Q)} + \|\nabla \gamma\|_{L^\infty(Q; \mathbb{R}^d)} + \|P\|_{L^\infty(0, T; \mathbb{R}^k)} \leq R. \quad (5.4)$$

#### 5.3.1 The individual problem as a stochastic optimal control problem

Let  $(B_s)_{s \in [0, T]}$  denote a Brownian motion and let  $Y$  be a random variable, independent of  $(B_s)_{s \in [0, T]}$ , with probability distribution  $m_0$ . Let  $\mathbb{F}$  denote the filtration generated by the Brownian motion  $(B_s)_{s \in [0, T]}$  and the initial random variable  $Y$ . We denote by  $L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)$  (resp.  $L_{\mathbb{F}}^{2,K}(t, T; \mathbb{R}^d)$ , for some constant  $K > 0$ ) the set of progressively measurable stochastic processes  $\nu$  on  $[t, T]$  with value in  $\mathbb{R}^d$  such that  $\mathbb{E} \left[ \int_t^T |\nu_s|^2 ds \right] < +\infty$  (resp.  $\mathbb{E} \left[ \int_t^T |\nu_s|^2 ds \right] \leq K$ ). For all  $\nu \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)$ , we denote by  $(X_s^\nu)_{s \in [0, T]}$  the solution to the stochastic differential equation

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_0 = Y.$$

We define the individual cost  $Z_{\gamma, P} : L_{\mathbb{F}}^2(0, T; \mathbb{R}^d) \rightarrow \mathbb{R}$ ,

$$Z_{\gamma, P}(\nu) = \mathbb{E} \left[ \int_0^T L(X_s^\nu, s, \nu_s) + \langle A^*[P](X_s^\nu, s), \nu_s \rangle + \gamma(X_s^\nu, s) ds + g(X_T^\nu) \right]. \quad (5.5)$$

We consider the following stochastic individual control problem

$$\inf_{\nu \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)} Z_{\gamma, P}(\nu). \quad (\mathbf{P}_{\gamma, P})$$

This problem will play an important role in the following, in particular in learning procedures: at each step, a representative player assumes the behavior of the others to be given and solves  $(\mathbf{P}_{\gamma, P})$ .

We define the mapping  $J_{\gamma,P}: Q \times L_{\mathbb{F}}^2(0, T; \mathbb{R}^d) \rightarrow \mathbb{R}$ ,

$$J_{\gamma,P}(x, t, \nu) = \mathbb{E} \left[ \int_t^T L(X_s, s, \nu_s) + \langle A^*[P](X_s, s), \nu_s \rangle + \gamma(X_s, s) ds + g(X_T) \right],$$

where  $(X_s)_{s \in [t, T]}$  is the solution to

$$dX_s = \nu_s ds + \sqrt{2} dB_s, \quad X_t = x.$$

We define by  $u[\gamma, P]: Q \rightarrow \mathbb{R}$  the value function associated with the individual control problem  $(P_{\gamma,P})$ ,

$$u[\gamma, P](x, t) = \inf_{\nu \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)} J_{\gamma,P}(x, t, \nu). \quad (5.6)$$

**Lemma 5.3.1.** *Let  $u = u[\gamma, P]$ . There exists a constant  $C > 0$ , only depending on  $R$ , such that*

$$u(x, t) = \inf_{\nu \in L_{\mathbb{F}}^{2,C}(t, T; \mathbb{R}^d)} J_{\gamma,P}(x, t, \nu)$$

for a.e.  $(x, t) \in Q$ , i.e. the optimization set in (5.6) can be restricted to  $L_{\mathbb{F}}^{2,C}(t, T; \mathbb{R}^d)$  (the set is defined in the beginning of section 5.3.1).

*Proof.* We first derive a lower bound of  $L$ . By assumption (A6),  $L(x, t, 0)$  and  $L_v(x, t, 0)$  are bounded. It follows then from the strong convexity assumption (A1) that there exists a constant  $C > 0$  such that

$$\frac{1}{C} |\nu|^2 - C \leq L(x, t, \nu), \quad \text{for all } (x, t, \nu) \in Q \times \mathbb{R}^d. \quad (5.7)$$

Then, for any  $(x, s) \in Q$  and for any  $\nu \in \mathbb{R}^d$ , we have the following estimates:

$$\begin{aligned} L(x, s, \nu) + \langle A^*[P](x, s), \nu \rangle &\geq \frac{1}{C} |\nu|^2 - \|a\|_{L^\infty(Q; \mathbb{R}^{k \times d})} |P(s)| |\nu| - C \\ &\geq \frac{1}{C} (|\nu|^2 - |P(s)|^2 - 1) \geq \frac{1}{C} (|\nu|^2 - 1). \end{aligned}$$

Let  $t \in [0, T]$ , let  $\varepsilon \in (0, 1)$  and let  $\tilde{\nu} \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)$  be an  $\varepsilon$ -optimal process. Using the bound on  $g$  given in Assumption (A9) and using inequality (5.4), we deduce from the above inequality that

$$\begin{aligned} \mathbb{E} \left[ \int_t^T |\tilde{\nu}_s|^2 ds \right] &\leq C \left( \inf_{\nu \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)} J_{\gamma,P}(x, t, \nu) + \varepsilon + 1 \right) \\ &\leq C (u[\gamma, P](x, t) + 2) \leq C, \end{aligned}$$

where the constant  $C$  does not depend on  $t$  and  $\varepsilon$ . Thus any  $\varepsilon$ -optimal process lies in  $L_{\mathbb{F}}^{2,C}(t, T; \mathbb{R}^d)$ , which concludes the proof.  $\square$

We now consider the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^*P] &= \gamma, & (x, t) \in Q, \\ u(x, T) &= g(x), & x \in \mathbb{T}^d. \end{aligned} \quad (5.8)$$

By the classical dynamic programming theory, we know that  $u[\gamma, P]$  is the unique viscosity solution to (5.8).

**Lemma 5.3.2.** *There exists  $\alpha \in (0, 1)$ , depending on  $\gamma$  and  $P$ , such that  $u[\gamma, P] \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ . In addition there exists a constant  $C > 0$ , only depending on  $R$ , such that*

$$\|u[\gamma, P]\|_{W^{2,1,p}(Q)} + \|\nabla u[\gamma, P]\|_{W^{2,1,p}(Q)} \leq C.$$

*Proof.* The proof is given in Appendix 5.7.4.  $\square$

### 5.3.2 The individual problem as a PDE optimal control problem

We consider in this subsection an equivalent formulation of  $(\mathbf{P}_{\gamma,P})$  as an optimal control problem of the Fokker-Planck equation. To this purpose, we consider the mapping  $\mathbf{m} : W^{1,0,\infty}(Q) \rightarrow W^{2,1,p}(Q)$  which associates to any  $v \in W^{1,0,\infty}(Q)$  the solution to the Fokker-Planck equation

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (vm) &= 0, & (x, t) \in Q, \\ m(x, 0) &= m_0(x), & x \in \mathbb{T}^d. \end{aligned} \quad (5.9)$$

**Lemma 5.3.3.** *The mapping  $\mathbf{m}$  is well defined. Moreover, for any  $v \in W^{1,0,\infty}(Q)$ , we have  $\mathbf{m}[v](x, t) > 0$ , for any  $(x, t) \in Q$ .*

*Proof.* Direct consequence of Lemma 5.7.9.  $\square$

We define  $\mathcal{B}^p = W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$  (recall that  $p > d + 2$  is fixed) and we define

$$\begin{aligned} \mathcal{R} &= \{(m, v) \in \mathcal{B}^p, \partial_t m - \Delta m + \nabla \cdot (vm) = 0, m(0) = m_0, (x, t) \in Q\}, \\ \tilde{\mathcal{R}} &= \{(m, w) \in \mathcal{B}^p, \partial_t m - \Delta m + \nabla \cdot w = 0, m(0) = m_0, m(x, t) > 0, (x, t) \in Q\}. \end{aligned}$$

**Lemma 5.3.4.** *The mapping  $\chi : \mathcal{R} \rightarrow \tilde{\mathcal{R}}$  given by  $\chi(m, v) = (m, mv)$  is well-posed and bijective. Its inverse is given by  $\chi^{-1}(m, w) = (m, w/m)$ .*

*Proof.* Let  $(m, v) \in \mathcal{R}$ . We have that  $m = \mathbf{m}[v] \in W^{2,1,p}(Q)$ , thus  $m \in L^\infty(Q)$  and  $\nabla m \in L^\infty(Q; \mathbb{R}^d)$ , by Lemma 5.7.6. It follows that  $w := \mathbf{m}[v]v \in W^{1,0,\infty}(Q)$ . Moreover,  $m > 0$ , by Lemma 5.3.3. Therefore  $(m, w) \in \tilde{\mathcal{R}}$ , that is,  $\chi$  is well defined. Similarly, for any  $(m, w) \in \tilde{\mathcal{R}}$ , we have that  $w/m \in W^{1,0,\infty}(Q)$  and  $\mathbf{m}[w/m] \in W^{2,1,p}(Q)$ . Obviously we have  $\chi \circ \chi^{-1} = id$  and  $\chi^{-1} \circ \chi = id$ , which concludes the proof.  $\square$

**Remark 5.3.5.** *Let  $(m, v) \in \mathcal{R}$  and let  $(m, w) = \chi(m, v) \in \tilde{\mathcal{R}}$ . Recalling the definition of the perspective function (5.1), we have*

$$\int_Q \mathbf{L}[v]m dx dt = \int_Q \tilde{\mathbf{L}}[m, w] dx dt.$$

*This fact, together with the existence of a bijection between  $\mathcal{R}$  and  $\tilde{\mathcal{R}}$ , will allow to prove the equivalence of the optimal control problems, introduced later, posed over  $\mathcal{R}$  and  $\tilde{\mathcal{R}}$ .*

We define the individual cost  $\mathcal{Z}_{\gamma,P} : \mathcal{R} \rightarrow \mathbb{R}$ ,

$$\mathcal{Z}_{\gamma,P}(m, v) = \int_Q (\mathbf{L}[v] + \gamma) m dx dt + \int_0^T \langle A[mv], P \rangle dt + \int_{\mathbb{T}^d} gm(T) dx.$$

We define the following individual control problem

$$\inf_{(m,v) \in \mathcal{R}} \mathcal{Z}_{\gamma,P}(m, v). \quad (\mathcal{P}_{\gamma,P})$$

Here the state equation of the agent is a Fokker-Planck equation with controlled drift  $v$ . We define the individual cost  $\tilde{\mathcal{Z}}_{\gamma,P} : \tilde{\mathcal{R}} \rightarrow \mathbb{R}$ ,

$$\tilde{\mathcal{Z}}_{\gamma,P}(m, w) = \int_Q (\tilde{\mathbf{L}}[m, w] + \gamma m) dx dt + \int_0^T \langle A[w], P \rangle dt + \int_{\mathbb{T}^d} gm(T) dx,$$

where  $\tilde{\mathbf{L}}$  is the perspective function of  $\mathbf{L}$  (see the definition (5.1)), and the following control problem

$$\inf_{(m,w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma,P}(m,w). \quad (\tilde{\mathcal{P}}_{\gamma,P})$$

Given  $v \in W^{1,0,\infty}(Q)$ , we denote  $(X_s^v)_{s \in [0,T]}$  the solution to the following stochastic differential equation

$$dX_s = v(X_s, s)ds + \sqrt{2}dB_s, \quad X_0 = Y. \quad (5.10)$$

We further consider the associated control  $\nu_s^v \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$  defined by  $\nu_s^v = v(s, X_s^v)$ .

**Lemma 5.3.6.** *For any  $v \in W^{1,0,\infty}(Q, \mathbb{R}^d)$ , we have*

$$Z_{\gamma,P}(\nu^v) = \mathcal{Z}_{\gamma,P}(\mathbf{m}[v], v) = \tilde{\mathcal{Z}}_{\gamma,P} \circ \chi(\mathbf{m}[v], v).$$

*Proof.* It is clear that  $\mathcal{Z}_{\gamma,P}(\mathbf{m}[v], v) = \tilde{\mathcal{Z}}_{\gamma,P} \circ \chi(\mathbf{m}[v], v)$ , see Remark 5.3.5. Since  $v \in W^{1,0,\infty}(Q, \mathbb{R}^d)$ , the process  $\nu^v$  lies in  $L_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$  and  $Z_{\gamma,P}(\nu^v) < +\infty$ . For any  $t \in [0, T]$ ,  $\mathbf{m}[v](\cdot, t)$  is the probability density of the distribution of  $X_t^v$ . In addition we have by definition that  $\nu_t^v = v(t, X_t^v)$ , which yields that  $Z_{\gamma,P}(\nu^v) = \mathcal{Z}_{\gamma,P}(\mathbf{m}[v], v)$ .  $\square$

**Lemma 5.3.7.** *Let  $u = \mathbf{u}[\gamma, P]$  and let  $v = -\mathbf{H}_p[\nabla u + A^*P]$ . Let  $m = \mathbf{m}[v]$  and let  $(m, w) = \chi(m, v)$ .*

1. *There exists  $\alpha \in (0, 1)$ , depending on  $\gamma$  and  $P$ , such that*

$$v \in \mathcal{C}^{1+\alpha,\alpha}(Q; \mathbb{R}^d), \quad m \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q), \quad w \in \mathcal{C}^{1+\alpha,\alpha}(Q; \mathbb{R}^d).$$

2. *There exists  $C > 0$ , depending only on  $R$ , such that*

$$\|v\|_{W^{1,0,\infty}(Q; \mathbb{R}^d)} \leq C, \quad \|m\|_{W^{2,1,p}(Q)} \leq C, \quad \|w\|_{W^{1,0,\infty}(Q; \mathbb{R}^d)} \leq C.$$

3. *The stochastic process  $(\nu_s^v)_{s \in [0,T]}$  is the solution to  $(\mathcal{P}_{\gamma,P})$ .*

4. *The pair  $(m, v)$  is a solution to  $(\mathcal{P}_{\gamma,P})$  and  $(m, w)$  is a solution to  $(\tilde{\mathcal{P}}_{\gamma,P})$ .*

*Proof.* *Point 1.* We know that  $H_p$  is Hölder continuous (Lemma 5.7.1),  $\nabla u$  is Hölder continuous (Lemma 5.3.2), and  $P$  is Hölder continuous by assumption. Thus  $v$  is Hölder continuous. Now we show that  $D_x v \in \mathcal{C}^\alpha(Q, \mathbb{R}^{d \times d})$ . The derivative of  $v$  is given by

$$D_x v = -\mathbf{H}_{px}[\nabla u + A^*P] - \mathbf{H}_{pp}[\nabla u + A^*P](D_{xx}^2 u + D_x A^*P). \quad (5.11)$$

Assumption (A6) yields  $D_x A^*P \in \mathcal{C}^\alpha(Q; \mathbb{R}^{d \times d})$ . In addition we have that  $\nabla u \in \mathcal{C}^\alpha(Q; \mathbb{R}^d)$  and  $D_{xx}^2 u \in \mathcal{C}^\alpha(Q; \mathbb{R}^{d \times d})$ . Finally the Hölder continuity of  $H_{pp}$  (see Lemma 5.7.1) yields  $v \in \mathcal{C}^{1+\alpha,\alpha}(Q)$ . It follows that  $m \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q; \mathbb{R}^d)$ , by Theorem 5.7.7 and  $w = mv \in \mathcal{C}^{1+\alpha,\alpha}(Q; \mathbb{R}^d)$ , as was to be proved.

