ABSTRACT

Mean Field Games of Controls and Moderate Interactions

Alan B. Mullenix, Ph.D.

Mentor: P. Jameson Graber, Ph.D.

Mean Field Game theory, a quickly growing field emerging around 2007, seeks to answer qualitative and quantitative questions about high population competitive dynamics. A typical Mean Field Game is comprised of a large pool of identical, weak, rational agents, each seeking to maximize a payout (or, equivalently, minimize a cost). When an agent selects a strategy, it alters the state of the playing field for all other agents, that is, the actions of all agents must be factored into the strategy choice of any agent. The main object of interest is the Nash Equilibria of the game: a strategy choice that, if chosen by all players, no single player may improve their outcome by adjusting their own strategy. The question of these Nash Equilibria can be cast as solutions of a coupled system of partial differential equations: a backward in time Hamilton Jacobi Equation that encapsulates the strategy costs, the unknown being the value function $u(x,t)$, and a forward in time Fokker-Planck type equation governing the evolution $m(x,t)$:

the probability distribution of player states. We look at two adaptations of the traditional Mean Field Game. First, a Potential Mean Field Game of Controls, in which agents have knowledge of not only the current states of other agents, but also their current adjustment to strategy. While the potential structure allows for the use of convex analysis techniques, we introduce a new, spatially non-local coupling component that depends on both the distribution of player states and the feedback. We also consider a Mean Field Game of Moderate Interactions, in which agents are more intensely affected by those in their immediate vicinity. This augments the standard Mean Field Game by introducing a local coupling term that has interaction with the gradient of the value function. We establish, in each case, well posedness results under generic assumptions on the data. In the conclusion we remark on the possible generalizations and further directions to take each work.

Mean Field Games of Controls and Moderate Interactions

by

Alan B. Mullenix, B.S., M.S.

A Dissertation

Approved by the Department of Mathematics

Dorina Mitrea, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by the Dissertation Committee

P. Jameson Graber, Ph.D., Chairperson

David Ryden, Ph.D.

Rob Kirby, Ph.D.

Qin Sheng, Ph.D.

David Kahle, Ph.D.

Accepted by the Graduate School

May 2022

J. Larry Lyon, Ph.D., Dean

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Jameson Graber, a source of expertise in seemingly endless arenas. The topic of this dissertation is at the confluence of many fields, and the guidance through what often seems like a hopelessly tangled mess was invaluable to the process. To provide all of this is worth thanks enough, but to do so also with patience, insight, and generosity, I could not have asked for more in a mentor.

I would also like to thank the committee for their time and involvement throughout my doctoral journey. Additionally, I'm grateful for the Baylor University Department of Mathematics, in particular chairpersons and directors Dr. Dorina Mitrea, Dr. Lance Littlejohn, and Dr. Mark Sepanski, for their unwavering backing over the past six years.

A special thanks go to Renea Randle, Gail Queen, and Dr. Michael Green, who shed a light on and nurtured my growth in mathematics so many years ago. The trajectory of my life was changed in ways I would have never considered possible up to that moment, and things accomplished I would have never felt qualified to pursue.

Finally, I would like to thank Dr. Alexandra Pasi, not, as it were, for being formative in this work, but for sharing wholly in life's conjectures, propositions, and proofs; in all the unwritten chapters yet to follow those here.

To Renea, Gail, Michael

CHAPTER ONE

Introduction

A Mean Field Game (MFG) for our purposes is a very high population differential game between identical rational agents who are relatively weak (an individual has little power to effect the entire population). MFG with non-identical or singular powerful agents are an active field of study as well, such as a monopolistic corporation competing in a market with smaller players, but we concentrate on the former. We motivate the main ideas of our formulation of a mean field game with a short scenario. Suppose a large group of several hundred people have just exited a building in a tight clump and notice they are all going to miss the bus a few blocks away if they don't hurry. We know each person now has desire to reach the bus stop *quickly*. Given this desire, the optimal "strategy" for each person to take is travel as fast as possible in a straight line for the bus stop, but everyone doing this en masse may cause congestion, slowing the group, resulting in many or all individuals to miss out on the bus. Each person choosing their individual optimal strategy resulted not only in a poor outcome for the group, but for the person themselves. A single enterprising person might expect the congestion clumping and opt to travel a longer, but unimpeded path around the group. Recall from above, however, that we consider all players equally enterprising, and so even players being clever may cause sub-optimal outcomes when that cleverness is in the aggregate. We have then a situation where each player attempting to optimize their end result must contend with an ever changing environment that is evolving with the collective choices of all players. A natural question to ask is whether or not there is a Nash Equilibrium strategy. Such a strategy adopted by all players, is individually protective in the following sense: assuming all players follow the strategy, any individual player cannot do better by adjusting their own away from the equilibrium. An example of a Nash Equilibrium can be seen in the classical game theoretic “Prisoner’s Dilemma”, where two separated prisoners each must choose to confess or stay silent. If both confess, they each receive a four year sentence. If they both stay silent, each receives only two years. If one confesses and the other remains silent, the confessing prisoner receives one year while the silent receives eight. Clearly, the best outcome for both prisoners occurs when they each stay silent, but individually the best outcome is to confess while hoping the other stays silent. Assuming that the other might confess, and fearing the eight year sentence, both prisoners choose to confess. The outcome is sub-optimal both for the group and the individual, but there is a protective mutually assured outcome. In what follows, we will be seeking solutions to a coupled system of partial differential equations. These solutions are the Nash Equilibrium strategies.

1.1 Mean Field Games

A typical mean field game consists of competitors seeking to minimize a “cost”:

$$J(\alpha, 0, T) = \mathbb{E}\left[\int_0^T \left(\underbrace{L(X_t, \alpha)}_{\text{Running Cost}} + \underbrace{f(X_t, m(X_t, t))}_{\text{“Congestion Penalty”}}\right) \mathrm{d}t + \underbrace{u_T(X_T)}_{\text{Final Cost}}\right]$$

with dynamics $\mathrm{d}X_t = \alpha\,\mathrm{d}t +$ “noise” (a Brownian Motion) driven by control α. The function $m(x, t)$ describes the distribution of player states. As noted in the motivating example, a vital feature is that the cost function for each player depends upon the state of *all* players. Examples of running costs range from literal monetary costs along the way, to fuel expenditure, distance travelled, or considered a “negative” cost, and looked at as end payout, or the maximization of a desired outcome. A final cost (or alternatively, payout) might be a scaling penalty/reward based on proximity to a goal outcome. The term $f(x, m)$ in the integrand, labeled colloquially “congestion penalty” to align with the motivating example, is the mechanism by which the individual player’s optimal control problem is coupled with the crowd at large. If for example, f is monotonically increasing with respect to m, that would model a penalty for traveling through a high density area. However, f could introduce reward for clustering up, penalty/reward for uniformly spreading out, even repulsion/attraction for a particular state. One could further let f assume a value of $+\infty$ out side of a particular range to impose an “infinite penalty”. For our purposes, f will be increasing with respect to the density m.

1.2 From Mean Field Game to PDE

The *Hamiltonian* $H(x, p)$ is the convex conjugate with respect to the second variable of the running cost (termed the Lagrangian), that is:

$$H(x, p) := \sup_{v} \left[p \cdot v - L(x, v) \right]$$

The Lagrangian is taken to be convex in the second variable, and thus H is as well. This convexity turns out to be vital to well posedness considerations of the Nash Equilibria in many treatments of the field.

A *value function* $u(x,t)$ of such a finite time horizon optimal control problem can be defined as the infimum of the above costs over all appropriate controls:

$$u(x,t) := \inf_{\alpha} J(\alpha, t, T)$$

This value function u satisfies, in a viscosity sense, a backward in time Hamilton Jacobi equation:

$$-u_t - \sigma\Delta u + H(x, \nabla u) = f(x,m); \quad u(T) = u_T.$$

The probability density of player states, $m(x,t)$ evolves via a forward in time Fokker-Planck type equation, where m_0 is the given initial state of players:

$$m_t - \sigma\Delta u - \nabla \cdot (mD_pH(x, \nabla u)) = 0; \quad m(x,0) = m_0.$$

Together, these form a coupled system of PDE, the solutions of which are the value function and player density evolution for Nash Equilibrium strategies.