*Point 2.* The constants  $C$  used for proving the second point only depend on  $C$ . By Lemma 5.7.1,  $H_p$ ,  $H_{pp}$ , and  $H_{px}$  are Hölder continuous. By (5.4) and Lemma 5.3.2, there exists  $C > 0$  only depending on  $R$  such that  $\|v\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$ .

We use again formula (5.11) for proving that  $D_x v$  is uniformly bounded. We know that  $a$  and  $D_x a$  are bounded (Assumption (A6)) and by Lemma 5.3.2,  $\nabla u$  and  $D_x x^2 u$  are bounded in  $L^\infty$  by some constant depending on  $R$ . We conclude that  $\|v\|_{W^{1,0,\infty}(Q; \mathbb{R}^d)} \leq$

$C$ , for some  $C$  depending only on  $R$ . Now we have that  $m$  is the solution to the Fokker-Planck equation

$$\begin{aligned} \partial_t m - \Delta m + m(\nabla \cdot v) + \nabla m \cdot v &= 0, & (x, t) \in Q, \\ m(0, x) &= m_0(x), & x \in \mathbb{T}^d. \end{aligned}$$

Since  $\|v\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C$ , we have that  $m$  is the solution of a parabolic PDE with bounded coefficients, which implies that  $\|m\|_{W^{2,1,p}(Q)} \leq C$ , by Theorem 5.7.3. By Lemma 5.7.6, we have  $\|m\|_{L^\infty(Q)} \leq C$  and  $\|\nabla m_k\|_{L^\infty(Q;\mathbb{R}^d)} \leq C$ . It follows that  $\|w\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C$  since  $w = mv$ .

*Point 3.* The statement holds by a classical verification argument.

*Point 4.* This is a direct consequence of Point 3 and Lemma 5.3.6. Indeed, for any  $(m', v') \in \mathcal{R}$ , we have

$$\mathcal{Z}_{\gamma,P}(m', v') = Z_{\gamma,P}(v^{v'}) \geq Z_{\gamma,P}(v^v) = \mathcal{Z}_{\gamma,P}(m[v], v),$$

which proves the optimality of  $(m[v], v)$ . The optimality of  $\chi(m[v], v)$  follows then from Remark 5.3.5.  $\square$

## 5.4 Properties of the solution to the mean field game system

We first recall the main result of [21] concerning the existence and uniqueness of a solution  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$  to (MFG). Then we establish a quadratic growth property (inequality (5.13)) which is at the heart of our convergence analysis in Section 5.6. It allows to show that  $(\bar{m}, \bar{v})$  is the unique solution to an optimization problem (P) and that  $(\bar{m}, \bar{m}\bar{v})$  is the unique solution to an equivalent convex potential problem ( $\tilde{P}$ ). With an analogous reasoning, we prove the uniqueness of the solutions to problems  $(\mathcal{P}_{\gamma,P})$  and  $(\tilde{\mathcal{P}}_{\gamma,P})$ .

**Theorem 5.4.1.** *There exists  $\alpha \in (0, 1)$  such that (MFG) has a unique classical solution  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$ , with*

$$\begin{cases} \bar{m} \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q), \\ \bar{v} \in \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d) \\ \bar{u} \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q), \\ \bar{\gamma} \in \mathcal{C}^\alpha(Q), \\ \bar{P} \in \mathcal{C}^\alpha(0, T; \mathbb{R}^k). \end{cases} \quad (5.12)$$

*Proof.* Direct application of [21, Theorem 1, Proposition 2].  $\square$

We define the following primal problem

$$\inf_{(m,v) \in \mathcal{R}} \mathcal{J}(m, v) := \int_Q L[v]m dx dt + \int_0^T (\mathbf{F}[m] + \Phi[A[mv]]) dt + \int_{\mathbb{T}^d} gm(T) dx. \quad (\text{P})$$

**Lemma 5.4.2.** *Let  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$  be the solution to (MFG). Then there exists a constant  $C > 0$  such that for any  $(m, v) \in \mathcal{R}$  we have the following estimate:*

$$\mathcal{J}(m, v) - \mathcal{J}(\bar{m}, \bar{v}) \geq \frac{1}{C} \int_Q |v - \bar{v}|^2 m dx dt. \quad (5.13)$$

*Proof.* By [21, Proposition 2], we have that  $(\bar{m}, \bar{v})$  is solution to Problem (P). By (MFG,ii) we have that  $\bar{v} = -\mathbf{H}_p[\nabla\bar{u} + A^*\bar{P}]$ . Then by Lemma 5.7.2,

$$\begin{aligned} \mathbf{L}[v](x, t)m(x, t) - \mathbf{L}[\bar{v}](x, t)\bar{m}(x, t) &\geq -\mathbf{H}[\nabla\bar{u} + A^*\bar{P}](x, t)(m(x, t) - \bar{m}(x, t)) \\ &\quad - \langle (\nabla\bar{u} + A^*\bar{P})(x, t), w(x, t) - \bar{w}(x, t) \rangle + \frac{1}{C}|v(x, t) - \bar{v}(x, t)|^2m(x, t), \end{aligned} \quad (5.14)$$

for all  $(x, t) \in Q$ ,  $v \in \mathcal{C}^{1+\alpha}(Q; \mathbb{R}^d)$  where  $(w, \bar{w}) = (mv, \bar{m}\bar{v})$ . By (MFG,i),

$$\int_Q -\mathbf{H}[\nabla\bar{u} + A^*\bar{P}](m - \bar{m})dxdt = \int_Q (-\partial_t\bar{u} - \Delta\bar{u} - \bar{\gamma})(m - \bar{m})dxdt. \quad (5.15)$$

By (MFG,iv) we have that  $\bar{\gamma}(x, t) = f(x, t, \bar{m}(t))$  thus by convexity of  $\mathbf{F}$ ,

$$\int_0^T (\mathbf{F}[m] - \mathbf{F}[\bar{m}]) dt \geq \int_Q \bar{\gamma}(m - \bar{m})dxdt. \quad (5.16)$$

By (MFG,v) we have that  $\bar{P} = \phi[A\bar{w}]$  thus by convexity of  $\Phi$ ,

$$\int_0^T (\Phi[A[w]] - \Phi[A[\bar{w}]]) dt \geq \int_0^T \langle \bar{P}, A[w - \bar{w}] \rangle dt = \int_Q \langle A^*\bar{P}, w - \bar{w} \rangle dxdt. \quad (5.17)$$

Combining (5.14), (5.15), (5.16), and (5.17) and integrating by parts we obtain that

$$\begin{aligned} \mathcal{J}(m, v) - \mathcal{J}(\bar{m}, \bar{v}) &\geq \int_Q ((\partial_t\bar{u} - \Delta\bar{u})(m - \bar{m}) - \nabla\bar{u}(w - \bar{w})) dxdt \\ &\quad + \int_{\mathbb{T}^d} (m(T) - \bar{m}(T))gdx + \frac{1}{C} \int_Q |v - \bar{v}|^2m dxdt \\ &\geq \int_Q \bar{u} (\partial_t(m - \bar{m}) - \Delta(m - \bar{m}) + \nabla \cdot (w - \bar{w})) dxdt \\ &\quad + \int_{\mathbb{T}^d} \bar{u}(0)(m(0) - m_0)dx + \frac{1}{C} \int_Q |v - \bar{v}|^2m dxdt. \end{aligned}$$

Then (5.13) holds since  $(m, w)$  and  $(\bar{m}, \bar{w})$  lie in  $\tilde{\mathcal{R}}$ .  $\square$

We next consider the problem

$$\inf_{(m, w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w) := \int_Q \tilde{\mathbf{L}}[m, w]dxdt + \int_0^T (\mathbf{F}[m] + \Phi[Aw]) dt + \int_{\mathbb{T}^d} gm(T)dx. \quad (\tilde{\mathbf{P}})$$

**Corollary 5.4.3.** *Let  $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$  be the unique solution to (MFG). Then  $(\bar{m}, \bar{v})$  is the unique solution to Problem (P) and  $(\bar{m}, \bar{w}) := \chi(\bar{m}, \bar{v})$  is the unique solution to Problem  $(\tilde{\mathbf{P}})$ .*

*Proof.* Let  $(m, v), (m', v') \in \mathcal{R}$  be two solutions to Problem (P). Then by Lemma 5.4.2 we have  $\int_Q |v - v'|^2m dxdt = 0$  which yields  $v = v'$  since  $m$  is positive. Then  $m$  and  $m'$  are solution to the same Fokker-Planck equation and thus  $m = m'$ . Finally,  $(\bar{m}, \bar{w})$  is the unique solution to  $(\tilde{\mathbf{P}})$ , by Remark 5.3.5.  $\square$

**Lemma 5.4.4.** *Let  $\beta \in (0, 1)$ , let  $(\gamma, P) \in \mathcal{U}^\beta$ , let  $u = \mathbf{u}[\gamma, P]$  and let  $v = -\mathbf{H}_p[\nabla u + P]$ . Then  $(\mathbf{m}[v], v)$  is the unique solution to Problem  $(\mathcal{P}_{\gamma, P})$  and  $\chi(\mathbf{m}[v], v)$  is the unique solution to Problem  $(\tilde{\mathcal{P}}_{\gamma, P})$ .*

*Proof.* The optimality of  $(m[v], v)$  and  $\chi(m[v], v)$  has been established in Lemma 5.3.7. Following the proof of Lemma 5.4.2, one can easily show that

$$\mathcal{Z}_{\gamma, P}(m', v') - \mathcal{Z}_{\gamma, P}(m[v], v) \geq \frac{1}{C} \int_Q |v' - v|^2 m' dx dt,$$

for any  $(m', v') \in \mathcal{R}$ . Applying the same reasoning as in the proof of Corollary 5.4.3 and using Remark 5.3.5 allows to conclude the proof.  $\square$

## 5.5 Generalized conditional gradient

In this section we first present the generalized conditional gradient method in an abstract framework. Then we present a generalized conditional gradient method for our potential mean field game. We show that this procedure is linked with the fictitious play method, a learning procedure. The generalized conditional gradient point of view allows us to link two notions from different areas: the notion of exploitability from game theory and the notion of duality gap defined in (generalized) conditional gradient theory.

**Abstract framework** We present here the main ideas of the generalized conditional gradient method in a finite dimensional setting. Consider the optimization problem

$$\min_{x \in K} f(x) = f_1(x) + f_2(x), \quad (\text{P}_f)$$

where  $K$  is a convex and compact subset of  $\mathbb{R}^n$  of finite diameter  $D$ ,  $f_1$  is a (possibly non-smooth) convex function and  $f_2$  a continuous differentiable function with  $L$ -Lipschitz gradient. We consider the mapping  $h: K \times K \rightarrow \mathbb{R}$  defined by

$$h(x, y) = f_1(y) - f_1(x) + \langle \nabla f_2(x), y - x \rangle.$$

The mapping  $h$  is a kind of first-order approximation of  $f(y) - f(x)$ , where only  $f_2$  is linearized. Let  $(\delta_k)_{k \in \mathbb{N}} \in [0, 1]$  be a sequence of step sizes. The method generates iteratively two sequences  $(\bar{x}_k)_{k \in \mathbb{N}}$  and  $(x_k)_{k \in \mathbb{N}}$  in  $K$ . At iteration  $k$ ,  $\bar{x}_k$  is available and  $(x_k, \bar{x}_{k+1})$  is obtained as follows:

$$\begin{aligned} x_k &\in \arg \min_{y \in K} h(\bar{x}_k, y), \\ \bar{x}_{k+1} &= (1 - \delta_k) \bar{x}_k + \delta_k x_k. \end{aligned}$$

We also consider the mapping  $\sigma: K \rightarrow \mathbb{R}$  defined by

$$\sigma(x) = - \min_{y \in K} h(x, y) \geq 0.$$

We call  $\sigma(x)$  the primal-dual gap at  $x \in K$ . This terminology is motivated by the following. Consider the Lagrangian  $\mathcal{L}: K \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\mathcal{L}(x, y, \lambda) = f_1(x) + f_2(y) + \langle \lambda, x - y \rangle.$$

It is easy to verify that  $(\text{P}_f)$  can be formulated as follows:

$$\inf_{x \in K, y \in \mathbb{R}^d} \sup_{\lambda \in \mathbb{R}^d} \mathcal{L}(x, y, \lambda).$$

In particular, for  $x \in K$ , we have  $f(x) = \sup_{\lambda' \in \mathbb{R}^d} \mathcal{L}(x, x, \lambda')$ . The dual problem writes

$$\sup_{\lambda \in \mathbb{R}^d} \inf_{x \in K, y \in \mathbb{R}^d} \mathcal{L}(x, y, \lambda).$$

Given  $x \in K$ , a candidate for the dual problem is  $\lambda = \nabla f_2(x)$ . The dual cost is then

$$\begin{aligned} \inf_{x' \in K, y' \in \mathbb{R}^d} \mathcal{L}(x', y', \lambda) &= \inf_{x' \in K} f_1(x') + \langle \lambda, x' \rangle + \inf_{y' \in \mathbb{R}^d} f_2(y') - \langle \nabla f_2(x), y' \rangle \\ &= f(x) + \inf_{x' \in K} h(x, x') = f(x) - \sigma(x). \end{aligned}$$

Thus  $\sigma(x)$  is nothing but the difference between the primal cost at  $x$ , and the dual cost at  $\nabla f_2(x)$ . We will later see that it coincides with the notion of exploitability in the context of mean field games.

Under the previous assumptions, one can show that (see [107, Lemma 2.4])

$$0 \leq f(\bar{x}_k) - f(\bar{x}) \leq \sigma(\bar{x}_k), \quad (5.18)$$

where  $\bar{x}$  is a solution to problem  $(P_f)$ . In words, any point  $x \in K$  is  $\sigma(x)$ -optimal.

**Application to potential mean field games** Our framework is infinite dimensional, we aim at minimizing the potential  $\tilde{\mathcal{J}}(m, w)$  under the constraint  $(m, w) \in \tilde{\mathcal{R}}$ . Following the ideas presented in the previous paragraph, we define a mapping  $h : \tilde{\mathcal{R}} \times \tilde{\mathcal{R}} \rightarrow \mathbb{R}$ ,

$$\begin{aligned} h((m, w), (m', w')) &= \tilde{\mathcal{Z}}_{\gamma, P}(m', w') - \tilde{\mathcal{Z}}_{\gamma, P}(m, w) \\ &= \int_Q \left( \tilde{\mathbf{L}}[m', w'] - \tilde{\mathbf{L}}[m, w] \right) dx dt + \int_{\mathbb{T}^d} g(m' - m)(T) dx \\ &\quad + \int_Q \gamma(m' - m) dx dt + \int_0^T \langle A[w' - w], P \rangle dt \end{aligned} \quad (5.19)$$

where  $\gamma(x, t) = f(x, t, m(t))$  and  $P(t) = \phi(t, Aw(t))$  for any  $(x, t) \in Q$ . By analogy with the previous abstract framework, we can interpret  $h((m, w), (m', w'))$  as a partial linearization of  $\tilde{\mathcal{J}}(m', w') - \tilde{\mathcal{J}}(m, w)$ : we have a non-linearized part composed of the perspective function  $\tilde{\mathbf{L}}$  (analogous to the term  $f_1$ ) and a linearized part composed of all the other terms (analogous to the term  $f_2$ ): the congestion  $\gamma$ , the price  $P$  and the terminal cost  $g$ . Two reasons motivates this choice of linearization:

1. In general the perspective function  $\tilde{\mathbf{L}}$  is not differentiable.
2. This particular choice of linearization allows to link the generalized conditional gradient method with the fictitious play algorithm, as explained in the end of this section.

We define the following generalized conditional gradient algorithm for potential mean field games as follows:

---

**Algorithm 5** Generalized conditional gradient

---

Choose  $(\bar{m}_0, \bar{w}_0) \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$  with  $\bar{m}_0(x, t) > 0$  for any  $(x, t) \in Q$  and choose a sequence  $(\delta_k)_{k \in \mathbb{N}} \in [0, 1]$ .

**for**  $0 \leq k < N$  **do**

Find  $(m_k, w_k) = \arg \min_{(m, w) \in \tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w))$

Actualise  $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1 - \delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$

**end for**

**return**  $(\bar{m}_N, \bar{w}_N)$ .

---



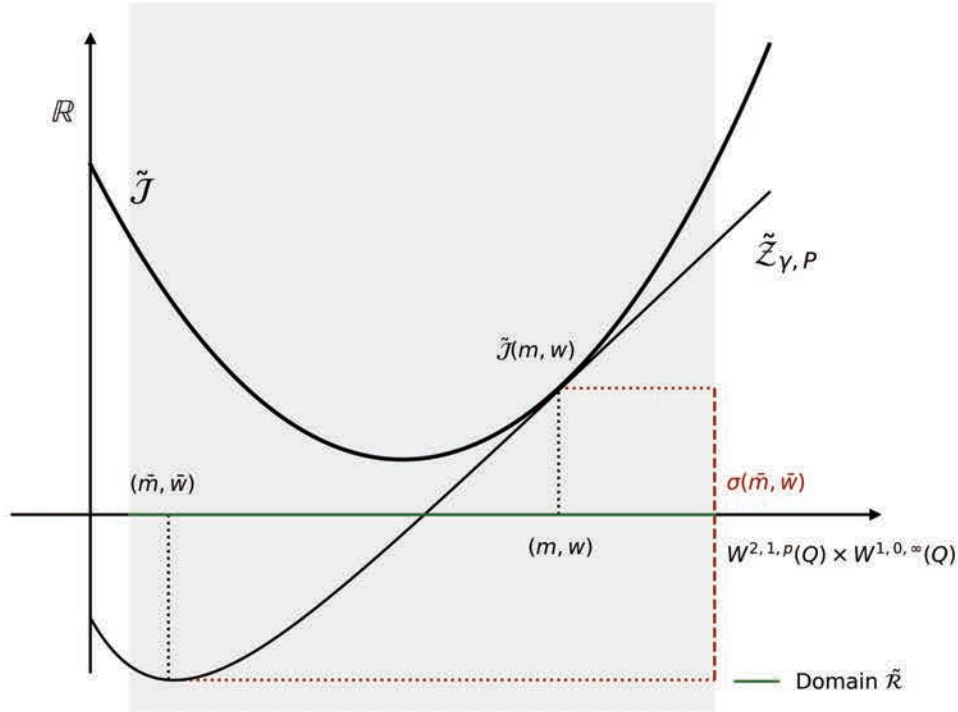


Figure 5.1: Illustration of the potential cost  $\tilde{J}$ , the individual cost  $\tilde{Z}_{\gamma,P}$  and the exploitability  $\sigma$ .

We first justify the well-posedness of the algorithm (in particular, we need to justify the existence and uniqueness of  $(m_k, w_k)$ ). To this goal, we introduce the following sequences

$$\begin{aligned} P_k(t) &= \phi(t, A\bar{w}_k(t)), & \gamma_k(x, t) &= f(x, t, \bar{m}_k(t)), \\ u_k(x, t) &= u[\gamma_k, P_k](x, t), & v_k(x, t) &= -\mathbf{H}_p[\nabla u_k + A^*P_k](x, t), \end{aligned}$$

for any  $(x, t) \in Q$ . For future reference, we define

$$\bar{v}_k = \bar{w}_k / \bar{m}_k.$$

In the next lemma, we provide an explicit formula to the minimization step, directly derived from Lemma 5.3.7.

**Lemma 5.5.1.** *For all  $k \in \mathbb{N}$ , we have  $(m_k, w_k) = \chi(\mathbf{m}[v_k], v_k)$ . Moreover, there exists  $\alpha_k \in (0, 1)$  such that*

$$\left\{ \begin{array}{l} m_k, \bar{m}_k \in \mathcal{C}^{2+\alpha_k, 1+\alpha_k/2}(Q), \\ v_k, w_k, \bar{w}_k \in \mathcal{C}^{1+\alpha_k, \alpha_k}(Q; \mathbb{R}^d), \\ u_k \in \mathcal{C}^{2+\alpha_k, 1+\alpha_k/2}(Q), \\ \gamma_k \in \mathcal{C}^{\alpha_k}(Q), \\ P_k \in \mathcal{C}^{\alpha_k}(0, T; \mathbb{R}^k). \end{array} \right. \quad (5.20)$$

*Proof.* We prove the result by induction. Let  $k \in \mathbb{N}$ . Assume that there exists  $\alpha \in (0, 1)$  such that  $\bar{m}_k \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ ,  $\bar{w}_k \in \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$ .

*Step 1:*  $P_k \in \mathcal{C}^\alpha(0, T; \mathbb{R}^k)$  and  $\gamma_k \in \mathcal{C}^\alpha(Q)$ . By assumptions (A6) and (A8),

$$\begin{aligned} |P_k(t_2) - P_k(t_1)| &= |\phi[A\bar{w}_k](t_2) - \phi[A\bar{w}_k](t_1)| \\ &\leq C (|t_2 - t_1|^\alpha + |A\bar{w}_k(t_2) - A\bar{w}_k(t_1)|) \\ &\leq C \left( |t_2 - t_1|^\alpha + \|a\|_{L^\infty(Q; \mathbb{R}^k \times d)} \int_{\mathbb{T}^d} |\bar{w}_k(t_2) - \bar{w}_k(t_1)| dx \right), \end{aligned}$$

for all  $t_1, t_2 \in [0, T]$ . It follows that  $P_k$  is Hölder continuous, since by induction assumption,  $\bar{w}_k \in \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$ . The announced regularity on  $\gamma_k$  is a direct consequence of the induction assumption ( $\bar{m}_k \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ ) and Assumption (A7).

*Step 2:*  $u_k \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ . The regularity of  $P_k$  and  $\gamma_k$  obtained in the previous steps allows us to apply 5.3.2, which yields the announced regularity on  $u_k$ .

*Step 3:*  $(m_k, w_k) = \chi(\mathbf{m}[v_k], v_k)$ . By Lemma 5.3.7 and Lemma 5.4.4,  $\chi(\mathbf{m}[v_k], v_k)$  is the unique minimizer of  $\mathcal{Z}_{\gamma_k, P_k}$ , thus the unique minimizer of  $h((\bar{m}_k, \bar{w}_k), \cdot)$  on  $\tilde{\mathcal{R}}$ .

*Step 4:*  $v_k \in \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$ ,  $m_k \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ , and  $w_k \in \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$ . Direct consequence of the previous steps and Point 1 of Lemma 5.3.7.

*Conclusion.* By Step 4 and by the induction assumption, we have that  $(\bar{m}_{k+1}, \bar{w}_{k+1}) \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \mathcal{C}^{1+\alpha, \alpha}(Q; \mathbb{R}^d)$ . Thus the induction assumption holds at  $k+1$ , which concludes the proof.  $\square$

**Link with the fictitious play** Let us consider the primal-dual gap

$$\sigma_k = - \min_{(m, w) \in \tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w)). \quad (5.21)$$

As mentioned earlier,  $\sigma_k$  is a primal gap certificate; it provides us with an upper bound of  $\tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \tilde{\mathcal{J}}(\bar{m}, \bar{w})$  (this will be proved in Lemma 5.6.2). In the current mean field game context, it coincides with the notion of exploitability: it is the largest decrease in cost that a representative agent can reach by playing its best response, assuming that all other agents use the feedback  $\bar{v}_k := \bar{w}_k / \bar{m}_k$ . Indeed, we have

$$\begin{aligned} \sigma_k &= \tilde{\mathcal{Z}}_{\gamma_k, P_k}(\bar{m}_k, \bar{w}_k) - \inf_{(m, w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma_k, P_k}(m, w) \\ &= \mathcal{Z}_{\gamma_k, P_k}(\nu^{\bar{v}_k}) - \inf_{\nu \in L^2_{\mathbb{F}}(t, T; \mathbb{R}^d)} \mathcal{Z}_{\gamma_k, P_k}(\nu), \end{aligned}$$

by Lemma 5.3.7 and Lemma 5.5.1.

We provide now an interpretation of the generalized gradient algorithm as a learning procedure called fictitious play. A definition and a study of the latter learning algorithm in the context of mean field games can be found in [35, 77]. Each iteration  $k$  of Algorithm 5 relies on the following steps:

For  $k \in \mathbb{N}$  let  $(\bar{m}_k, \bar{w}_k)$  be a given belief and  $\gamma_k$  and  $P_k$  the resulting beliefs on congestion and price. Then there are four main steps:

1. Given  $(\bar{m}_k, \bar{w}_k)$  compute the congestion terms  $P_k$  and  $\gamma_k$ . In words, the agents make a prediction of the congestion term and the price at equilibrium, based on the belief  $(\bar{m}_k, \bar{w}_k)$ .
2. Find the value function  $u_k$  solution to the Hamilton-Jacobi-Bellman equation parametrized by  $(\gamma_k, P_k)$ . Then compute the optimal control  $v_k$ , given the value function  $u_k$  and the price  $P_k$ . This step can be interpreted as follows: for a given belief on the distributions of the others  $(m_k, w_k)$ , a representative agent computes its best response  $v_k$ .

3. Find the solution  $m_k$  to the Fokker-Planck equation for the given drift  $v_k$  and compute the associated distribution of controls  $w_k$ .
4. The actualization step of  $(\bar{m}_{k+1}, \bar{w}_{k+1})$  can be interpreted as a learning step. The learning rule consists in averaging the past realizations of the distribution and flow at a rate determined by the sequence  $(\delta_k)_{k \in \mathbb{N}}$ .

## 5.6 Convergence Results

In this section, the generic constants  $C$  and  $\alpha$  depend on the data of the problem (introduced in Section 5.2.2) and depend on the pair  $(\bar{m}_0, \bar{w}_0)$  chosen to initialize Algorithm 5.

**Lemma 5.6.1.** *There exists  $C > 0$  such that for any  $k \in \mathbb{N}$ ,*

$$\begin{aligned}
\|\gamma_k\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} &\leq C & \|m_k\|_{W^{2,1,p}(Q)} &\leq C \\
\|P_k\|_{L^\infty(0,T;\mathbb{R}^k)} &\leq C & \|w_k\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} &\leq C \\
\|u_k\|_{W^{2,1,p}(Q)} &\leq C & \|\bar{m}_k\|_{W^{2,1,p}(Q)} &\leq C \\
\|\nabla u_k\|_{W^{2,1,p}(Q;\mathbb{R}^d)} &\leq C & \|\bar{w}_k\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} &\leq C \\
\|v_k\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} &\leq C & &
\end{aligned}$$

In addition, we have

$$m_k(x, t) \geq 1/C, \quad \bar{m}_k(x, t) \geq 1/C, \quad \|\bar{v}_k\|_{L^\infty(Q;\mathbb{R}^d)} \leq C,$$

for all  $(x, t) \in Q$ .

*Proof.* Let  $k \in \mathbb{N}$ . Assume that there exists  $C > 0$  such that the bounds hold for all  $i \in \{0, \dots, k-1\}$ .

*Step 1: Bounds of  $\gamma_k$  and  $P_k$ .* These bounds directly follow from assumptions (A4), (A5), and (A7). They imply the existence of  $C > 0$  such that

$$\|\gamma_k\|_{L^\infty(Q)} + \|\nabla \gamma_k\|_{L^\infty(Q;\mathbb{R}^d)} + \|P_k\|_{L^\infty(0,T;\mathbb{R}^k)} \leq C,$$

so that we can employ the technical Lemmas of Section 5.3 to prove the other announced bounds.

*Step 2: Bounds of  $u_k$  and  $\nabla u_k$ .* Direct consequence of Step 1 and Lemma 5.3.2.

*Step 3: Bounds of  $v_k$ ,  $m_k$  and  $w_k$ .* Direct consequence of the previous steps and Point 2 of Lemma 5.3.7.

*Step 4: Bounds of  $\bar{m}_k$  and  $\bar{w}_k$ .* This is a direct consequence of the fact that  $(\bar{m}_k, \bar{w}_k)$  can be expressed as a convex combination of  $(m_k, w_k)_{i=0,\dots,k-1}$  and  $(\bar{m}_0, \bar{w}_0)$ .