1.3 Historical Development and Related Current Works

The theory of Mean Field Games (a term borrowed from physics), was developed simultaneously in 2006-2007 by Lasry, Lions [39], and Huang, Malhamé, Caines [33] (under the name "Nash certainty equivalence principle"). The main idea (that the impossibility of players to assess all the relevant details of competitors in fact doesn't *matter*) was already studied in the case of more traditional, static games by Aumann [1]. The speed at which this field has grown means contributions are both numerous and intertwined, but we attempt to lay out the major highlights. The probabalistic approach has been well studied by Carmona and Delarue, especially in the two volumes [14] and [15]. See also the Paris-Princeton lectures taught by Lasry [32] and include work by Lions and Guéant. A thorough PDE based introduction to MFG can be found in Cardaliaguet and Poretta's monograph [13] with contributed work from Santambrogio (lecture notes on variational mean field games), Delarue (master equation considerations, see below), Achdou and Laurière (applications and numerics). Weak solution existence and uniqueness is proved for first and second order MFG with local coupling by Cardaliaguet, Graber in [10] and Cardaliaguet, Graber, Poretta, Tonon in [11]. Duality techniques that inspire Chapter Two can be found in [30] by Graber and Mouzouni.

Before moving on to the context of the manuscripts that appear here, we briefly mention the other large way in which PDE manifest in the study of Mean Field Games: the so-called "master equation". As outlined by Cardaliaguet and Poretta in [13], the master equation is a PDE in the space of measures (infinite space dimension) which does not have some of the primary weaknesses of the mean field game system of PDE we've described. In particular, it allows one to deal with global noise in the game system, which is clearly important to modeling real world situations. We do not discuss the master equation here, but it would be folly to not nod to it when discussing PDE in MFG theory. In exchange for overcoming these weaknesses however, it is of course a challenging object of study. See [25], [8].

Turning to the context of Chapter Two (Mean Field Games of Controls), both stationary and deterministic iterations of the problem have been treated by Gomes et. al in [22] and [24] respectively. The suffix we employ: "of controls" was introduced in 2017 [12] by Cardialiaguet and LeHalle, who looked at an application to trade crowding where the mean field contained only the actions and not states of traders, showing that nonetheless traders attempting to optimize their individual outcomes ended up in a MFG Nash Equilibrium strategy. We quote from that article to stress the interesting phenomena of the aggregating interaction despite the individual's efforts of optimal control: *"Let us underline that, in our model, the market participants do not have access to the distribution of the trading positions of the other participants; they do not necessarily have the same estimate of the permanent market impact; they are not even aware that they are "playing a game". Nevertheless, the configuration after stabilization is an MFG equilibrium.".* In the 2019 article [5], Bonnans, Hadikhanloo, and Pfeiffer gave existence of a unique classical solution for a time-dependent, non-deterministic model of the problem, in which the limitation of bounded controls in [12] was done away with. To note, Mean Field Games of Controls are also referred to as Extended Mean Field Games, and returning to the probabilistic approaches aforementioned have been studied in [14] by Carmona and Delarue, and in [16] by Carmona and Lacker.

As to the context of Chapter Three (Mean Field Games of Moderate Interactions), we investigate the model studied by Flandoli, Ghio, and Livieri in 2021 [21]. The motivating consideration of a Mean Field Game of Moderate Interactions is that in certain models, such as physical congestion, it makes sense that agents interact only with other nearby agents. On the other hand, one doesn't expect this in, say, a trading scenario (where traders may have uniform access to information). However, one might speculate a situation where purchasing power/interest concentrates mostly on "nearby" trader interactions, i.e. those with similar assets and aims. In the

Flandoli article, the existence of weak solutions are given on any finite time horizon, with uniqueness results holding only when the time horizon is sufficiently small. In this work we show the existence of classical solutions on any finite time horizon. A prior contribution to this area in the context of pedestrian crowd modeling is that of Aurell, Djehiche [2].

1.4 Main Results of this Work

Two alterations that will be discussed in this work are in Chapter Two with introduction of a new, spatially non-local coupling element in the second argument of the Hamiltonian, which will appear as

$$-u_t - \sigma\Delta u + H\left(x, \nabla u - G(x)^{\intercal}\Psi\left(\int_{\mathbb{T}^d} G(x)vm\,\mathrm{d}x\right)\right) = f(x,m)$$

$$v = -D_pH\left(x, \nabla u - G(x)^{\intercal}\Psi\left(\int_{\mathbb{T}^d} G(x)vm\,\mathrm{d}x\right)\right),$$

the second line being a new fixed point condition that must be satisfied that does not appear in a traditional MFG presentation. In Chapter Three an additional local coupling term is introduced in the Hamiltonian as the form of a vector field $b(x,m)$, appearing as

$$-u_t - \sigma\Delta u + \big[H(x,\nabla u) - b(x,m)\cdot\nabla u\big] = f(x,m).$$

In Chapter Two we exploit the potential structure of the MFG to apply convex analysis techniques, viewing the game as two minimization problems in duality to each other. Relaxation of the dual problem then leads to existence and uniqueness of weak solutions to the MFG. Further space and time regularity results are given provided some additional assumptions of the behaviors of the coupling f and Hamiltonian H.

In Chapter Three we use regularity bootstrapping (using regularity of the Hamilton Jacobi equation to obtain results for the Fokker-Planck and vice versa) to prove the a priori estimates required to apply the Leray-Schauder Fixed Point Theorem. From this we obtain existence of classical solutions for any finite time horizon. We go on in Chapter Four to allude to further obtainable results applying various techniques in the spirit of Gomes [23].

1.5 Notation and Preliminary Definitions

A few common spaces and pieces of notation will be used throughout that we collect here.

In both manuscripts we will be working over the set

$$Q := \mathbb{T}^d \times [0, T],$$

a space-time cylinder with $\mathbb{T}^d$ being the flat d-dimensional torus.

We recall that the *weak derivative* of a function f on a set A is a function g such that for any smooth function ϕ, compactly supported in A, we have

$$\int_A f\phi' \, dx = -\int_A g\phi \, dx.$$

Let $\alpha \in (0, 1)$ and let A be a compact set. A function f is said to be uniformly Hölder continuous on A with index α if the quantity

$$[f]_{\alpha;A} := \sup_{x,y \in A;\ x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

is finite. The Hölder spaces, denoted $\mathcal{C}^{k,\alpha}(A)$, are subspaces of $\mathcal{C}^k(A)$ whose k-th order partial derivatives are uniformly Hölder continuous on A with index α. These are Banach spaces, with norm

$$\|u\|_{\mathcal{C}^{k,\alpha}(A)} = \|u\|_{\mathcal{C}^k(A)} + \sum_{|\beta|=k} [D^\beta u]_{\alpha;A},$$

where β is a multi-index, and the $\mathcal{C}^k$ norm above is given by

$$\|u\|_{\mathcal{C}^k} = \sum_{|\beta|=k} \|D^\beta u\|_\infty.$$

These spaces enjoy the useful inclusion that if $0 \leqslant k+\alpha \leqslant m+\beta$ with k, m integers and $\alpha, \beta \in (0,1)$,

$$\mathcal{C}^{m,\beta} \to \mathcal{C}^{k,\alpha} \quad \text{is a continuous embedding.}$$

In the case that $k = m$ and $\alpha < \beta$, bounded sets are precompact in the embedding. The Sobolev spaces, denoted $W^{n,k}(A)$, consist of k times weakly differentiable functions on A, with each weak derivative being a member of $L^n(A)$. These are Banach spaces with norm

$$\|f\|_{W^{n,k}(A)} = \sum_{|\alpha|\leqslant k} \|\partial_\alpha f\|_{L^n(A)},$$

where α is a multi-index.

1.6 Attribution

The contents of Chapter Two are in collaboration with Dr. P. Jameson Graber and Dr. Laurent Pfeiffer. We confirm that all authors contributed equally to this work, in the matters of planning, organization, notation, proof, applications, typesetting, and revision. Further, the chapter was reproduced from Nonlinear Differential Equations and Applications (NoDEA) and Springer.

CHAPTER TWO

Weak Solutions for Potential Mean Field Games of Controls

Graber, P.J., Mullenix, A. & Pfeiffer, L. Weak solutions for potential mean field games of controls. Nonlinear Differ. Equ. Appl. 28, 50 (2021).