*Step 5:  $m_k(x, t), \bar{m}_k(x, t) \geq 1/C$  for any  $(x, t) \in Q$ .* Since  $m_k = \mathbf{m}[v_k]$  with  $\|v_k\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C$  and  $m_0(x) \geq \varepsilon_0$  for any  $x \in \mathbb{T}^d$ , therefore  $m_k(x, t) \geq 1/C$  by Lemma 5.7.9. Then  $\bar{m}_k(x, t) \geq 1/C$  as a convex combination of  $(m_k)_{i=0,\dots,k-1}$ .

*Step 6:  $\|\bar{v}_k\|_{L^\infty(Q;\mathbb{R}^d)} \leq C$ .* By Step 4 and Step 5,

$$\|\bar{v}_k\|_{L^\infty(Q;\mathbb{R}^d)} = \|\bar{w}_k/\bar{m}_k\|_{L^\infty(Q;\mathbb{R}^d)} \leq C.$$

*Conclusion.* Since  $(\bar{m}_0, \bar{w}_0) \in \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \mathcal{C}^{1+\alpha, \alpha}(Q;\mathbb{R}^d)$  with  $\bar{m}_0(x, t) > 0$  for any  $(x, t) \in Q$ , the conclusion follows by induction.  $\square$

Recall the definition of the exploitability  $\sigma_k$ , given in (5.21). We define the sequence of primal gaps  $(\epsilon_k)_{k \in \mathbb{N}}$  as follows

$$\epsilon_k = \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \tilde{\mathcal{J}}(\bar{m}, \bar{w}).$$

We recall that  $(\bar{m}, \bar{w}) = \arg \min_{(m, w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w)$ . The following Lemma is a certificate result, similar to inequality (5.18).

**Lemma 5.6.2.** *We have that  $\epsilon_k \leq \sigma_k$ .*

*Proof.* For any  $(m, w) \in \tilde{\mathcal{R}}$  we have that

$$h((\bar{m}_k, \bar{w}_k), (m, w)) = \tilde{\mathcal{J}}(m, w) - \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) + a + b,$$

where

$$\begin{aligned} a &= \int_0^T \mathbf{F}[\bar{m}_k] - \mathbf{F}[m] dt + \int_Q f(x, t, \bar{m}_k(t))(m(x, t) - \bar{m}_k(x, t)) dx dt \leq 0, \\ b &= \int_0^T \Phi[A\bar{w}_k] - \Phi[Aw] dt + \int_0^T \langle \phi(t, Aw_k(t)), A[w - \bar{w}_k](t) \rangle dt \leq 0, \end{aligned}$$

by convexity of  $\mathbf{F}$  and  $\Phi$ . Then we have that

$$\inf_{(m, w) \in \tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w)) \leq \inf_{(m, w) \in \tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w) - \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k), \quad (5.22)$$

and the conclusion follows.  $\square$

**Lemma 5.6.3.** *There exists  $C > 0$  such that for any  $\delta \in [0, 1]$ , it holds:*

$$\tilde{\mathcal{J}}(\bar{m}_k^\delta, \bar{w}_k^\delta) \leq \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \delta \sigma_k + \delta^2 C, \quad (5.23)$$

where  $(\bar{m}_k^\delta, \bar{w}_k^\delta) = \delta(m_k, w_k) + (1 - \delta)(\bar{m}_k, \bar{w}_k)$ .

*Proof.* The convexity of  $\tilde{\mathcal{L}}$  yields

$$\int_Q \tilde{\mathcal{L}}[\bar{m}_k^\delta, \bar{w}_k^\delta] dx dt \leq \int_Q \tilde{\mathcal{L}}[\bar{m}_k, \bar{w}_k] + \delta \left( \tilde{\mathcal{L}}[\bar{m}_k, \bar{w}_k] - \tilde{\mathcal{L}}[m_k, w_k] \right) dx dt. \quad (5.24)$$

Using that  $F$  is the primitive of  $f$  in the sense of (5.2), we have for all  $t \in [0, T]$ ,

$$\begin{aligned} \mathbf{F}[\bar{m}_k^\delta](t) &= \mathbf{F}[\bar{m}_k](t) \\ &+ \delta_k \int_0^1 \int_{\mathbb{T}^d} f(x, t, \bar{m}_k(t) + s\delta(m_k(t) - \bar{m}_k(t))) (m_k(x, t) - \bar{m}_k(x, t)) dx ds, \end{aligned} \quad (5.25)$$

For any  $(x, t) \in Q$ , the Lipschitz-continuity of  $f$  yields

$$\begin{aligned} f(x, t, \bar{m}_k(t) + s\delta(m_k(t) - \bar{m}_k(t))) &\leq f(x, t, \bar{m}_k(t)) + s\delta C \|m_k(t) - \bar{m}_k(t)\|_{L^2(\mathbb{T}^d)} \\ &\leq f(x, t, \bar{m}_k(t)) + s\delta C, \end{aligned}$$

since  $\bar{m}_k, m_k$  are uniformly bounded by Lemma 5.6.1. Plugging into (5.25) yields

$$\mathbf{F}[\bar{m}_k^\delta](t) = \mathbf{F}[\bar{m}_k](t) + \delta \int_{\mathbb{T}^d} f(x, t, \bar{m}_k(t))(m_k(x, t) - \bar{m}_k(x, t)) dx + \delta^2 C. \quad (5.26)$$

Now using that  $\Phi$  is the primitive of  $\phi$ , we have

$$\Phi[A\bar{w}_k^\delta](t) \leq \Phi[A\bar{w}_k](t) + \delta \langle \phi(t, A\bar{w}_k(t)), A[w_k - \bar{w}_k](t) \rangle + \delta^2 |A[w_k - \bar{w}_k](t)|^2$$

by Assumption (A8). Using that  $\bar{w}_k, w_k$  are uniformly bounded by Lemma 5.6.1 yields

$$|A[w_k - \bar{w}_k](t)| \leq \|a(t)\|_{L^\infty(\mathbb{T}^d)} \|w_k(t) - \bar{w}_k(t)\|_{L^\infty(\mathbb{T}^d)} \leq C.$$

Combining the two last inequalities yields

$$\Phi[A\bar{w}_k^\delta](t) \leq \Phi[A\bar{w}_k](t) + \delta \langle \phi(t, A\bar{w}_k(t)), A[w_k - \bar{w}_k](t) \rangle + \delta^2 C. \quad (5.27)$$

Then inequality (5.23) holds combining the Assumption (A9) on  $g$  and inequalities (5.24), (5.26), and (5.27) which concludes the proof.  $\square$

**Lemma 5.6.4.** *We have that*

$$\epsilon_{k+1} \leq (1 - \delta_k)\epsilon_k + \delta_k^2 C.$$

*Proof.* A direct application of Lemma 5.6.3 yields,

$$\tilde{\mathcal{J}}(\bar{m}_{k+1}, \bar{w}_{k+1}) \leq \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \delta_k \sigma_k + \delta_k^2 C.$$

Thus  $\epsilon_{k+1} \leq \epsilon_k - \delta_k \sigma_k + \delta_k^2 K$  and the conclusion follows by Lemma 5.6.2 since  $-\sigma_k \leq -\epsilon_k$ .  $\square$

**Lemma 5.6.5.** *Let  $L_0 := \max\{\epsilon_0/2, C\}$  and  $L_1 := \max\{2\epsilon_1, C\}/\ln(2)$ . We have that*

$$\begin{cases} \text{(i)} & \epsilon_k \leq \frac{4L_0}{k+2} & \text{if } \delta_k = \frac{2}{k+2}, & \text{for any } k \in \mathbb{N}, \\ \text{(ii)} & \epsilon_k \leq \frac{\ln(k+1)L_1}{k+1} & \text{if } \delta_k = \frac{1}{k+1}, & \text{for any } k \in \mathbb{N} \setminus \{0\}. \end{cases} \quad (5.28)$$

The above Lemma summarizes the rate of convergence of the sequence  $(\epsilon_k)_{k \in \mathbb{N}}$  for two learning rates. The first result (5.28,i) is classical in the context of conditional gradient algorithm (see [53, 59]). For the sake of completeness we recall how to derive this result in the following proof. The second result (5.28,ii) corresponds to the classical fictitious play learning rate.

*Proof. Step 1: (5.28,i) holds.* Let  $\delta_k = 2/(k+2)$  for any  $k \in \mathbb{N}$ . For  $k=0$ , it is clear that (5.28,i) holds. For  $k > 0$ , assume that  $\epsilon_k$  satisfies the inequality (5.28,i). By Lemma 5.6.4 we have that

$$\epsilon_{k+1} \leq \left(1 - \frac{2}{k+2}\right) \frac{4L_0}{k+2} + \frac{4C}{(k+2)^2} \leq \frac{4L_0(k+1)}{(k+2)^2} \leq \frac{4L_0}{(k+3)},$$

and by induction the step 1 is proved.

*Step 2: (5.28,ii) holds.* Let  $\delta_k = 1/(k+1)$  for any  $k \in \mathbb{N}$ . For  $k=1$ , it is clear that (5.28,ii) holds by Lemma 5.6.4. For  $k > 1$  assume that  $\epsilon_k$  satisfies the inequality (5.28,ii) then by Lemma 5.6.4 we have

$$\epsilon_{k+1} \leq \left(1 - \frac{1}{k+1}\right) \frac{\ln(k+1)L_1}{k+1} + \frac{C}{(k+1)^2}.$$

Then to prove (5.28,ii) it is enough to check

$$\frac{k \ln(k+1)L_1 + C}{(k+1)^2} \leq \frac{\ln(k+2)}{(k+2)}.$$

Multiplying both side by  $(k+1)^2(k+2)$ , the inequation (5.28,ii) holds if

$$\begin{aligned} (k+2)\frac{C}{L_1} &\leq (k+1)^2 \ln(k+2) - k(k+2) \ln(k+1) \\ &\leq k(k+2) \ln\left(1 + \frac{1}{k+1}\right) + \ln(k+2). \end{aligned} \quad (5.29)$$

The concavity of the logarithm yields  $\ln(1 + 1/(k+1)) \geq \ln(2)/k$ . Thus the inequality (5.29) holds whenever

$$\frac{C}{L_1} \leq \ln(2) + \frac{\ln(k+2)}{k+2},$$

which holds by definition of  $L_1$ . Then Step 2 is proved, which concludes the proof.  $\square$

**Lemma 5.6.6.** *There exists  $C > 0$  such that  $\sigma_k \leq C\epsilon_k^{1/2}$  for all  $k \in \mathbb{N}$ .*

*Proof.* For any  $\delta \in [0, 1]$ , Lemma 5.6.3 yields  $\tilde{\mathcal{J}}(\bar{m}, \bar{w}) \leq \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \delta\sigma_k + \delta^2C$ . It follows that

$$\sigma_k \leq \epsilon_k/\delta + \delta C, \quad \forall \delta \in (0, 1], \quad (5.30)$$

by Lemma 5.6.5. The optimal choice of  $\delta \in (0, 1]$  in the latter inequality is given by  $\delta = \min\{\sqrt{\epsilon_k/C}, 1\}$ . Since the sequence  $(\epsilon_k)_{k \in \mathbb{N}}$  is uniformly bounded from above, we can increase the constant  $C$ , so that one can choose  $\delta = \sqrt{\epsilon_k/C} \in (0, 1]$ . For this choice of  $\delta$ , inequality (5.30) yields the announced result.  $\square$

For any  $k \in \mathbb{N}$  we denote

$$\begin{aligned} \delta\bar{m}_k &= \bar{m}_k - \bar{m}, & \delta\bar{w}_k &= \bar{w}_k - \bar{w}, & \delta\bar{v}_k &= \bar{v}_k - \bar{v}, \\ \delta P_k &= P_k - \bar{P}, & \delta\gamma_k &= \gamma_k - \bar{\gamma}, & \delta u_k &= u_k - \bar{u}. \end{aligned}$$

**Theorem 5.6.7.** *There exists  $C > 0$  such that for all  $k \in \mathbb{N}$ ,*

$$\begin{aligned} &\|\delta\bar{v}_k\|_{L^2(Q; \mathbb{R}^d)} + \|\delta\bar{m}_k\|_{L^\infty(0, T; L^2(\mathbb{T}^d))} + \|\delta\bar{w}_k\|_{L^2(Q; \mathbb{R}^d)} \\ &\quad + \|\delta P_k\|_{L^2(0, T; \mathbb{R}^k)} + \|\delta\gamma_k\|_{L^\infty(Q)} + \|\delta u_k\|_{L^\infty(Q)} \leq C\epsilon_k^{1/2}. \end{aligned}$$

*Proof. Step 1:*  $\|\delta\bar{v}_k\bar{m}_k\|_{L^2(Q; \mathbb{R}^d)} \leq C\epsilon_k^{1/2}$ . By Lemma 5.4.2, we have

$$\frac{1}{C} \int_Q |\delta\bar{v}_k|^2 \bar{m}_k dx dt \leq \mathcal{J}(\bar{m}_k, \bar{v}_k) - \mathcal{J}(\bar{m}, \bar{v}) = \epsilon_k.$$

Combining the above inequality with  $\|\bar{m}_k\|_{L^\infty(Q)} \leq C$  yields the desired estimate.

*Step 2:*  $\|\delta\bar{v}_k\|_{L^2(Q; \mathbb{R}^d)} \leq C\epsilon_k^{1/2}$ . By Step 1 and Lemma 5.6.1,

$$\|\delta\bar{v}_k\|_{L^2(Q; \mathbb{R}^d)} \leq \|\delta\bar{v}_k\bar{m}_k\|_{L^2(Q; \mathbb{R}^d)} \|1/\bar{m}_k\|_{L^\infty(Q)} \leq C\epsilon_k^{1/2},$$

and Step 2 holds.

*Step 3:*  $\|\delta\bar{m}_k\|_{C(0, T; L^2(\mathbb{T}^d))} \leq C\epsilon_k^{1/2}$ . We have that  $\delta\bar{m}_k$  satisfies

$$\begin{aligned} \partial_t \delta\bar{m}_k - \Delta \delta\bar{m}_k + \nabla \cdot (\bar{v} \delta\bar{m}_k) &= -\nabla \cdot (\delta\bar{v}_k \bar{m}_k), & (x, t) &\in Q, \\ \delta m_k(0, x) &= 0, & x &\in \mathbb{T}^d. \end{aligned}$$

We define the space  $V = W^{2,1}(\mathbb{T}^d)$  and its dual  $V^*$ . Then  $\delta m_k$  is solution of a parabolic equation of the form

$$\begin{aligned}\partial_t m(t) + B(t)m(t) &= f(t), & (x, t) \in Q, \\ m(0, x) &= 0, & x \in \mathbb{T}^d,\end{aligned}$$

where  $B(t) \in L(V, V^*)$  and  $f(t) \in V^*$ . It is easy to verify that since  $\bar{v} \in W^{1,0,\infty}(Q; \mathbb{R}^d)$ , there exists a constant  $C$  such that  $\langle B(t)y, y' \rangle_V \leq C\|y\|_V\|y'\|_V$ , for a.e.  $t \in (0, T)$  and for all  $y$  and  $y'$  in  $V$ . For any  $y \in V$  we further have that

$$\begin{aligned}\langle B(t)y, y \rangle_V &= \int_{\mathbb{T}^d} (-\Delta y + \nabla \cdot \bar{v}(t)y + \langle \bar{v}(t), \nabla y \rangle) y dx \\ &\geq \int_{\mathbb{T}^d} |\nabla y|^2 + C|y|^2 - C|\nabla y||y| dx \geq \frac{1}{2}\|y\|_V^2 - \frac{C}{2}\|y\|_{L^2(\mathbb{T}^d)}^2,\end{aligned}$$

where we have used that  $-\int_{\mathbb{T}^d} |\nabla y||y| dx \geq -\frac{1}{2}\int_{\mathbb{T}^d} |\nabla y|^2/C + C|y|^2 dx$ . Then  $B(t)$  is semi-coercive, uniformly in time. Thus by [92, Chapter 3, Theorem 1.2] we have

$$\begin{aligned}\|\delta \bar{m}_k\|_{L^2(0,T;V)} + \|\partial_t \delta \bar{m}_k\|_{L^2(0,T;V^*)} &\leq C\|f\|_{L^2(0,T;V^*)} \\ &\leq C\|\nabla \cdot \delta \bar{v}_k \bar{m}_k\|_{L^2(0,T;V^*)} \\ &\leq C\|\delta \bar{v}_k \bar{m}_k\|_{L^2(Q;\mathbb{R}^d)} \leq C\epsilon_k^{1/2}.\end{aligned}$$

We conclude Step 3 with the continuous inclusion (see [92, Chapter 3, Theorem 1.1])

$$\{m \in L^2(0, T; V); \partial_t m \in L^2(0, T; V^*)\} \subseteq C(0, T; L^2(\mathbb{T}^d)).$$

*Step 4:*  $\|\delta \bar{w}_k\|_{L^2(Q;\mathbb{R}^d)} \leq C\epsilon_k^{1/2}$ . By definition of  $\delta \bar{w}_k$  we have

$$\|\delta \bar{w}_k\|_{L^2(Q;\mathbb{R}^d)} \leq \|\delta \bar{v}_k \bar{m}_k\|_{L^2(Q;\mathbb{R}^d)} + \|\bar{v} \delta \bar{m}_k\|_{L^2(Q;\mathbb{R}^d)} \leq C\epsilon_k^{1/2},$$

where the last inequality follows from Step 1 and Step 3.

*Step 5:*  $\|\delta P_k\|_{L^2(0,T;\mathbb{R}^k)} \leq C\epsilon_k^{1/2}$  and  $\|\delta \gamma_k\|_{L^\infty(Q)} \leq C\epsilon_k^{1/2}$ . Using that  $\phi$  is Lipschitz with respect to its second variable (see Assumption (A8)),

$$|\delta P_k(t)| = |\phi(t, A\bar{w}_k(t)) - \phi(t, A\bar{w}(t))| \leq C|A\delta \bar{w}_k(t)|$$

for almost every  $t \in [0, T]$ . Since

$$|A\delta \bar{w}_k(t)| = \left| \int_{\mathbb{T}^d} a(x, t) \delta \bar{w}_k(x, t) dx \right| \leq \|a(t)\|_{L^\infty(\mathbb{T}^d; \mathbb{R}^k \times d)} \|\delta \bar{w}_k(t)\|_{L^1(\mathbb{T}^d; \mathbb{R}^d)},$$

Since  $\|a\|_{L^\infty(Q;\mathbb{R}^k \times d)} \leq C$ , Step 4 yields the desired estimate

$$\|\delta P_k\|_{L^2(0,T;\mathbb{R}^k)} \leq C\|\delta \bar{w}_k\|_{L^2(Q;\mathbb{R}^d)} \leq C\epsilon_k^{1/2}.$$

Using that  $f$  is Lipschitz with respect to its third variable (see Assumption (A7)) yields

$$|\delta \gamma_k(x, t)| = |f(x, t, \bar{m}_k(t)) - f(x, t, \bar{m}(t))| \leq C\|\delta \bar{m}_k(t)\|_{L^2(\mathbb{T}^d)},$$

for any  $(x, t) \in Q$ . Taking the supremum over  $(x, t) \in Q$  both sides of the inequality yields that  $\|\delta \gamma_k\|_{C(Q)} \leq C\epsilon_k^{1/2}$  by Step 3, which concludes the step.

*Step 6:*  $\|\delta u_k\|_{C(Q)} \leq C\epsilon_k^{1/2}$ . Since  $\|\gamma_k\|_{L^\infty(Q)} \leq C$ ,  $\|P_k\|_{L^2(0,T;\mathbb{R}^k)} \leq C$ ,  $\|\bar{\gamma}\|_{L^\infty(Q)} \leq C$ , and  $\|\bar{P}\|_{L^2(0,T;\mathbb{R}^k)} \leq C$ , Lemma 5.3.1 yields

$$\delta u_k(x, t) = \inf_{\alpha \in L_{\mathbb{F}}^{2,C}(t,T;\mathbb{R}^d)} J_{\gamma_k, P_k}(x, t, \alpha) - \inf_{\alpha' \in L_{\mathbb{F}}^{2,C}(t,T;\mathbb{R}^d)} J_{\bar{\gamma}, \bar{P}}(x, t, \alpha'),$$

for any  $(x, t) \in Q$ . We denote  $(X_s^\alpha)_{s \in [t, T]}$  the solution to the stochastic differential equation  $dX_s = \alpha_s ds + \sqrt{2} dB_s$  with  $X_t^\alpha = x$ , for any  $\alpha \in L_{\mathbb{F}}^2(t, T; \mathbb{R}^d)$ . Then

$$\begin{aligned} |\delta u_k(x, t)| &\leq \sup_{\alpha \in L_{\mathbb{F}}^{2,C}(t,T;\mathbb{R}^d)} |J_{\gamma_k, P_k}(x, t, \alpha) - J_{\bar{\gamma}, \bar{P}}(x, t, \alpha)| \\ &\leq \sup_{\alpha \in L_{\mathbb{F}}^{2,C}(t,T;\mathbb{R}^d)} \mathbb{E} \left[ \int_t^T |\langle A^*[\delta P_k](X_s^\alpha, s), \alpha_s \rangle| + |\delta \gamma_k(X_s^\alpha, s)| ds \right]. \end{aligned}$$

For any  $(x, s) \in Q$  and  $\alpha \in \mathbb{R}^d$ , the Cauchy-Schwarz inequality yields

$$\begin{aligned} |\langle A^*[\delta P_k](x, s), \alpha \rangle| &\leq |\langle a(x, s) \delta P_k(s), \alpha \rangle| \\ &\leq \|a\|_{L^\infty(Q; \mathbb{R}^k \times d)} |\delta P_k(s)| |\alpha|. \end{aligned}$$

Since  $\|a\|_{L^\infty(Q; \mathbb{R}^k \times d)} \leq C$ , we finally have

$$|\delta u_k(x, t)| \leq C \left( \|\delta P_k\|_{L^2(0,T;\mathbb{R}^k)} + \|\delta \gamma_k\|_{L^\infty(Q)} \right).$$

Thus Step 6 holds by Step 5, which concludes the proof.  $\square$

Let us comment our last convergence results: Lemma 5.6.6 and Theorem 5.6.7. For the fictitious play learning rate  $\delta_k = 1/(k+1)$ , we have proved that the primal gap sequence  $(\epsilon_k)_{k \in \mathbb{N}}$  converges in  $O(\ln(k)/k)$  and the exploitability sequence  $(\sigma_k)_{k \in \mathbb{N}}$  and the sequence of variables  $(\bar{m}_k, \bar{w}_k, \bar{v}_k, P_k, \gamma_k, u_k)_{k \in \mathbb{N}}$  converge in  $O(\sqrt{\ln(k)/k})$ . We have obtained a sharper convergence result for the Frank-Wolfe learning rate  $\delta_k = 2/(k+2)$ . For this choice, we have shown that the primal gap sequence  $(\epsilon_k)_{k \in \mathbb{N}}$  converges in  $O(1/k)$  and the exploitability sequence  $(\sigma_k)_{k \in \mathbb{N}}$  and the sequence of variables  $(\bar{m}_k, \bar{w}_k, \bar{v}_k, P_k, \gamma_k, u_k)_{k \in \mathbb{N}}$  converge in  $O(\sqrt{1/k})$ . The convergence results for the variables of the problem  $(\bar{m}_k, \bar{w}_k, \bar{v}_k, P_k, \gamma_k, u_k)$  are new in the mean field game literature.

We conclude this section with a discussion on our results. The results concerning the convergence of the primal gap and the exploitability (Lemmas 5.6.5 and 5.6.6) are the same as those obtained in [62] for different mean field game models, with a discrete structure. These results are indeed general, since they only rely on the convexity structure of the potential problem and the regularity properties of the coupling terms. Therefore, they could certainly be adapted to other models, for example first order mean field games.

We also expect that similar convergence results, for the coupling terms, the value function, and the distribution, could be obtained in a different framework. A key step in the proof would be to establish a quadratic growth property (as the one obtained in Lemma 5.4.2), under a strong convexity assumption on the running cost  $L$ .

## 5.7 Appendix

### 5.7.1 Regularity of the Hamiltonian

Some properties of the Hamiltonian can be deduced from the convexity assumption (A1) and the Hölder continuity of  $L$  and its derivatives (Assumption (A6)). They are collected in the following lemmas. whose proofs can be found in [21].



**Lemma 5.7.1.** *The Hamiltonian  $H$  is differentiable with respect to  $p$  and  $H_p$  is differentiable with respect to  $x$  and  $p$ . Moreover, for all  $R > 0$ , there exists  $\alpha \in (0, 1)$  such that  $H \in C^\alpha(B_R)$ ,  $H_p \in C^\alpha(B_R, \mathbb{R}^d)$ ,  $H_{px} \in C^\alpha(B_R, \mathbb{R}^{d \times d})$ , and  $H_{pp} \in C^\alpha(B_R, \mathbb{R}^{d \times d})$ .*

*Proof.* See [21, Lemma 1].  $\square$

**Lemma 5.7.2.** *There exists a constant  $C > 0$  such that for all  $(x, t) \in Q$ , for all  $p \in \mathbb{R}^d$  and for all  $v \in \mathbb{R}^d$ ,*

$$H(x, t, p) + L(x, t, v) + \langle p, v \rangle \geq \frac{1}{C} |v + H_p(x, t, p)|^2. \quad (5.31)$$

*In addition for any  $m, \bar{m} \geq 0$  and  $\bar{v} = -H_p(x, t, p)$  we have that*

$$L(x, t, v)m - L(x, t, \bar{v})\bar{m} \geq -H(x, t, p)(m - \bar{m}) - \langle p, w - \bar{w} \rangle + \frac{1}{C} |v - \bar{v}|^2 m, \quad (5.32)$$

where  $(w, \bar{w}) := (mv, \bar{m}\bar{v})$ .

*Proof.* See [21, Proof of Proposition 2].  $\square$

## 5.7.2 A priori bounds for parabolic equations

In this appendix we provide estimates for the following parabolic equation:

$$\begin{aligned} \partial_t u - \sigma \Delta u + \langle b, \nabla u \rangle + cu &= h, & (x, t) \in Q, \\ u(x, 0) &= u_0(x), & x \in \mathbb{T}^d, \end{aligned} \quad (5.33)$$

for different assumptions on  $b$ ,  $c$ ,  $h$ , and  $u_0$ . The proofs of the following results can be found in the Appendix of [21]; they largely rely on [87]. We recall that  $p$  is a fixed parameter and  $p > d + 2$ .

**Theorem 5.7.3.** *For all  $R > 0$ , there exists  $C > 0$  such that for all  $u_0 \in W^{p, 2-2/p}(\mathbb{T}^d)$ , for all  $b \in L^p(Q; \mathbb{R}^d)$ , for all  $c \in L^p(Q)$ , for all  $h \in L^p(Q)$ , satisfying*

$$\|u_0\|_{W^{p, 2-2/p}(\mathbb{T}^d)} \leq R, \quad \|b\|_{L^p(Q; \mathbb{R}^d)} \leq R, \quad \|c\|_{L^p(Q)} \leq R, \quad \|h\|_{L^p(Q)} \leq R,$$

*equation (5.33) has a unique solution  $u$  in  $W^{2, 1, p}(Q)$ . Moreover,  $\|u\|_{W^{2, 1, p}(Q)} \leq C$ .*

**Theorem 5.7.4.** *For  $q \in (1, \infty)$ , the trace at time  $t = 0$  of elements of  $W^{2, 1, q}(Q)$  belongs to  $W^{q, 2-2/q}(\Omega)$ .*

**Theorem 5.7.5.** *There exists  $C > 0$  such that for all  $u_0 \in W^{2-2/p, p}(\mathbb{T}^d)$  and for all  $h \in L^p(Q)$ , the unique solution  $u$  to (5.33) (with  $b = 0$  and  $c = 0$ ) satisfies the following estimate:*

$$\|u\|_{W^{2, 1, p}(Q)} \leq C \left( \|u_0\|_{W^{2-2/p, p}(\mathbb{T}^d)} + \|h\|_{L^p(Q)} \right).$$

**Lemma 5.7.6.** *There exists  $\delta \in (0, 1)$  and  $C > 0$  such that for all  $u \in W^{2, 1, p}(Q)$ ,*

$$\|u\|_{C^\delta(Q)} + \|\nabla u\|_{C^\delta(Q; \mathbb{R}^d)} \leq C \|u\|_{W^{2, 1, p}(Q)}.$$

**Theorem 5.7.7.** *For all  $\beta \in (0, 1)$ , for all  $R > 0$ , there exist  $\alpha \in (0, 1)$  and  $C > 0$  such that for all  $u_0 \in C^{2+\beta}(\mathbb{T}^d)$ ,  $b \in C^{\beta, \beta/2}(Q; \mathbb{R}^d)$ ,  $c \in C^{\beta, \beta/2}(Q)$  and  $h \in C^{\beta, \beta/2}(Q)$  satisfying*

$$\|u_0\|_{C^{2+\beta}(\mathbb{T}^d)} \leq R, \quad \|b\|_{C^{\beta, \beta/2}(Q; \mathbb{R}^d)} \leq R, \quad \|c\|_{C^{\beta, \beta/2}(Q)} \leq R, \quad \text{and} \quad \|h\|_{C^{\beta, \beta/2}(Q)} \leq R,$$

*the solution to (5.33) lies in  $C^{2+\alpha, 1+\alpha/2}(Q)$  and satisfies  $\|u\|_{C^{2+\alpha, 1+\alpha/2}(Q)} \leq C$ .*

### 5.7.3 Maximum principle

In this appendix we establish a maximum principle for the Fokker-Planck equation. We study the parabolic equation (5.33) with  $h = 0$ ,

$$\begin{aligned} \partial_t m - \sigma \Delta m + \langle b, \nabla m \rangle + cm &= 0, & (x, t) \in Q, \\ m(x, 0) &= m_0(x), & x \in \mathbb{T}^d. \end{aligned} \quad (5.34)$$

We assume that  $m_0$  satisfies Assumption (A9) and define the mapping  $\bar{m}: L^\infty(Q; \mathbb{R}^d) \times L^\infty(Q) \rightarrow W^{2,1,p}(Q)$  which associates to any  $(b, c)$  the solution to (5.34). By Theorem 5.7.3 the mapping  $\bar{m}$  is well-defined.