2.1 Abstract

We analyze a system of partial differential equations that model a potential mean field game of controls, briefly MFGC. Such a game describes the interaction of infinitely many negligible players competing to optimize a personal value function that depends in aggregate on the state and, most notably, control choice of all other players. A solution of the system corresponds to a Nash Equilibrium, a strategy for which no one player can improve by altering only their own action. We investigate the second order, possibly degenerate, case with non-strictly elliptic diffusion operator and local coupling function. The main result exploits potentiality to employ variational techniques to provide a unique weak solution to the system, with additional space and time regularity results under additional assumptions. New analytical subtleties occur in obtaining a priori estimates with the introduction of an additional coupling that depends on the state distribution as well as feedback.

2.2 Introduction

Mean Field Games (MFG), introduced simultaneously in 2006-7 by J.-M. Lasry, P.-L. Lions [38] and M. Huang, R. Malhamé, P. Caines [33], have seen swift development into a vibrant and substantial subfield of partial differential equations. See, for instance, the monographs [14, 15, 4]. Considered are high population games of homogeneous, negligibly powerful players all attempting to optimize a cost while contending with the effects of the choices of all other players.

The term Mean Field, inspired by physics, relates to each player viewing the remaining players as one large entity. The cost functional that has to be optimized by each player typically incorporates an interaction term $f(m)$, where m denotes the distribution of player states. Mean Field Games of Controls (briefly, MFGC), also called Extended Mean Field Games, introduce a control element into the Mean Field, so that not only can players "detect" (via the Mean Field) the positions of others, but also their control choices. Such an extension naturally arises in many applications, for example in economics [32, 17, 18, 26, 30, 31, 27]. MFGC have been studied by D. Gomes and V. Voskanyan, who have results on classical solutions with S. Patrizi in the stationary (time independent) second order case where the diffusion is explicitly the Laplacian [22], and also in the time dependent first order case [24]. In the second order uniformly parabolic time dependent case, Z. Kobeissi has proved the existence of classical solutions under sufficient structural and smoothness assumptions, with uniqueness under additional assumptions, as well as results on approximate solutions [35, 36]. P. Cardaliaguet and C.-A. Lehalle have provided a theorem giving the existence of weak solutions to a general system of MFGC, under the assumption that the Lagrangian is monotone with respect to the measure variable and that the Hamiltonian is sufficiently smooth; in particular it must depend on the density of players nonlocally [12]. See also R. Carmona and F. Delarue [14] for a probabalistic approach to Extended Mean Field Games.

In this article, we investigate the second order degenerate case (which can, in particular, be first order) featuring a non-strictly elliptic diffusion operator with space dependence. The MFGC system to be studied is

$$\begin{cases} (i) \quad -\partial_t u - A_{ij}\partial_{ij} u + H\left(x, Du(x,t) + G(x)^\intercal P(t)\right) = f\left(x, m(x,t)\right) & (x,t) \in Q, \\ (ii) \quad \partial_t m - \partial_{ij}\left(A_{ij} m\right) + \nabla \cdot (vm) = 0 & (x,t) \in Q, \\ (iii) \quad P(t) = \Psi\Big(\int_{\mathbb{T}^d} G(x) v(x,t) m(x,t)\,\mathrm{d}x\Big) & t \in [0,T], \\ (iv) \quad v(x,t) = -D_\xi H\left(x, Du(x,t) + G(x)^\intercal P(t)\right) & (x,t) \in Q, \\ (v) \quad m(x,0) = m_0(x), \quad u(x,T) = u_T(x), & x \in \mathbb{T}^d \end{cases} \tag{2.1}$$

where u, m are scalar functions, v is a vector field in $\mathbb{R}^d$, $P = P(t) \in \mathbb{R}^k$, $Q := \mathbb{T}^d \times [0,T]$, and $A(x) = [A_{ij}(x)]_{1 \leqslant i,j \leqslant d}$ is a given matrix-valued function on $\mathbb{T}^d$ whose values are symmetric and non-negative.

The heuristic interpretation of the above system is the following. The variable m describes the distribution of the state of the players and P is a time-dependent price variable. These two variables are interaction terms; they describe the collective behavior of the agents. Isolated agents have no impact of them. Each agent controls the following dynamical system in $\mathbb{T}^d$: $\mathrm{d}X_t = \alpha_t\,\mathrm{d}t + \sqrt{2}\Sigma(X_t)\,\mathrm{d}B_t$ where $(B_t)_{t\in[0,T]}$ is a standard Brownian motion in $\mathbb{R}^D$, α is an adapted process in $\mathbb{R}^d$, and $\Sigma\colon \mathbb{T}^d \to \mathbb{R}^{d\times D}$ is such that $A(x) = \Sigma(x)\Sigma(x)^\intercal$, for all $x \in \mathbb{T}^d$. Given the interaction terms m and P, the associated cost (to be minimized) is given by

$$\mathbb{E}\Big[\int_0^T \Big(H^*(X_t, -\alpha_t) + \langle P(t), G(X_t)\alpha_t\rangle + f(X_t, m(X_t,t))\Big)\mathrm{d}t + u_T(X_T)\Big].$$

Equation (2.1)(i) is a Hamilton-Jacobi-Bellman equation and characterizes the value function u associated with the optimal control problem to be solved by each agent, given m and P. Equation (2.1)(iv) gives the corresponding solution in feedback form:

$$\alpha_t = v(X_t, t) = -D_p H(X_t, Du(X_t, t) + G(X_t)^\intercal P(t)).$$

Assuming now that all agents employ the feedback v, the evolution of their distribution is given by the Fokker-Planck equation (2.1)(ii) (the initial distribution m_0) is fixed). Then equation (2.1)(iii) gives the price in function of m and v. When $G(x)$ is the identity matrix, P is simply a function Ψ of the average control of the agents. Let us emphasize that P induces an interaction of the agents not only through their states but also via their controls. In summary, a (mean field) Nash equilibrium is attained when v is a best reponse with respect to the interaction terms m and P (equations (i) and (iv)), and when m and P can be deduced from v via equations (ii) and (iii).

A natural economic interpretation of P is as negative market price (the negative sign is introduced so that we have a minimization problem). In this interpretation, P increases (i.e. demand decreases) in each dimension along which the total market supply of a particular good increases. See [6] for more details. Alternatively, one could interpret System (2.1) as a model of congestion penalization with an additional dispersive forcing term given by P. For example, P may be proportional to average velocity. In this case, whereas f imposes a cost corresponding to population density, P pushes agents to move in a direction opposite to that general motion of the crowd, thus encouraging the crowd to disperse.

The basic structural assumptions are

1. $\mathbb{T}^d \times \mathbb{R}^d \ni (x,\xi) \mapsto H(x,\xi) \in \mathbb{R}$ is convex in ξ
2. $\mathbb{T}^d \times [0,\infty) \ni (x,m) \mapsto f(x,m) \in \mathbb{R}$ is monotone increasing in m
3. $\mathbb{R}^k \ni z \mapsto \Psi(z) \in \mathbb{R}^k$ is monotone in z, i.e. $\langle \Psi(t,z_1) - \Psi(t,z_2), z_1 - z_2 \rangle \geqslant 0$ for all $z_1, z_2 \in \mathbb{R}^k$.

See section 2.4 for more detailed assumptions on the data.

We will focus in the article on the MFG system obtained after performing the Benamou-Brenier change of variables $w = mv$ [3]:

$$\begin{cases} (i) \quad -\partial_t u - A_{ij}\partial_{ij} u + H\left(x, Du(x,t) + G(x)^\intercal P(t)\right) = f\left(x, m(x,t)\right) & (x,t) \in Q, \\ (ii) \quad \partial_t m - \partial_{ij}\left(A_{ij} m\right) + \nabla \cdot w = 0 & (x,t) \in Q, \\ (iii) \quad P(t) = \Psi\left(\int_{\mathbb{T}^d} G(x) w(x,t)\,\mathrm{d}x\right) & t \in [0,T], \\ (iv) \quad w(x,t) = -D_\xi H\left(x, Du(x,t) + G(x)^\intercal P(t)\right) m(x,t) & (x,t) \in Q, \\ (v) \quad m(x,0) = m_0(x), \quad u(x,T) = u_T(x), & x \in \mathbb{T}^d. \end{cases} \tag{2.2}$$

It is worth mentioning that although the function w as defined above is determined by u, m, and P, labeling greatly reduces clutter in the statements and calculations to follow. For the same reason, we keep P, although it is determined by w.