**Lemma 5.7.8.** *The mapping  $\bar{m}$  is continuous.*

*Proof.* Consider the mapping  $\varphi: W^{2,1,p}(Q) \times L^\infty(Q; \mathbb{R}^d) \times L^\infty(Q) \rightarrow W^{p,2-2/p}(\mathbb{T}^d) \times L^p(Q)$  defined by

$$\varphi[m, b, c] = (m(0, \cdot) - m_0(\cdot), \partial_t m - \sigma \Delta m + \langle b, \nabla m \rangle + cm).$$

We define

$$\begin{aligned} \varphi_0[m] &= m(0, \cdot), & \varphi_2[m, b] &= \langle b, \nabla m \rangle, \\ \varphi_1[m] &= \partial_t m - \sigma \Delta m, & \varphi_3[m, c] &= cm, \end{aligned}$$

so that  $\varphi[m, b, c] = (\varphi_0[m] - m_0(\cdot), \varphi_1[m] + \varphi_2[m, b] + \varphi_3[m, c])$ . By Theorem 5.7.3 and Theorem 5.7.4, there exists a constant  $C > 0$  such that

$$\begin{aligned} \|\varphi_0[m]\|_{W^{p,2-2/p}(\mathbb{T}^d)} &\leq C \|m\|_{W^{2,1,p}(Q)}, & \|\varphi_2[m, b]\|_{L^p(Q)} &\leq \|b\|_{L^\infty(Q)} \|m\|_{W^{2,1,p}(Q)}, \\ \|\varphi_1[m]\|_{L^p(Q)} &\leq C \|m\|_{W^{2,1,p}(Q)}, & \|\varphi_3[m, c]\|_{L^p(Q)} &\leq \|c\|_{L^\infty(Q)} \|m\|_{W^{2,1,p}(Q)}. \end{aligned}$$

Thus  $\varphi_0$  and  $\varphi_1$  (resp.  $\varphi_2$  and  $\varphi_3$ ) are  $C^\infty$  as bounded linear (resp. bi-linear) applications. It follows that  $\varphi$  is  $C^\infty$ . Let  $(m, b, c) \in W^{2,1,p}(Q) \times L^p(Q; \mathbb{R}^d) \times L^p(Q)$  be such that  $\varphi[m, b, c] = 0$ . For any direction  $z \in W^{2,1,p}(Q)$ , we have

$$D_m \varphi[m, b, c]z = (z(0, \cdot), \partial_t z - \sigma \Delta z + \langle b, \nabla z \rangle + cz).$$

For any  $(h_0, h_1) \in W^{p,2-2/p}(\mathbb{T}^d) \times L^p(Q)$ , the equation

$$\begin{aligned} \partial_t z - \sigma \Delta z + \langle b, \nabla z \rangle + cz &= h_1, & (x, t) \in Q, \\ z(x, 0) &= h_0, & x \in \mathbb{T}^d, \end{aligned}$$

has a unique solution  $z \in W^{2,1,p}(Q)$ , by Theorem 5.7.3. Then  $D_m \varphi[m, b, c]$  is bijective and thus invertible. The conclusion follows by the implicit function theorem.  $\square$

**Lemma 5.7.9.** *Let  $v \in W^{1,0,\infty}(Q; \mathbb{R}^d)$  and let  $m = \bar{m}[v, \nabla \cdot v] \in W^{2,1,p}(Q)$  be the solution to (5.34) with  $(b, c) = (v, \nabla \cdot v)$ . Assume that  $m_0(x) \geq \varepsilon_0 > 0$  for any  $x \in \mathbb{T}^d$ . Then*

$$m(x, t) \geq \varepsilon_0 \exp(-T \|\nabla \cdot v\|_{L^\infty(Q)}), \quad \forall (x, t) \in Q. \quad (5.35)$$

*Proof.* We first prove the result when  $v \in \mathcal{C}^{1+\alpha, \alpha/2}(Q; \mathbb{R}^d)$ , for some  $\alpha \in (0, 1)$ . By Theorem 5.7.7,  $m \in C^{2,1}(Q)$ . Let  $\kappa > \|\nabla \cdot v\|_{L^\infty(Q)}$ . We define

$$y(x, t) = e^{-\kappa t} (m(x, t) - \varepsilon_0 e^{-\kappa t}),$$

for all  $(x, t) \in Q$ . By a direct computation we have

$$\begin{aligned} \partial_t y(x, t) &= -y(x, t)(\kappa - \nabla \cdot v(x, t)) + \Delta y(x, t) + \langle v(x, t), \nabla y(x, t) \rangle \\ &\quad + \varepsilon_0 e^{-2\kappa t} (\kappa + \nabla \cdot v(x, t)). \end{aligned} \quad (5.36)$$

Next we show that  $y(x, t) \geq 0$  for all  $(x, t) \in Q$ . Let  $(x_0, t_0) \in \arg \min_{(x,t) \in Q} y(x, t)$ . Let us assume, by a way of contradiction, that  $y(x_0, t_0) < 0$ . Since  $y(0, x) \geq 0$  for any  $x \in \mathbb{T}^d$ , we have that  $t_0 > 0$  and thus  $\partial_t y(x_0, t_0) \leq 0$ . Since  $x_0 \in \mathbb{T}^d$ , we have that  $\nabla y(x_0, t_0) = 0$ . Moreover, since  $m$  is twice differentiable with respect to its second variable, we have that  $\Delta y(x_0, t_0) \geq 0$ . Then it follows from (5.36) that

$$\partial_t y(x_0, t_0) \geq -y(x_0, t_0)(\kappa - \nabla \cdot v(x_0, t_0)) + \varepsilon_0 e^{-2\kappa t_0} (\kappa + \nabla \cdot v(x_0, t_0)).$$

The right-hand side is positive since  $\kappa > \|\nabla \cdot v\|_{L^\infty(Q)}$ . This contradicts the inequality  $\partial_t y(x_0, t_0) \leq 0$  and proves that  $y(x, t) \geq 0$ , for any  $(x, t) \in Q$ . It follows then from the definition of  $y$  that  $m(x, t) \geq \varepsilon_0 e^{-\kappa t}$ , for any  $(x, t) \in Q$ . Passing to the limit when  $\kappa \rightarrow \infty$  yields (5.35).

We now consider the general case when  $v \in W^{1,0,\infty}(Q; \mathbb{R}^d)$  and proceed by density. Let  $(\rho_k)_{k \in \mathbb{N}}$  be a sequence of regularizing kernels in  $C^\infty(Q)$ . We define  $v_k = \rho_k * v \in C^\infty(Q; \mathbb{R}^d)$ , where  $*$  is the convolution product. We next define  $m_k = \bar{m}[v_k, \nabla \cdot v_k]$  for any  $k \in \mathbb{N}$ . Applying (5.35) to  $m_k$ , we obtain that

$$m_k(x, t) \geq \varepsilon_0 \exp(-T \|\nabla \cdot v_k\|_{L^\infty(Q)}), \quad \forall (x, t) \in Q.$$

Since  $v_k$  (resp.  $\nabla v_k$ ) uniformly converges to  $v$  (resp.  $\nabla v$ ) and since  $\bar{m}$  is continuous for the uniform topology, we deduce that  $m_k$  converges to  $m$  in  $W^{2,1,p}(Q)$  and finally that  $m_k$  uniformly converges to  $m$ , by Lemma 5.7.6. This allows us to pass to the limit in the above inequality, which concludes the proof of the lemma.  $\square$

#### 5.7.4 Existence and uniqueness of a classical solution to the Hamilton-Jacobi-Bellman equation

In this appendix we prove Lemma 5.3.2, that is, we establish the existence of a solution to the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^* P] &= \gamma, & (x, t) \in Q, \\ u(x, T) &= g(x), & x \in \mathbb{T}^d, \end{aligned} \quad (5.37)$$

in  $C^{2,1}(Q)$ . By classical, we mean that (5.37) can be understood in a pointwise manner. We recall that  $g \in \mathcal{C}^{2+\alpha}(\mathbb{T}^d)$  and that  $(\gamma, P) \in \mathcal{U}^\beta$  (defined in (5.3)). Moreover, the constant  $R > 0$  is such that

$$\|\gamma\|_{L^\infty(Q)} + \|\nabla \gamma\|_{L^\infty(Q; \mathbb{R}^d)} + \|P\|_{L^\infty(0, T; \mathbb{R}^k)} \leq R. \quad (5.38)$$

The proof of the lemma relies on a fixed point approach. To this purpose, we introduce the mapping  $\mathcal{T}: W^{2,1,p}(Q) \times [0, 1] \rightarrow W^{2,1,p}(Q)$  which associates to any  $u \in W^{2,1,p}(Q)$  the classical solution  $\mathcal{T}[u, \tau]$  to the linear parabolic equation

$$\begin{aligned} -\partial_t \tilde{u} - \Delta \tilde{u} + \tau \mathbf{H}[\nabla u + A^* P] &= \tau \gamma & (x, t) \in Q, \\ \tilde{u}(x, T) &= \tau g(x) & x \in \mathbb{T}^d. \end{aligned}$$

For any  $(u, \tau) \in W^{2,1,p}(Q) \times [0, 1]$ , we have  $\tau(\gamma - \mathbf{H}[\nabla u + A^* P]) \in L^\infty(Q)$ , thus  $\mathcal{T}[u, \tau]$  lies in  $W^{2,1,p}(Q)$ , by Theorem 5.7.3, proving that  $\mathcal{T}$  is well-defined.

**Lemma 5.7.10.** *The mapping  $\mathcal{T}: W^{2,1,p}(Q) \times [0, 1] \rightarrow W^{2,1,p}(Q)$  is continuous and compact. In addition, for all  $K > 0$ , there exists  $\alpha \in (0, 1)$  and  $C > 0$  depending on  $K$ ,  $\gamma$ , and  $P$  such that  $\|u\|_{W^{2,1,p}(Q)} \leq K$  implies  $\|\mathcal{T}[u, \tau]\|_{\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)} \leq C$ .*

*Proof. Step 1: Continuity of  $\mathcal{T}$ .* Let  $(u_k, \tau_k) \in W^{2,1,p}(Q) \times [0, 1]$  be a sequence converging to  $(u, \tau) \in W^{2,1,p}(Q) \times [0, 1]$ . Then  $\nabla u_k \rightarrow \nabla u$  in  $L^\infty(Q; \mathbb{R}^d)$  by Lemma 5.7.6. Then  $\tau_k(\gamma - \mathbf{H}[\nabla u_k + A^*P]) \rightarrow \tau(\gamma - \mathbf{H}[\nabla u + A^*P])$  in  $L^\infty(Q; \mathbb{R}^d)$  by continuity of the Hamiltonian (see Lemma 5.7.1). Finally the continuity of  $\mathcal{T}$  follows by Theorem 5.7.5.

*Step 2: Compactness of  $\mathcal{T}$ .* Let  $K > 0$  and let  $(u, \tau) \in W^{2,1,p}(Q) \times [0, 1]$  be such that  $\|u\|_{W^{2,1,p}(Q)} + |\tau| < K$ . Combining Lemma 5.7.6 and Lemma 5.7.1 there exist  $\alpha \in (0, 1)$  and  $C > 0$  such that  $\|\tau(\gamma - \mathbf{H}[\nabla u + A^*P])\|_{\mathcal{C}^\alpha(Q)} < C$ . Then applying Theorem 5.7.7, there exist  $\alpha \in (0, 1)$  and  $C > 0$  such that  $\|\mathcal{T}[u, \tau]\|_{\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)} < C$ . By the Arzela-Ascoli Theorem the centered ball of  $\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$  of radius  $C > 0$  is a relatively compact subset of  $W^{2,1,p}(Q)$ . As a consequence  $\mathcal{T}[u, \tau]$  is a compact mapping and the conclusion follows.  $\square$

**Theorem 5.7.11.** (Leray-Schauder) *Let  $X$  be a Banach space and let  $T: X \times [0, 1] \rightarrow X$  be a continuous and compact mapping. Assume that  $T(x, 0) = 0$  for all  $x \in X$  and assume there exists  $C > 0$  such that  $\|x\|_X < C$  for all  $(x, \tau) \in X \times [0, 1]$  such that  $T(x, \tau) = x$ . Then, there exists  $x \in X$  such that  $T(x, 1) = x$ .*

*Proof.* See [63, Theorem 11.6].  $\square$

*Proof of Lemma 5.3.2. Step 1: Existence of a classical solution.* We have that  $\mathcal{T}[u, 0] = 0$  for all  $u \in W^{2,1,p}(Q)$ . Now let  $(u, \tau) \in W^{2,1,p}(Q) \times [0, 1]$  such that  $\mathcal{T}[u, \tau] = u$ . From Lemma 5.7.10, the mapping  $\mathcal{T}$  is continuous and compact, in addition  $u$  is a classical solution and thus the viscosity solution to the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\partial_t u - \Delta u + \tau \mathbf{H}[\nabla u + A^*P] &= \tau \gamma & (x, t) \in Q, \\ u(x, T) &= \tau g(x) & x \in \mathbb{T}^d, \end{aligned}$$

and can be interpreted as the value function associated to the following stochastic control problem

$$\inf_{\nu \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^d)} \tau \mathbb{E} \left[ \int_0^T L(X_s^\tau, s, \nu_s) + \langle A^*[P](X_s^\tau, s), \nu_s \rangle + \gamma(X_s^\tau, s) ds + g(X_T^\tau) \right],$$

where  $(X_s^\tau)_{s \in [t, T]}$  is the solution to  $dX_s = \tau \nu_s ds + \sqrt{2} dB_s$ ,  $X_0 = Y$ . Following [21, Proposition 1, Step 2], there exists a constant  $C > 0$ , depending only on  $R$ , such that  $\|u\|_{L^\infty(Q)} + \|\nabla u\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$ . Assumption (5.38) yields that  $\|\mathbf{H}[\nabla u + A^*P] - \gamma\|_{L^\infty(Q)} \leq C$ . Then  $u$  is the solution to a parabolic PDE with bounded coefficients and thus  $\|u\|_{W^{2,1,p}(Q)} \leq C$ , by Theorem 5.7.3. Again,  $C$  only depends on  $R$ . By the Leray-Schauder Theorem 5.7.11, there exists a classical solution to (5.37).