In [5] the authors prove the existence of classical solutions to (2.1) when A is the identity matrix and the congestion term f is nonlocal. They also showed that the game is "potential," which means that a Nash equilibrium can be interpreted as a critical point of a suitably chosen functional, which we may call the potential. Cf. [38, Section 6.2]. When the potential is strictly convex, we formally have that the Nash equilibrium is unique, and under suitable assumptions one can show that the PDE system is well-posed. In what follows we will provide the existence and uniqueness of a suitably defined “weak solution” to the MFGC system with local coupling and provide additional regularity results involving the solution u and the distribution evolution m. The method follows that of Cardaliaguet, Graber, Porretta, and Tonon in [11]–see also [9, 10]–in the study of the case of first and second order “classical” MFG systems, which are also potential. The nonlocal interaction term $P(t)$ introduces new subtleties into the analysis, especially as it does not introduce any a priori gain of regularity. On the contrary, a priori estimates on solutions to the Hamilton-Jacobi Equation (2.2)(i) are highly sensitive to the L^p norms of $P(t)$. See Section 2.6.

We first lay out the required assumptions on the data (Section 2.4). We then view the MFGC system as a system of optimality for the minimization of a suitably defined convex potential. Next, we consider the dual problem and show that the correct relaxation of it provides existence and a.e. uniqueness of an adjoint state (Section 2.6). The solutions to these minimization problems are then shown to be proper candidates for the weak solution to the MFGC, whose existence is then proved (Section 2.7). Finally, with some additional assumptions on the data, we include some space and time regularity results for the weak solution based on previous techniques of Graber and Meśzáros [28] (Section 2.8). We now lay out the notation and assumptions to hold throughout the paper.

2.3 Notation

We denote by $\langle x, y\rangle$ the Euclidean scalar product of two vectors $x, y \in \mathbb{R}^d$ and by $|x|$ the Euclidean norm of x. We use conventions on repeated indices: for instance, if

$a, b \in \mathbb{R}^d$, we often write $a_i b_i$ for the scalar product $\langle a, b \rangle$. More generally, if A and B are two square symmetric matrices of size $d \times d$, we write $A_{ij} B_{ij}$ for $\mathrm{Tr}(AB)$.

To avoid further difficulties arising from boundary issues, we work in the flat $d-$dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with $Q := \mathbb{T}^d \times [0, T]$ for some fixed time $T > 0$. Our methods can be applied in a more or less straightforward way when the boundary conditions are of Neumann type; with some further technical assumptions they may be applied on the whole space. Other boundary conditions, which may be more appropriate for economics applications, tend to introduce greater technicalities; this is a subject for future research. We denote by $P(\mathbb{T}^d)$ the set of Borel probability measures over $\mathbb{T}^d$. It is endowed with the weak convergence. For $k, n \in \mathbb{N}$ and $T > 0$, we denote by $\mathcal{C}^k(Q, \mathbb{R}^n)$ the space of maps $G = G(t, x)$ of class $\mathcal{C}^k$ in time and space with values in $\mathbb{R}^n$. For $p \in [1, \infty]$ and $T > 0$, we denote by $L^p(\mathbb{T}^d)$ and $L^p(Q)$ the set of $p-$integrable maps over $\mathbb{T}^d$ and Q respectively. We often abbreviate $L^p(\mathbb{T}^d)$ and $L^p(Q)$ into L^p. We denote by $\|f\|_p$ the L^p-norm of a map $f \in L^p$. The conjugate of a real $p > 1$ is denoted by p', i.e. $1/p + 1/p' = 1$.

2.4 Assumptions

We now collect the assumptions on the "congestion coupling" f, the "aggregate control coupling" Ψ, the Hamiltonian H, and the initial and terminal conditions m_0 and u_T. Along the article, we assume that there exist some constants $C_1 > 0$, $C_2 > 0$, $C_3 > 0$, $C_4 > 0$, $q > 1$, $r > 1$, and $s > 1$ such that the following hypotheses hold true. We denote

$$p = q'.$$

(H1) (Conditions on the coupling)

- The map $f\colon \mathbb{T}^d \times [0, +\infty) \to \mathbb{R}$ is continuous in both variables, increasing with respect to the second variable m, and satisfies

$$\frac{1}{C_1}|m|^{q-1} - C_1 \leqslant f(x,m) \leqslant C_1|m|^{q-1} + C_1 \qquad \forall m \geqslant 0\ . \tag{2.3}$$

  Moreover $f(x,0) = 0$ for all $x \in \mathbb{T}^d$.

- The map $\Psi\colon \mathbb{R}^k \to \mathbb{R}^k$ is the continuous gradient of some convex function $\Phi\colon \mathbb{R}^k \to \mathbb{R}$. Without loss of generality, we assume that $\Phi(0) = 0$. Moreover,

$$\Phi(z) \leqslant C_2|z|^s + C_2 \qquad \forall z \in \mathbb{R}^k. \tag{2.4}$$

  Changing C_2 if necessary, we have

$$\Phi^*(z) \geqslant \frac{1}{C_2}|z|^{s'} - C_2 \qquad \forall z \in \mathbb{R}^k.$$

  If $\frac{1}{s} + \frac{1}{pr} < 1$, we assume that

$$\frac{1}{C_2}|z|^s - C_2 \leqslant \Phi(z) \qquad \forall z \in \mathbb{R}^k. \tag{2.5}$$

- The map $G\colon \mathbb{T}^d \to \mathcal{L}(\mathbb{R}^d; \mathbb{R}^k)$ is continuously differentiable. If $\frac{1}{s} + \frac{1}{pr} < 1$, we assume that it is constant.

(H2) (Conditions on the Hamiltonian) The Hamiltonian $H\colon \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous in both variables, convex and differentiable in the second variable, with $D_\xi H$ continuous in both variables, and has a superlinear growth in the gradient variable:

$$\frac{1}{C_3}|\xi|^r - C_3 \leqslant H(x,\xi) \leqslant C_3|\xi|^r + C_3 \qquad \forall (x,\xi) \in \mathbb{T}^d \times \mathbb{R}^d. \tag{2.6}$$

We note for later use that the Fenchel conjugate H^* of H with respect to the second variable, given by

$$H^*(x,\zeta) = \sup_{\gamma} \left[\langle \zeta, \gamma \rangle - H(x,\gamma)\right],$$

is continuous and satisfies similar inequalities to H (changing C_3 if necessary):

$$\frac{1}{C_3}|\xi|^{r'} - C_3 \leqslant H^*(x,\xi) \leqslant C_3|\xi|^{r'} + C_3 \qquad \forall (x,\xi) \in \mathbb{T}^d \times \mathbb{R}^d. \tag{2.7}$$

(H3) (Conditions on A) There exists a Lipschitz continuous map $\Sigma : \mathbb{T}^d \to \mathbb{R}^{d\times D}$ such that $\Sigma\Sigma^T = A$ and such that

$$|\Sigma(x) - \Sigma(y)| \leqslant C_4|x-y| \qquad \forall x, y \in \mathbb{T}^d. \tag{2.8}$$

(H4) (Conditions on the initial and terminal conditions) $u_T : \mathbb{T}^d \to \mathbb{R}$ is of class $\mathcal{C}^2$, while $m_0 : \mathbb{T}^d \to \mathbb{R}$ is a C^1 positive density (namely $m_0(x) > 0$ and $\int_{\mathbb{T}^d} m_0 \,\mathrm{d}x = 1$).

(H5) (Restrictions on the exponents). We consider 4 cases, depending on whether $s' < r$ or $s' \geqslant r$ and whether A is constant or not. In the case that A is not constant,

$$s' < r \implies s' \geqslant p; \qquad s' \geqslant r \implies r \geqslant p,$$

while when A is constant,

$$s' < r \implies \frac{s'(d+1)}{d} \geqslant p$$

$$s' \geqslant r \implies \left(s' \geqslant 1+d\right) \text{ or } \left(s' < 1+d \text{ and } \frac{s'(1+d)}{d-s'+1} > p\right)$$

As a remark, the condition $f(x,0)=0$ is just a normalization condition, which we may assume without loss of generality, as explained in [11, Section 2]. Let us compare the natures of Assumption (H5).

(a) Assumption (H5) is stronger in cases 1A and 2A than in cases 1B and 2B, respectively, that is, Assumption (H5) is stronger in the case of a non-constant A than in the constant case.

(b) If A is not constant (cases 1A and 2A), then Assumption (H5) can be summarized by $\min(s', r) \geqslant p$.

(c) If A is constant, it is easy to verify that Assumption (H5) is stronger in the case 1B ($s' < r$) than in case 2B ($s' \geqslant r$).

1. If $\Psi = 0$, then we are back to the framework of [11] and our assumptions coincide. Indeed, (2.4) is then satisfied with any $s > 1$. Taking s sufficiently close to 1, we have $1/s + 1/(rp) \geqslant 1$, so that (2.5) is not necessary, and we have $s' \geqslant r$, so that we are either in case 2A or 2B in hypothesis (H5). If A is constant, we must choose s close enough to 1, so that $s' \geqslant 1 + d$.

Let us set

$$F(x,m) = \begin{cases} \displaystyle\int_0^m f(x,\tau)d\tau & \text{if } m \geqslant 0 \\ \\ +\infty & \text{otherwise.} \end{cases}$$

Then F is continuous on $\mathbb{T}^d \times (0, +\infty)$, differentiable and strictly convex in m and satisfies

$$\frac{1}{C_1}|m|^q - C_1 \leqslant F(x,m) \leqslant C_1|m|^q + C_1 \qquad \forall m \geqslant 0, \tag{2.9}$$

changing C_1 if necessary. Let F^* be the Fenchel conjugate of F with respect to the second variable. Note that $F^*(x,a) = 0$ for $a \leqslant 0$ because $F(x,m)$ is nonnegative, equal to $+\infty$ for $m < 0$, and null at zero. Moreover,

$$\frac{1}{C_1}|a|^p - C_1 \leqslant F^*(x,a) \leqslant C_1|a|^p + C_1 \qquad \forall a \geqslant 0, \tag{2.10}$$

changing again C_1 if necessary. Most of the results in this paper hold also for time-dependent data, in particular when f and H depend on t. It suffices to have the estimates in this subsection hold uniformly with respect to t.

2.5 Two Problems in Duality

The approach that we follow consists in viewing the MFG system as an optimality condition for two convex problems, which we introduce now.

Let $\mathcal{K}_0$ be the set of all triples $(u, P, \gamma) \in \mathcal{C}^2(Q) \times \mathcal{C}^0([0,T];\mathbb{R}^k) \times \mathcal{C}^0(Q)$ satisfying

$$\begin{cases} -\partial_t u - A_{ij}\partial_{ij}u + H\left(x, Du(x,t) + G(x)^\intercal P(t)\right) &= \gamma, \\ u(x,T) &= u_T(x). \end{cases} \tag{2.11}$$

The associated cost is given by

$$D(u,P,\gamma) = -\int_{\mathbb{T}^d} u(x,0)m_0(x)\,\mathrm{d}x + \int_0^T \Phi^*\left(P(t)\right)\mathrm{d}t + \iint\limits_Q F^*\left(x,\gamma(x,t)\right)\mathrm{d}x\,\mathrm{d}t. \tag{2.12}$$

The first problem is

$$\inf_{(u,P,\gamma)\in\mathcal{K}_0} D(u,P,\gamma). \tag{2.13}$$

We consider now the set $\mathcal{K}_1$ of all pairs $(m, w) \in L^1(Q) \times L^1(Q; \mathbb{R}^d)$ such that $m \geqslant 0$ a.e., $\int_{\mathbb{T}^d} m(x,t)\,\mathrm{d}x = 1$ for a.e. $t \in (0,T)$, and such that the continuity equation

$$\begin{cases} \partial_t m - \partial_{ij}\left(A_{ij} m\right) + \nabla \cdot w &= 0, \\ m(x,0) &= m_0(x) \end{cases} \tag{2.14}$$

holds in the sense of distributions. For $(m, w) \in \mathcal{K}_1$, consider the following cost functional:

$$\begin{aligned} B(m,w) = & \iint_Q \left(H^*\left(x, -\frac{w(x,t)}{m(x,t)}\right) m(x,t) + F\left(x, m(x,t)\right)\right) \mathrm{d}x\,\mathrm{d}t \\ & + \int_0^T \Phi\left(\int_{\mathbb{T}^d} G(x) w(x,t)\,\mathrm{d}x\right)\mathrm{d}t + \int_{\mathbb{T}^d} u_T(x) m(x,T)\,\mathrm{d}x, \end{aligned} \tag{2.15}$$

where for $m(t,x) = 0$, we impose that

$$m(t,x) H^*\left(x, -\frac{w(t,x)}{m(t,x)}\right) = \begin{cases} +\infty & \text{if } w(t,x) \neq 0 \\ 0 & \text{if } w(t,x) = 0. \end{cases}$$

Since H^* and F are bounded from below and $m \geqslant 0$, the first integral in B is well defined in $\mathbb{R} \cup \{+\infty\}$. Concerning the second term in B, we simply need to observe that since Φ is convex, there exists a constant $C > 0$ such that $\Phi(z) \geqslant -C(1 + |z|)$, for all $z \in \mathbb{R}^k$. For $w \in L^1(Q; \mathbb{R}^d)$, the term

$$\int_{\mathbb{T}^d} G(x) w(x, \cdot)\,\mathrm{d}x$$

is integrable in time and therefore

$$\int_0^T \Phi\left(\int_{\mathbb{T}^d} G(x) w(x,t)\,\mathrm{d}x\right)\mathrm{d}t$$

is well-defined in $\mathbb{R} \cup \{+\infty\}$. For the third term, we refer the reader to [11, Lemma 4.1], where it is proved that for $(m, w) \in \mathcal{K}_1$, m can be seen as a continuous map from $[0,T]$ to $P(\mathbb{T}^d)$ for the Rubinstein-Kantorovich distance $\mathbf{d}_1$. Finally, the second optimization problem is the following:

$$\inf_{(m,w)\in\mathcal{K}_1} B(m,w). \tag{2.16}$$

2.5.1 Lemma. *We have*

$$\inf_{(u,P,\gamma)\in\mathcal{K}_0} D(u,P,\gamma) = -\min_{(m,w)\in\mathcal{K}_1} B(m,w).$$

Moreover, the minimum in the right-hand side is achieved by a unique pair $(m,w)\in\mathcal{K}_1$ *satisfying*

$$m \in L^q(Q), \quad w \in L^{\frac{r'q}{r'+q-1}}(Q;\mathbb{R}^d), \quad \textit{and} \quad \int_{\mathbb{T}^d} G(x)w(x,\cdot)\,\mathrm{d}x \in L^s((0,T);\mathbb{R}^k). \tag{2.17}$$

Proof. Following previous papers [9, 10, 11], we look to apply the Fenchel-Rockafellar duality theorem. In order to do so, we reformulate the first optimization problem into a more suitable form.

Let $E_0 = \mathcal{C}^2(Q)\times\mathcal{C}^0([0,T],\mathbb{R}^k)$ and $E_1 = \mathcal{C}^0(Q)\times\mathcal{C}^0(Q;\mathbb{R}^d)$. Define on E_0 the functional

$$\mathcal{F}(u,P) = \int_0^T \Phi^*(P(t))\,\mathrm{d}t - \int_{\mathbb{T}^d} u(0,x)m_0(x)\ \mathrm{d}x + \chi_S(u),$$

where χ_S is the convex characteristic function of the set $S = \left\{u\in E_0, u(T,\cdot) = u_T\right\}$, i.e., $\chi_S(u) = 0$ if $u\in S$ and $+\infty$ otherwise. For $(a,b)\in E_1$, we define

$$\mathcal{G}(a,b) = \iint_Q F^*\left(x, -a + H(x,b)\right)\mathrm{d}x\,\mathrm{d}t.$$

The functional $\mathcal{F}$ is convex and lower semi-continuous on E_0 while $\mathcal{G}$ is convex and continuous on E_1. Let $\Lambda : E_0 \to E_1$ be the bounded linear operator defined by

$\Lambda(u,P) = (\partial_t u + A_{ij}\partial_{ij}u, Du + G^\intercal P)$. We can observe that

$$\inf_{(u,P,\gamma)\in\mathcal{K}_0} D(u,P,\gamma) = \inf_{(u,P)\in E_0} \left\{\mathcal{F}(u,P) + \mathcal{G}(\Lambda(u,P))\right\}.$$