*Step 2: Uniqueness.* Let  $u_1$  and  $u_2$  be two classical solutions to (5.37). Then  $u_1$  and  $u_2$  are viscosity solutions to (5.37). Since the viscosity solution is unique,  $u_1 = u_2$ .

*Step 3:*  $\|u\|_{W^{2,1,p}(Q)} + \|\nabla u\|_{W^{2,1,p}(Q)} \leq C$ . We have already obtained a bound on  $\|u\|_{W^{2,1,p}(Q)}$  in Step 1. It remains to show that  $\|\nabla u\|_{W^{2,1,p}(Q)} \leq C$ . Let  $i \in \{1, \dots, d\}$ . We have that  $u^i := \partial_{x_i} u$  is the solution to the following equation

$$\begin{aligned} -\partial_t u^i - \Delta u^i + \partial_{x_i} \mathbf{H}[\nabla u + A^*P] + \mathbf{H}_p[\nabla u + A^*P] \cdot (\nabla u^i + \partial_{x_i} A^*P) &= \partial_{x_i} \gamma, \\ u^i(T) &= \partial_{x_i} g, \end{aligned}$$

for any  $(x, t) \in Q$ . By Lemma 5.7.1,  $\partial_{x_i} \mathbf{H}$  and  $\mathbf{H}_p$  are continuous, thus

$$\|\partial_{x_i} \mathbf{H}[\nabla u + A^*P]\|_{L^\infty(Q)} \leq C, \quad \|\mathbf{H}_p[\nabla u + A^*P]\|_{L^\infty(Q; \mathbb{R}^d)} \leq C,$$

since  $\|\nabla u\|_{C^\alpha(Q; \mathbb{R}^d)} \leq C$  and  $\|A^*P\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$ . By Assumption (A6),  $\partial_{x_i} a$  is continuous, therefore  $\|\partial_{x_i} A^*P\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$ . We further have  $\|\nabla \gamma_k\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$  and  $\|\partial_{x_i} g\|_{L^\infty(\mathbb{T}^d)} \leq C$ , by Assumption (A9). It follows that  $u_k^i$  is the solution of a parabolic PDE with bounded coefficients, thus by Theorem 5.7.3,  $\|u^i\|_{W^{2,1,p}(Q)} \leq C$  and the Step 3 is proved which concludes the proof.  $\square$



# Bibliography

- [1] Yves Achdou, Francisco J. Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Partial differential equation models in macroeconomics. *Phil. Trans. R. Soc. A*, 372(2028):19 pages, 2014. [41](#)
- [2] Yves Achdou, Francisco J. Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Partial differential equation models in macroeconomics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2028):20130397, 2014. [97](#)
- [3] Yves Achdou, Fabio Camilli, and Italo Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM Journal on Control and Optimization*, 50(1):77–109, 2012. [12](#), [23](#), [35](#)
- [4] Yves Achdou and Italo Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010. [12](#), [23](#)
- [5] Yves Achdou and Ziad Kobeissi. Mean field games of controls: Finite difference approximations, 2020. [11](#), [12](#), [23](#), [24](#), [68](#), [98](#)
- [6] Yves Achdou and Mathieu Laurière. Mean field games and applications: Numerical aspects. *arXiv preprint arXiv:2003.04444*, 2020. [11](#), [23](#), [30](#), [68](#), [98](#)
- [7] Clémence Alasseur, Imen Ben Tahar, and Anis Matoussi. An extended mean field game for storage in smart grids. *Journal of Optimization Theory and Applications*, 184(2):644–670, 2020. [97](#)
- [8] Clémence Alasseur, Imen Ben Taher, and Anis Matoussi. An extended mean field game for storage in smart grids. *Journal of Optimization Theory and Applications*, pages 1–27, 2020. [24](#), [42](#)
- [9] Andrea Angiuli, Jean-Pierre Fouque, and Mathieu Lauriere. Reinforcement learning for mean field games, with applications to economics. *arXiv preprint arXiv:2106.13755*, 2021. [12](#), [23](#)
- [10] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. [24](#), [41](#)
- [11] Heinz H. Bauschke and Patrick L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*, volume 408. Springer, 2011. [72](#)
- [12] Dario Bauso, Hamidou Tembine, and Tamer Basar. Opinion dynamics in social networks through mean-field games. *SIAM Journal on Control and Optimization*, 54(6):3225–3257, 2016. [97](#)

- [13] Dario Bauso, Hamidou Tembine, and Tamer Başar. Robust mean field games. *Dynamic games and applications*, 6(3):277–303, 2016. [25](#)
- [14] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000. [29](#), [68](#), [84](#)
- [15] Jean-David Benamou and Guillaume Carlier. Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, Oct 2015. [12](#), [23](#), [68](#), [98](#)
- [16] Jean-David Benamou, Guillaume Carlier, Simone Di Marino, and Luca Nenna. An entropy minimization approach to second-order variational mean-field games. *Mathematical Models and Methods in Applied Sciences*, 29(08):1553–1583, 2019. [12](#), [23](#), [68](#), [98](#)
- [17] Jean-David Benamou, Guillaume Carlier, and Filippo Santambrogio. Variational mean field games. In *Active Particles, Volume 1*, pages 141–171. Springer, 2017. [12](#), [23](#), [30](#), [68](#), [97](#), [98](#)
- [18] Charles Bertucci. A remark on Uzawa’s algorithm and an application to mean field games systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 54(3):1053–1071, 2020. [12](#), [23](#)
- [19] Charles Bertucci, Spyridon Vassilaras, Jean-Michel Lasry, Georgios S. Paschos, Mérouane Debbah, and Pierre-Louis Lions. Transmit strategies for massive machine-type communications based on mean field games. In *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–5. IEEE, 2018. [24](#), [42](#)
- [20] J. Frédéric Bonnans. *Convex and Stochastic Optimization*. Universitext Series. Springer-Verlag, 2019. [71](#)
- [21] J. Frédéric Bonnans, Saeed Hadikhanloo, and Laurent Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Applied Mathematics & Optimization*, 83:1431–1464, 2021. [8](#), [20](#), [30](#), [35](#), [36](#), [41](#), [68](#), [97](#), [102](#), [107](#), [108](#), [118](#), [119](#), [122](#)
- [22] J. Frédéric Bonnans, Pierre Lavigne, and Laurent Pfeiffer. Discrete potential mean field games. 2021. [12](#), [23](#), [24](#), [67](#), [98](#)
- [23] J. Frédéric Bonnans, Pierre Lavigne, and Laurent Pfeiffer. Discrete-time mean field games with risk-averse agents. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:44, 2021. [41](#)
- [24] J. Frédéric Bonnans, Pierre Lavigne, and Laurent Pfeiffer. Generalized conditional gradient and learning in potential mean field games, 2021. [12](#), [23](#), [24](#)
- [25] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013. [73](#)
- [26] Kristian Bredies, Dirk A. Lorenz, and Peter Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *Computational Optimization and Applications*, 42(2):173–193, 2009. [98](#)



- [27] Luis Briceño-Arias, Dante Kalise, Ziad Kobeissi, Mathieu Laurière, A. Mateos González, and Francisco J. Silva. On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings. *ESAIM: Proceedings and Surveys*, 65:330–348, 2019. [12](#), [23](#), [68](#), [98](#)
- [28] Luis M. Briceno-Arias, Dante Kalise, and Francisco J. Silva. Proximal methods for stationary mean field games with local couplings. *SIAM Journal on Control and Optimization*, 56(2):801–836, 2018. [12](#), [23](#), [68](#), [98](#)
- [29] George W. Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951. [98](#)
- [30] Simone Cacace, Fabio Camilli, Annalisa Cesaroni, and Claudio Marchi. An ergodic problem for mean field games: qualitative properties and numerical simulations. *arXiv preprint arXiv:1801.08828*, 2018. [12](#), [23](#)
- [31] Simone Cacace, Fabio Camilli, and Claudio Marchi. A numerical method for mean field games on networks. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51(1):63–88, 2017. [12](#), [23](#)
- [32] Pierre Cardaliaguet. Notes on mean field games. Technical report, Technical report, 2010. [27](#)
- [33] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games*. Princeton University Press, 2019. [29](#)
- [34] Pierre Cardaliaguet, P. Jameson Graber, Alessio Porretta, and Daniela Tonon. Second order mean field games with degenerate diffusion and local coupling. *Non-linear Differential Equations and Applications NoDEA*, 22(5):1287–1317, 2015. [9](#), [11](#), [20](#), [22](#), [30](#), [68](#), [97](#)
- [35] Pierre Cardaliaguet and Saeed Hadikhanloo. Learning in mean field games: the fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017. [12](#), [23](#), [38](#), [39](#), [68](#), [98](#), [112](#)
- [36] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018. [24](#)
- [37] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018. [41](#), [97](#)
- [38] Pierre Cardaliaguet, Alpár R. Mészáros, and Filippo Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016. [30](#), [68](#), [97](#)
- [39] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games with Applications II: Mean Field Games with Common Noise and Master Equations*, volume 84. Springer, 2018. [29](#)
- [40] René Carmona, François Delarue, and Daniel Lacker. Mean field games of timing and models for bank runs. *Applied Mathematics & Optimization*, 76(1):217–260, 2017. [24](#), [97](#)

- [41] René Carmona and Daniel Lacker. A probabilistic weak formulation of mean field games and applications. *The Annals of Applied Probability*, 25(3):1189–1231, 2015. [97](#)
- [42] René Carmona and Mathieu Laurière. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II—the finite horizon case. *arXiv preprint arXiv:1908.01613*, 2019. [12](#), [23](#)
- [43] René Carmona and Mathieu Laurière. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games i: The ergodic case. *SIAM Journal on Numerical Analysis*, 59(3):1455–1485, 2021. [12](#), [23](#)
- [44] René A. Carmona, Jean-Pierre Fouque, and Li Hsien Sun. Mean field games and systemic risk. *Communications in Mathematical Sciences*, 13(4):911–933, 2015. Publisher Copyright: © 2015 International Press. [24](#)
- [45] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145, 2011. [33](#), [68](#), [80](#), [81](#)
- [46] Antonin Chambolle and Thomas Pock. On the ergodic convergence rates of a first-order primal–dual algorithm. *Mathematical Programming*, 159(1):253–287, 2016. [33](#), [68](#), [80](#), [81](#)
- [47] Patrick Cheridito and Michael Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(01):137–162, 2011. [47](#)
- [48] Patrick L. Combettes. Perspective functions: Properties, constructions, and examples. *Set-Valued and Variational Analysis*, 26(2):247–264, 2018. [50](#), [73](#)
- [49] Romain Couillet, Samir M. Perlaza, Hamidou Tembine, and Mérouane Debbah. Electrical vehicles in the smart grid: A mean field game analysis. *IEEE Journal on Selected Areas in Communications*, 30(6):1086–1096, 2012. [97](#)
- [50] Kai Cui and Heinz Koepl. Approximately solving mean field games via entropy-regularized deep reinforcement learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1909–1917. PMLR, 2021. [12](#), [23](#), [38](#), [98](#)
- [51] François Delarue and Athanasios Vasileiadis. Exploration noise for learning linear-quadratic mean field games. *arXiv preprint arXiv:2107.00839*, 2021. [38](#), [98](#)
- [52] Josu Doncel, Nicolas Gast, and Bruno Gaujal. A mean field game analysis of SIR dynamics with vaccination. *Probability in the Engineering and Informational Sciences*, pages 1–18, 2020. [97](#)
- [53] Joseph C. Dunn and S. Harshbarger. Conditional gradient algorithms with open loop step size rules. *Journal of Mathematical Analysis and Applications*, 62(2):432–444, 1978. [98](#), [115](#)
- [54] Romuald Elie, Emma Hubert, and Gabriel Turinici. Contact rate epidemic control of covid-19: an equilibrium view. *Mathematical Modelling of Natural Phenomena*, 15:35, 2020. [97](#)

- [55] Romuald Elie, Julien Pérolat, Mathieu Laurière, Matthieu Geist, and Olivier Pietquin. Approximate fictitious play for mean field games. *arXiv preprint arXiv:1907.02633*, 2019. [12](#), [23](#), [68](#), [98](#)
- [56] Olivier Féron, Peter Tankov, and Laura Tinsi. Price formation and optimal trading in intraday electricity markets. *arXiv preprint arXiv:2009.04786*, 2020. [24](#), [97](#)
- [57] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011. [47](#)
- [58] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3):707–738, Aug 2015. [3](#), [15](#), [28](#), [42](#), [59](#)
- [59] Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956. [36](#), [98](#), [115](#)
- [60] Drew Fudenberg and David K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998. [11](#), [23](#), [98](#)
- [61] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & mathematics with applications*, 2(1):17–40, 1976. [34](#), [68](#)
- [62] Matthieu Geist, Julien Pérolat, Mathieu Laurière, Romuald Elie, Sarah Perrin, Olivier Bachem, Rémi Munos, and Olivier Pietquin. Concave utility reinforcement learning: the mean-field game viewpoint. *arXiv preprint arXiv:2106.03787*, 2021. [12](#), [23](#), [35](#), [36](#), [38](#), [98](#), [99](#), [118](#)
- [63] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. springer, 2015. [122](#)
- [64] Roland Glowinski and Americo Marroco. Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 9(R2):41–76, 1975. [34](#), [68](#)
- [65] Diogo A. Gomes, Joana Mohr, and Rafael Rigao Souza. Discrete time, finite state space mean field games. *Journal de mathématiques pures et appliquées*, 93(3):308–328, 2010. [30](#), [67](#)
- [66] Diogo A. Gomes and Joao Saude. A mean-field game approach to price formation. *Dynamic Games and Applications*, pages 1–25, 2020. [30](#), [35](#), [68](#)
- [67] P. Jameson Graber and Alain Bensoussan. Existence and uniqueness of solutions for Bertrand and Cournot mean field games. *Applied Mathematics & Optimization*, pages 1–25, 2015. [30](#), [41](#), [68](#), [97](#)
- [68] P. Jameson Graber, Vincenzo Ignazio, and Ariel Neufeld. Nonlocal bertrand and cournot mean field games with general nonlinear demand schedule. *Journal de Mathématiques Pures et Appliquées*, 148:150–198, 2021. [30](#), [41](#), [68](#), [97](#)
- [69] P. Jameson Graber, Alpár R. Mészáros, Francisco J. Silva, and Daniela Tonon. The planning problem in mean field games as regularized mass transport. *Calculus of Variations and Partial Differential Equations*, 58(3):1–28, 2019. [35](#)