In the interest of employing the Fenchel-Rockafellar duality theorem, note that $\mathcal{F}(u,P) < +\infty$ for $(u_T, 0)$ and $\mathcal{G}(\Lambda(u,P))$ is continuous at $(u_T, 0)$. This satisfies the duality theorem, and so

$$\inf_{(u,P)\in E_0} \left\{\mathcal{F}(u,P) + \mathcal{G}(\Lambda(u,P))\right\} = \max_{(m,w)\in E_1'} \left\{-\mathcal{F}^*(\Lambda^*(m,w)) - \mathcal{G}^*(-(m,w))\right\},$$

where E_1' is the dual space of E_1, i.e., the set of vector valued Radon measures (m,w) over Q with values in $\mathbb{R}\times\mathbb{R}^d$, E_0' is the dual space of E_0, $\Lambda^* : E_1' \to E_0'$ is the dual operator of Λ and $\mathcal{F}^*$ and $\mathcal{G}^*$ are the convex conjugates of $\mathcal{F}$ and $\mathcal{G}$ respectively. We now compute the relevant conjugate transforms.

$$\begin{aligned}
&\mathcal{F}^*(\Lambda^*(m,w)) = \sup_{(u,P)} \Big\{\langle (u,P), \Lambda^*(m,w)\rangle - \mathcal{F}(u,P)\Big\} \\
&= \sup_{\substack{(u,P)\\ u\in S}} \Big\{\langle \Lambda(u,P), (m,w)\rangle - \int_0^T \Phi^*(P)\,\mathrm{d}t + \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x\Big\} \\
&= \sup_{\substack{(u,P)\\ u\in S}} \Big\{\langle \partial_t u + A_{ij}\partial_{ij}u, m\rangle + \langle Du, w\rangle + \langle G^\intercal P, w\rangle \\
&- \int_0^T \Phi^*(P)\,\mathrm{d}t + \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x\Big\} \\
&= \sup_{\substack{(u,P)\\ u\in S}} \Big\{\langle -\partial_t m + \partial_{ij}(A_{ij}m) - \nabla\cdot w, u\rangle + \langle G^\intercal P, w\rangle \\
&- \int_0^T \Phi^*(P)\,\mathrm{d}t + \int_{\mathbb{T}^d} u_T(x)m(T,x)\,\mathrm{d}x\Big\}
\end{aligned}$$

It is evident here that if $-\partial_t m + \partial_{ij}(A_{ij}m) - \nabla\cdot w \neq 0$ in the sense of distributions, this supremum is infinite. If the condition does hold however, the supremum no longer depends on u, and so the calculation is reduced to

$$\begin{aligned}
&\sup_P \left\{ \int_0^T \Big(\textstyle\int_{\mathbb{T}^d} \langle G(x)^\intercal P(t), w(x,t) \rangle \,\mathrm{d}x \Big) - \Phi^*(P(t)) \,\mathrm{d}t \right\} \\
&\qquad = \sup_P \left\{ \int_0^T \Big\langle P(t), \textstyle\int_{\mathbb{T}^d} G(x) w(x,t) \,\mathrm{d}x \Big\rangle - \Phi^*(P(t)) \,\mathrm{d}t \right\} \\
&\qquad = \sup_P \left\{ \int_0^T \Phi\Big(\textstyle\int_{\mathbb{T}^d} G(x) w(x,t) \,\mathrm{d}x \Big) \,\mathrm{d}t \right\}.
\end{aligned}$$

Combined with the conditions above, we have

$$\mathcal{F}^*(\Lambda^*(m,w)) =$$
$$\begin{cases} \displaystyle\int_0^T \Phi\Big(\textstyle\int_{\mathbb{T}^d} G w \Big) + \displaystyle\int_{\mathbb{T}^d} u_T(x) m(T,x) \,\mathrm{d}x, & \text{if } -\partial_t m + A_{ij}\partial_{ij} m - \nabla \cdot w = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Following [9], we have that $\mathcal{G}^*(m,w) = +\infty$ if $(m,w) \notin L^1(Q) \times L^1(Q;\mathbb{R}^d)$ and

$$\mathcal{G}^*(m,w) = \iint_Q K^*(x, m(t,x), w(t,x)) \,\mathrm{d}x \,\mathrm{d}t$$

otherwise, where K^* is given by

$$K^*(x,m,w) = \begin{cases} F(x,-m) - m H^*\Big(x, -\frac{w}{m}\Big) & \text{if } m < 0 \\ 0 & \text{if } m = 0 \text{ and } w = 0 \\ +\infty & \text{otherwise,} \end{cases}$$

it is the convex conjugate of

$$K(x,a,b) = F^*(x, -a + H(x,b)) \qquad \forall (x,a,b) \in \mathbb{T}^d \times \mathbb{R} \times \mathbb{R}^d.$$

Therefore

$$
\begin{aligned}
&\max_{(m,w)\in E_1'} \big\{ -\mathcal{F}^*(\Lambda^*(m,w)) - \mathcal{G}^*(-(m,w))\big\} \\
&= \max\Big\{ \iint\limits_Q -F(x,m) - mH^*\Big(x, -\frac{w}{m}\Big)\,\mathrm{d}x\,\mathrm{d}t - \int_0^T \Phi\left(\textstyle\int_{\mathbb{T}^d} Gw\right)\mathrm{d}t \\
&- \int_{\mathbb{T}^d} u_T(x) m(T,x)\,\mathrm{d}x\Big\}
\end{aligned}
$$

with the last maximum taken over the L^1 maps (m,w) such that $m \geqslant 0$ a.e. and

$$
-\partial_t m + \partial_{ij}(A_{ij}m) - \nabla\cdot w = 0, \quad m(0) = m_0
$$

holds in the sense of distributions. Since $\int_{\mathbb{T}^d} m_0 = 1$, it follows that $\int_{\mathbb{T}^d} m(t) = 1$ for any $t \in [0,T]$. Thus the pair (m,w) belongs to the set $\mathcal{K}_1$ and the first part of the statement is proved.

It remains to show (2.17). Taking an optimal $(m,w) \in \mathcal{K}_1$ in the above problem, we have that $w(t,x) = 0$ for all $(t,x) \in [0,T]\times\mathbb{T}^d$ whenever $m(t,x) = 0$. By convexity of Φ, we have

$$
\int_0^T \Phi\big(\textstyle\int_{\mathbb{T}^d} Gw\,\mathrm{d}x\big)\,\mathrm{d}t \geqslant \displaystyle\int_0^T \Phi(0) + \big\langle \Psi(0), \textstyle\int_{\mathbb{T}^d} Gw\,\mathrm{d}x\big\rangle\,\mathrm{d}t \geqslant C - C\|w\|_1. \tag{2.18}
$$

Moreover, by Hölder's inequality,

$$
\|w\|_1 = \iint\limits_Q \frac{|w|}{m} m\,\mathrm{d}x\,\mathrm{d}t \leqslant \Big(\iint\limits_Q \Big(\frac{|w|}{m}\Big)^{r'} m\Big)^{1/r'}.
$$

It follows with Young's inequality that

$$
\|w\|_1 \leqslant \frac{\varepsilon}{r'} \iint\limits_Q |w|^{r'} m^{1-r'}\,\mathrm{d}x\,\mathrm{d}t + \varepsilon^{-(r-1)},
$$

for any $\varepsilon > 0$. Therefore,

$$\int_0^T \Phi\big(\textstyle\int_{\mathbb{T}^d} Gw \,\mathrm{d}x\big)\,\mathrm{d}t \geqslant C - \dfrac{\varepsilon C}{r'} \displaystyle\iint_Q |w|^{r'} m^{1-r'} \,\mathrm{d}x \,\mathrm{d}t - \varepsilon^{-(r-1)}.$$

Using successively the optimality of (m, w), the growth conditions on F and H^* and the above inequality, we obtain

$$\begin{aligned} C &\geqslant \iint_Q F(x,m) + mH^*\Big(x, -\frac{w}{m}\Big)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \Phi\big(\textstyle\int_{\mathbb{T}^d} Gw\big)\,\mathrm{d}t + \displaystyle\int_{\mathbb{T}^d} u_T(x) m(T,x)\,\mathrm{d}x \\ &\geqslant \frac{1}{C}\|m\|_q^q + \Big(\frac{1}{C} - \varepsilon C\Big) \iint_Q |w|^{r'} m^{1-r'}\,\mathrm{d}x\,\mathrm{d}t - \varepsilon^{-(r-1)} - C, \end{aligned}$$

for some constant $C > 0$ independent of ε. Choosing $\varepsilon > 0$ sufficiently small, we deduce that $m \in L^q(Q)$ and that $|w|^{r'} m^{1-r'} \in L^1(Q)$. To investigate the claim on w given in the statement, write, for some ρ to be determined,

$$\|w\|_{\rho r'}^{r'} = \left\| m^{r'-1} \frac{|w|^{r'}}{m^{r'-1}} \right\|_\rho \leqslant \left\| m^{r'-1} \right\|_{\frac{q}{r'-1}} \left\| \frac{|w|^{r'}}{m^{r'-1}} \right\|_1 = \|m\|_q^{r'-1} \left\| \frac{|w|^{r'}}{m^{r'-1}} \right\|_1 < \infty.$$

For this interpolative Hölder inequality to hold, we must have $\rho = \frac{q}{r'+q-1}$. Thus $w \in L^\sigma(Q;\mathbb{R}^d)$, with

$$\sigma = \frac{r'q}{r'+q-1}, \quad \text{i.e. } \sigma' = rp. \tag{2.19}$$

Two cases must be considered. If $\frac{1}{s} \geqslant 1 - \frac{1}{rp} = \frac{1}{\sigma}$, then we have $\sigma \geqslant s$, thus $w \in L^s(Q;\mathbb{R}^d)$ and (2.17) follows. In the other case, we have by Hypothesis (H1) the growth assumption $\Phi(z) \geqslant \frac{1}{C}|z|^s - C$. It can be employed to get a better bound from below in (2.18). We obtain (2.17) with a straightforward adaptation of the above proof.

$\square$

In general we do not expect problem (2.13) to have a solution. In this section we exhibit a relaxation for (2.13) (Proposition 2.6.8) and show that the obtained relaxed problem has at least one solution (Proposition 2.6.10).

2.6.1 Estimates on Subsolutions to HJ equations

In this subsection we prove estimates in Lebesgue spaces for subsolutions of Hamilton-Jacobi equations of the form

$$\begin{cases} (i) \quad -\partial_t u - A_{ij}\partial_{ij} u + H(Du + G^\intercal P) \leqslant \gamma \\ (ii) \quad u(x,T) \leqslant u_T(x) \end{cases} \tag{2.20}$$

in terms of Lebesgue norms of γ, u_T, and P. Equation (2.20) is understood in the sense of distributions. This means that $Du + G^\intercal P \in L^r(Q)$ and for any nonegative test function $\zeta \in C_c^\infty((0,T] \times \mathbb{T}^d)$,

$$-\int_{\mathbb{T}^d} \zeta(T) u_T + \iint_Q u\partial_t\zeta + \langle D\zeta, ADu\rangle + \zeta\big(\partial_i A_{ij}\partial_j u + H(Du + G^\intercal P)\big) \leqslant \iint_Q \gamma\zeta. \tag{2.21}$$

Let us introduce some notation. For all $\tilde{r} > 1$ and for all $\tilde{p} \geqslant 1$, let us define $\bar{\kappa}(\tilde{r},\tilde{p})$ and $\bar{\eta}(\tilde{r},\tilde{p})$ by

$$\bar{\eta}(\tilde{r},\tilde{p}) = \frac{d(\tilde{r}(\tilde{p}-1)+1)}{d - \tilde{r}(\tilde{p}-1)} \quad \text{and} \quad \bar{\kappa}(\tilde{r},\tilde{p}) = \frac{\tilde{r}\tilde{p}(1+d)}{d - \tilde{r}(\tilde{p}-1)}$$

if $\tilde{p} < 1 + \frac{d}{\tilde{r}}$ and

$$\bar{\eta}(\tilde{r},\tilde{p}) = \infty \quad \text{and} \quad \bar{\kappa}(\tilde{r},\tilde{p}) = \infty$$

if $\tilde{p} > 1 + \frac{d}{\tilde{r}}$. In the border line case $\tilde{p} = 1 + \frac{d}{\tilde{r}}$, $\bar{\eta}(\tilde{r},\tilde{p})$ and $\bar{\kappa}(\tilde{r},\tilde{p})$ can be fixed to arbitrarily large values. We let the reader verify that

$$\bar{\kappa}(\tilde{r},\tilde{p}) \geqslant \tilde{r} \quad \text{and} \quad \bar{\kappa}(\tilde{r},\tilde{p}) \geqslant \tilde{p}, \tag{2.22}$$

assuming that the assigned value to $\bar{\kappa}(\tilde{r},\tilde{p})$ is large enough in the border line case.

We now restate [11, Theorem 3.3], since it will prove useful below.

2.6.1 Theorem. *Let u satisfy*

$$\begin{cases} (i) \quad -\partial_t u - A_{ij}\partial_{ij}u + \frac{1}{K}|Du|^{\tilde{r}} \leqslant \gamma \\ \\ (ii) \quad u(x,T) \leqslant u_T(x) \end{cases} \tag{2.23}$$

in the sense of distributions, with $\gamma \in L^{\tilde{p}}(Q)$ for some $\tilde{p} \geqslant 1$ and $\tilde{r} > 1$. Then

$$\|u_+\|_{L^\infty((0,T),L^\eta(\mathbb{T}))} + \|u_+\|_{L^\kappa(Q)} \leqslant C,$$

where $u_+ := \max\{u,0\}$, $\kappa = \bar{\kappa}(\tilde{r},\tilde{p})$, $\eta = \bar{\eta}(\tilde{r},\tilde{p})$. The constant C depends only on $T, d, \tilde{r}, \tilde{p}$, K (appearing in (2.23)*), C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_{\tilde{p}}$, and $\|u_T\|_\eta$.*

2.6.2 Remark. *Although the case $\tilde{p} = 1$ is not explicitly mentioned in [11, Theorem 3.3], it is not hard to check that the theorem also applies in that case.*

2.6.3 Corollary. *Let $P \in \mathcal{C}^0([0,T];\mathbb{R}^k)$ and $\gamma \in \mathcal{C}^0(Q;\mathbb{R})$. Let u be the viscosity solution to the HJ equation*

$$\begin{cases} -\partial_t u - A_{ij}\partial_{ij}u + H\left(x, Du(x,t) + G(x,t)^\intercal P(t)\right) &= \gamma, \\ u(x,T) &= u_T(x). \end{cases} \tag{2.24}$$

Let $\tilde{r} > 1$. Define

$$\kappa = \bar{\kappa}(\tilde{r},1) = \frac{\tilde{r}(1+d)}{d}. \tag{2.25}$$

Then

$$\|u\|_{L^\infty((0,T),L^1(\mathbb{T}))} + \|u\|_{L^\kappa(Q)} \leqslant C,$$

where the constant C depends only on $T, d, \tilde{r}, \tilde{p}$, C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_1, \|Du\|_{\tilde{r}}$, $\|H(Du + G^\intercal P)\|_1$, and $\|u_T\|_1$.

Proof. By [34], u also satisfies the HJ equation in the sense of distributions. Observe that (2.24) can be rewritten

$$-\partial_t u - A_{ij}\partial_{ij}u + |Du|^{\tilde{r}} = \gamma - H(Du + G^\intercal P) + |Du|^{\tilde{r}}. \tag{2.26}$$

The L^1 norm of the right-hand side depends on $\|\gamma\|_1$, $\|H(Du + G^\intercal P)\|_1$, and $\|Du\|_{\tilde{r}}$. Similarly, $-u$ is a weak subsolution of a HJ equation with right-hand side

$$-\gamma + H(Du + G^\intercal P) + |Du|^{\tilde{r}}.$$

By applying Theorem 2.6.1 to both u and $-u$, we deduce the desired estimate.

□

When $s' \geqslant r$, the growth assumption on the Hamiltonian (Hypothesis (H2)) can be exploited to derive a more precise estimate on the solution to (2.24).

2.6.4 Corollary. *Let P, u, and γ be as in Corollary 2.6.3. Assume moreover that $\gamma \geqslant 0$ and $s' \geqslant r$. Take $\tilde{r} \in (1, r]$ and define*

$$\tilde{p} = \min\left(p, \frac{s'}{\tilde{r}}\right), \quad \kappa = \bar{\kappa}(\tilde{r}, \tilde{p}), \quad \text{and} \quad \eta = \bar{\eta}(\tilde{r}, \tilde{p}).$$

Then

$$\|u\|_{L^\infty((0,T),L^\eta(\mathbb{T}^d))} + \|u\|_{L^\kappa(Q)} \leqslant C.$$

The constant C depends only on $T, d, \tilde{r}, \tilde{p}$, C_3 (appearing in Hypothesis (H2)), C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_p, \|P\|_{s'}$, and $\|u_T\|_\eta$.