- [70] P. Jameson Graber and Charafeddine Mouzouni. Variational mean field games for market competition. In *PDE models for multi-agent phenomena*, pages 93–114. Springer, 2018. [30](#), [68](#), [97](#)
- [71] P. Jameson Graber and Charafeddine Mouzouni. On mean field games models for exhaustible commodities trade. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:11, 2020. [97](#)
- [72] P. Jameson Graber, Alan Mullenix, and Laurent Pfeiffer. Weak solutions for potential mean field games of controls. *Nonlinear Differential Equations and Applications NoDEA*, 28(5):1–34, 2021. [30](#), [68](#), [97](#)
- [73] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. *Mean Field Games and Applications*, pages 205–266. Springer, Berlin, Heidelberg, 2011. [41](#)
- [74] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer, 2011. [97](#)
- [75] Xin Guo, Anran Hu, Renyuan Xu, and Junzi Zhang. A general framework for learning mean-field games. *arXiv preprint arXiv:2003.06069*, 2020. [12](#), [23](#)
- [76] Saeed Hadikhanloo. Learning in anonymous nonatomic games with applications to first-order mean field games. *arXiv preprint arXiv:1704.00378*, 2017. [12](#), [23](#)
- [77] Saeed Hadikhanloo and Francisco J. Silva. Finite mean field games: Fictitious play and convergence to a first order continuous mean field game. *Journal de Mathématiques Pures et Appliquées*, 132:369 – 397, 2019. [12](#), [23](#), [38](#), [68](#), [98](#), [112](#)
- [78] Bingsheng He, Min Tao, and Xiaoming Yuan. Alternating direction method with gaussian back substitution for separable convex programming. *SIAM Journal on Optimization*, 22(2):313–340, 2012. [34](#), [69](#), [85](#)
- [79] Minyi Huang, Peter E Caines, and Roland P Malhamé. Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized  $\varepsilon$ -nash equilibria. *IEEE transactions on automatic control*, 52(9):1560–1571, 2007. [42](#)
- [80] Minyi Huang, Roland P. Malhamé, and Peter E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006. [1](#), [13](#), [41](#), [67](#), [97](#)
- [81] Martin Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *International Conference on Machine Learning*, pages 427–435. PMLR, 2013. [39](#), [98](#)
- [82] Thomas Kerdreux. *Accelerating conditional gradient methods*. PhD thesis, Université Paris sciences et lettres, 2020. [98](#)
- [83] Ziad Kobeissi. On classical solutions to the mean field game system of controls. *arXiv preprint arXiv:1904.11292*, 2019. [41](#)

- [84] Aimé Lachapelle, Jean-Michel Lasry, Charles-Albert Lehalle, and Pierre-Louis Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *Mathematics and Financial Economics*, 10(3):223–262, 2016. [97](#)
- [85] Aimé Lachapelle and Marie-Therese Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation research part B: methodological*, 45(10):1572–1589, 2011. [97](#)
- [86] Daniel Lacker. On the convergence of closed-loop nash equilibria to the mean field game limit. *The Annals of Applied Probability*, 30(4):1693–1761, 2020. [29](#)
- [87] Olga Aleksandrovna Ladyzhenskaia, Vsevolod Alekseevich Solonnikov, and Nina N Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988. [5](#), [17](#), [100](#), [119](#)
- [88] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i—the cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006. [1](#), [13](#), [41](#), [67](#), [97](#)
- [89] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. ii—horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006. [1](#), [13](#), [30](#), [41](#), [67](#), [68](#), [97](#)
- [90] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007. [1](#), [13](#), [41](#), [67](#), [97](#)
- [91] Mathieu Lauriere. Numerical methods for mean field games and mean field type control. *arXiv preprint arXiv:2106.06231*, 2021. [11](#), [23](#)
- [92] Jacques-Louis Lions. *Optimal control of systems governed by partial differential equations*, volume 170. Springer Verlag, 1971. [117](#)
- [93] Alpár Richárd Mészáros and Francisco J. Silva. A variational approach to second order mean field games with density constraints: the stationary case. *Journal de Mathématiques Pures et Appliquées*, 104(6):1135–1159, 2015. [30](#), [68](#), [97](#)
- [94] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996. [98](#)
- [95] Jun Moon and Tamer Başar. Linear quadratic risk-sensitive and robust mean field games. *IEEE Transactions on Automatic Control*, 62(3):1062–1077, March 2017. [25](#), [42](#)
- [96] Charafeddine Mouzouni. *Topic in mean field games theory & applications in economics and quantitative finance*. PhD thesis, Université de Lyon, 2019. [12](#), [23](#), [24](#)
- [97] Levon Nurbekyan et al. Fourier approximation methods for first-order nonlocal mean-field games. *Portugaliae Mathematica*, 75(3):367–396, 2019. [12](#), [23](#)
- [98] Basel Committee on Banking Supervision. Messages from the academic literature on risk measurement for the trading book, 2011. [24](#), [42](#)
- [99] Carlo Orrieri, Alessio Porretta, and Giuseppe Savaré. A variational approach to the mean field planning problem. *Journal of Functional Analysis*, 277(6):1868–1957, 2019. [35](#)

- [100] Mario V. F. Pereira and Leontina M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1):359–375, 1991. [29](#)
- [101] Julien Perolat, Sarah Perrin, Romuald Elie, Mathieu Laurière, Georgios Piliouras, Matthieu Geist, Karl Tuyls, and Olivier Pietquin. Scaling up mean field games with online mirror descent. *arXiv preprint arXiv:2103.00623*, 2021. [12](#), [23](#), [38](#), [98](#)
- [102] Sarah Perrin, Mathieu Laurière, Julien Pérolat, Matthieu Geist, Romuald Élie, and Olivier Pietquin. Mean field games flock! the reinforcement learning way. *arXiv preprint arXiv:2105.07933*, 2021. [38](#), [98](#)
- [103] Sarah Perrin, Julien Pérolat, Mathieu Laurière, Matthieu Geist, Romuald Elie, and Olivier Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. *arXiv preprint arXiv:2007.03458*, 2020. [12](#), [23](#), [38](#), [98](#)
- [104] Laurent Pfeiffer. Optimality conditions in variational form for non-linear constrained stochastic control problems. *Mathematical Control & Related Fields*, 10(3):493–526, 2020. [56](#)
- [105] Pieper, Konstantin and Walter, Daniel. Linear convergence of accelerated conditional gradient algorithms in spaces of measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:38, 2021. [98](#)
- [106] Adam Proskinski and Filippo Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. *Stochastics*, 89(6-7):923–942, 2017. [30](#), [68](#), [97](#)
- [107] Alain Rakotomamonjy, Rémi Flamary, and Nicolas Courty. Generalized conditional gradient: analysis of convergence and applications. *arXiv preprint arXiv:1510.06567*, 2015. [98](#), [110](#)
- [108] Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pages 296–301, 1951. [98](#)
- [109] R. Tyrrell Rockafellar. *Convex analysis*, volume 36. Princeton university press, 1970. [11](#), [22](#), [76](#)
- [110] Andrzej Ruszczyński. Risk-averse dynamic programming for markov decision processes. *Mathematical Programming*, 125(2):235–261, October 2010. [29](#), [46](#), [47](#), [48](#)
- [111] Andrzej Ruszczyński and Alexander Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006. [24](#), [41](#)
- [112] Lars Ruthotto, Stanley J. Osher, Wuchen Li, Levon Nurbekyan, and Samy Wu Fung. A machine learning framework for solving high-dimensional mean field game and mean field control problems. *Proceedings of the National Academy of Sciences*, 117(17):9183–9193, 2020. [12](#), [23](#)
- [113] Naci Saldi, Tamer Basar, and Maxim Raginsky. Markov–Nash equilibria in mean-field games with discounted cost. *SIAM Journal on Control and Optimization*, 56(6):4256–4287, 2018. [25](#), [29](#), [42](#), [65](#)

- 
- [114] Filippo Santambrogio. A modest proposal for MFG with density constraints. *Networks & Heterogeneous Media*, 7(2):337–347, 2012. [30](#), [68](#)
- [115] Filippo Santambrogio. Crowd motion and evolution PDEs under density constraints. *ESAIM: Proceedings and Surveys*, 64:137–157, 2018. [30](#), [68](#)
- [116] Alexander Shapiro. Minimax and risk averse multistage stochastic programming. *European Journal of Operational Research*, 219(3):719 – 726, 2012. Feature Clusters. [24](#), [41](#)
- [117] Jayakumar Subramanian and Aditya Mahajan. Reinforcement learning in stationary mean-field games. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 251–259, 2019. [12](#), [23](#)
- [118] Hamidou Tembine, Quanyan Zhu, and Tamer Başar. Risk-sensitive mean-field games. *IEEE Transactions on Automatic Control*, 59(4):835–850, April 2014. [25](#), [42](#)
- [119] Muhammad Aneeq uz Zaman, Kaiqing Zhang, Erik Miehling, and Tamer Başar. Reinforcement learning in non-stationary discrete-time linear-quadratic mean-field games. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 2278–2284. IEEE, 2020. [12](#), [23](#)
- [120] Cédric Villani. *Optimal transport: Old and New*. Springer Verlag, 2008. [43](#)
- [121] Hong-Kun Xu. Convergence analysis of the Frank-Wolfe algorithm and its generalization in Banach spaces. *arXiv preprint arXiv:1710.07367*, 2017. [98](#)

**Titre :** Jeux à champ moyen : méthodes numériques et cas d'agents averses au risque

**Mots clés :** Jeux à champ moyen, mesures de risque, jeux potentiels, contraintes dures, gradient conditionnel généralisé, apprentissage.

**Résumé :**

Les jeux à champ moyen (abrégés MFG) sont à la fois une théorie mathématique et un outil de modélisation. Développés indépendamment en 2006 par Jean-Michel Lasry et Pierre-Louis Lions, et Minyi Huang, Roland P. Malhamé, et Peter E. Caines, les MFG offrent un cadre particulièrement adapté pour analyser les interactions stratégiques entre un grand nombre de joueurs rationnels et anonymes. Dans cette thèse, nous proposons plusieurs développements à cette théorie :

1) En utilisant le concept de mesure de risque composite, nous étudions un modèle MFG en temps discret impliquant des agents averses au risque. Nous montrons l'existence d'une solution via une approche de point fixe. Nous montrons qu'une politique optimale du MFG est  $\varepsilon(N)$ -optimale pour un certain jeu à  $N$  joueurs. La suite  $\varepsilon(N)$  converge vers zéro lorsque le nombre de joueurs tend vers l'infini.

2) Nous étudions des MFG potentiels (aussi appelés variationnels) en espace de temps discret et en espace d'état fini avec des contraintes dures, c'est-à-dire avec des potentiels convexes, éventuellement

non différentiables et à domaine borné. Nous étudions un problème primal et un problème dual, et nous montrons : un résultat de dualité, l'existence et l'unicité (dans le cas différentiable) d'une solution au système MFG. Ensuite, nous implémentons deux familles de méthodes numériques : des méthodes proximales primales-duales (Chambolle-Pock et Chambolle-Pock-Bregman) et des méthodes de Lagrangien augmenté (ADMM et ADM-G). Nous proposons un modèle de congestion et un modèle de prix que nous résolvons avec ces méthodes. Nous comparons les performances empiriques de chacune des méthodes pour chaque problème.

3) Nous appliquons l'algorithme du gradient conditionnel généralisé pour les MFG potentiels, dans un cadre EDP. Nous mettons en évidence le lien entre cet algorithme et une méthode d'apprentissage appelée fictitious play. On montre que pour le taux d'apprentissage  $\delta_k = 2/(k+2)$ , le coût potentiel converge en  $O(1/k)$ ; l'exploitabilité et les variables du problème convergent en  $O(1/\sqrt{k})$ , pour des normes spécifiques.

**Title :** Mean field games: numerical methods and case of risk-averse agents

**Keywords :** Mean field games, risk measures, potential games, hard constraints, generalized conditional gradient, learning.

**Abstract :**

Mean field games (abbreviated MFGs) are both a mathematical theory and a modeling tool. Developed in 2006 independently by Jean-Michel Lasry and Pierre-Louis Lions and Minyi Huang, Roland P. Malhamé, and Peter E. Caines, MFGs provide a framework to analyze interactions among a large number of rational and anonymous agents. In this thesis we provide several developments to this theory:

1) Using the concept of composite risk measure, we study a discrete-time MFG model involving risk-averse agents. We show the existence of a solution via a fixed point approach. We show that an optimal policy of the MFG is  $\varepsilon(N)$ -optimal for a certain  $N$ -player game. The sequence  $\varepsilon(N)$  converges to zero as the number of players tends to infinity.

2) We study discrete time and finite state space potential (also called variational) MFGs with hard constraints, that is with convex potentials, possi-

bly non-differentiable and with bounded domain. We study a primal and a dual problem, and we show: a duality result, the existence and uniqueness (in the differentiable case) of a solution to the MFG system. Then we implement two families of numerical methods: primal-dual proximal methods (Chambolle-Pock and Chambolle-Pock-Bregman) and augmented Lagrangian based methods (ADMM and ADM-G). We propose a congestion model and a price model that we solve with these methods. We compare the empirical performance of each method for each problem.

3) We apply the generalized conditional gradient algorithm for potential MFGs in a PDE framework. We highlight the connection between this algorithm and a learning method called fictitious play algorithm. We show that for the learning rate  $\delta_k = 2/(k+2)$ , the potential cost converges in  $O(1/k)$ ; the exploitability the variables of the problem converge in  $O(1/\sqrt{k})$ , for specific norms.