Proof. We have $\gamma \geqslant 0$ and the upper bound

$$H(Du + G^{\intercal}P) \leqslant C|Du|^r + C|P|^r,$$

therefore,

$$-\partial_t u - A_{ij}\partial_{ij}u + C|Du|^r \geqslant -C|P|^r - C, \quad u(T,x) \geqslant \Big(\min_{x' \in \mathbb{T}^d} u_T(x')\Big). \tag{2.27}$$

Let $\hat{u}$ be defined by

$$\hat{u}(x,t) = \Big(\min_{x' \in \mathbb{T}^d} u_T(x')\Big) - \int_t^T |P(t)|^r \,\mathrm{d}t - C(T-t),$$

which solves (2.27) with inequality replaced by equality. By the comparison principle, $u \geqslant \hat{u} \geqslant -C$, where C depends only on T, the growth of H, $\min u_T$, and $\|P\|_r$. Note that $\|P\|_r$ depends only on $\|P\|_{s'}$ and T because $s' \geqslant r$.

Next, by the growth condition of H we have

$$-\partial_t u - A_{ij}\partial_{ij}u + \frac{1}{C_3}|Du + G^{\intercal}P|^r - C_3 \leqslant \gamma.$$

Observe that by Young's inequality,

$$|Du|^{\tilde{r}} \leqslant 2^{\tilde{r}-1}|Du + G^{\intercal}P|^{\tilde{r}} + 2^{\tilde{r}-1}|G^{\intercal}P|^{\tilde{r}} \leqslant 2^{\tilde{r}-1}\Big(\frac{\tilde{r}}{r}|Du + G^{\intercal}P|^r + 1\Big) + C|P|^{\tilde{r}},$$

since $r \geqslant \tilde{r}$. It follows that

$$|Du + G^{\intercal}P|^r \geqslant \frac{1}{C}|Du|^{\tilde{r}} - C|P|^{\tilde{r}} - C,$$

therefore,

$$-\partial_t u - A_{ij}\partial_{ij}u + \frac{1}{C}|Du|^{\tilde{r}} \leqslant \gamma + C|P|^{\tilde{r}} + C. \tag{2.28}$$

Since $|P|^{\tilde{r}}$ lies in $L^{s'/\tilde{r}}(Q)$, we have that the right hand side of (2.28) is bounded in $L^{\tilde{p}}$. Combining this with the lower bound on u, the conclusion follows from Theorem 2.6.1.

□

We can now fix the values of the coefficients $\tilde{r} \in (1, r]$, $\kappa > 1$, and $\eta > 1$ to be employed in the sequel, consistently with Corollary 2.6.3 (if $s' < r$) and Corollary 2.6.4 (if $s' \geqslant r$). As will appear later in the proofs of Lemma 2.6.9 and Proposition 2.6.10 , these coefficients must satisfy the following:

$$\left[s' \geqslant \tilde{r}\right], \quad \left[\kappa \geqslant p\right], \quad \text{and} \quad \left[A \text{ is not constant} \Longrightarrow \tilde{r} \geqslant p\right].$$

This is the reason why four subcases have been introduced in Hypothesis (H5) and why we have a specific definition of the coefficients for each of the subcases. In order to deal with the case 2B, we need the following lemma.

2.6.5 Lemma. *Assume that $s' \geqslant r$ and A is constant, that is, consider the case 2B of Assumption (H5). Then the corresponding condition:*

$$\left[s' \geqslant 1 + d\right] \quad \textit{or} \quad \left[s' < 1 + d \quad \textit{and} \quad \frac{s'(1+d)}{d - s' + 1} > p\right] \tag{2.29}$$

is satisfied if and only if there exists $\tilde{r} \in (1, r]$ such that

$$\bar{\kappa}\Big(\tilde{r}, \min\left(p, \tfrac{s'}{\tilde{r}}\right)\Big) \geqslant p. \tag{2.30}$$

Proof. Several cases must be distinguished.

- Case (i): $s' > p$. In that case, either $s' \geqslant 1 + d$ or $s' < 1 + d$ and then

$$\frac{s'(1+d)}{d - s' + 1} > s' > p.$$

Thus, if $s' > p$, then (2.29) holds true. Then we can set $\tilde{r} = \frac{s'}{p} > 1$. We have $\min\left(p, \frac{s'}{\tilde{r}}\right) = p$ and therefore $\kappa = \bar{\kappa}(\tilde{r}, p) \geqslant p$, by inequality (2.22).

- Case (ii): $s' \leqslant p$. Then whatever the choice of $\tilde{r} \in (1, r]$, we have

$$\tilde{p}(\tilde{r}) := \min\left(p, \tfrac{s'}{\tilde{r}}\right) = \tfrac{s'}{\tilde{r}}.$$

  - Case (iia): $s' > 1 + d$. Then we can chose $\tilde{r}$ sufficiently close to 1, so that

$$\frac{s' - d}{\tilde{r}} > 1.$$

    Then we have $\tilde{p}(\tilde{r}) > 1 + \frac{d}{\tilde{r}}$, thus $\kappa = \bar{\kappa}(\tilde{r}, \tilde{p}(\tilde{r})) = \infty$.

  - Case (iib): $s' \leqslant 1 + d$. Then whatever the choice of $\tilde{r} \in (1, r]$, we have $\tilde{p}(\tilde{r}) < 1 + \frac{d}{\tilde{r}}$ and therefore, condition (2.30) is equivalent to:

$$\exists \tilde{r} \in (1, r], \quad \frac{s'(d+1)}{d + \tilde{r} - s'} \geqslant p.$$

    The above condition is clearly satisfied if and only if either $s' = 1 + d$ or $s' < 1 + d$ and $\frac{s'(d+1)}{d+1-s'} > p$.

$\square$

We can finally fix $\tilde{r}$, κ, and η.

- In cases 1A and 1B (i.e. $s' < r$), we set

$$\tilde{r} = s', \quad \kappa = \bar{\kappa}(\tilde{r}, 1), \quad \eta = \bar{\eta}(\tilde{r}, 1).$$

  Then we have $\kappa \geqslant \tilde{r} \geqslant p$. In case 1A, $\tilde{r} \geqslant p$.

- In case 2A (i.e. $s' \geqslant r$ and A is not constant), we set $\tilde{r} = p$. In case 2B (i.e. $s' \geqslant r$ and A is constant), we assign a value to $\tilde{r}$ so that (2.30) holds true. In both cases 2A and 2B, we set

$$\kappa = \bar{\kappa}\Big(\tilde{r}, \min\left(p, \tfrac{s'}{\tilde{r}}\right)\Big) \quad \text{and} \quad \eta = \bar{\eta}\Big(\tilde{r}, \min\left(p, \tfrac{s'}{\tilde{r}}\right)\Big).$$

In case 2A, we have $\kappa \geqslant \tilde{r} = p$ by inequality (2.22). In case 2B, we have $\kappa \geqslant p$ by definition.

2.6.6 Remark. *In case 2B, it is easy to deduce from the proof of Lemma 2.6.5 an explicit* $\tilde{r} \in (1, r]$ *such that* (2.30) *holds. Note that the obtained* $\tilde{r}$ *may not be the best one (i.e. the largest one). For example, if* $s' \geqslant pr$*, then one can take* $\tilde{r} = r$*. Then* $\frac{s'}{\tilde{r}} \geqslant p$ *and therefore* $\kappa = \bar{\kappa}(\tilde{r}, p) \geqslant p$*, by inequality* (2.22).

2.6.2 The Relaxed Problem

We propose in this subsection an appropriate relaxation of problem (2.13). Let $\mathcal{K}$ denote the set of triplets $(u, P, \gamma) \in L^{\kappa}(Q) \times L^{s'}(0,T) \times L^{p}(Q)$ such that $Du + G^{\intercal} P \in L^{r}(Q; \mathbb{R}^d)$, $Du \in L^{\tilde{r}}(Q; \mathbb{R}^d)$, and such that (2.20) holds in the sense of distributions. The following statement explains that u has a "trace" in a weak sense.

2.6.7 Lemma. *Let* $f \in L^1(Q)$ *and let* $u \in L^1(Q)$ *satisfy* $Du \in L^{\tilde{r}}(Q; \mathbb{R}^d)$ *and*

$$- \partial_t u - A_{ij} \partial_{ij} u \leqslant f, \quad u(T) \leqslant u_T \tag{2.31}$$

in the sense of distributions, i.e. for every non-negative function $\vartheta \in C_c^{\infty}(\mathbb{T}^d \times (0,T])$ *we have*

$$\iint\limits_{Q} \big(u \partial_t \vartheta + \partial_j (A_{ij} \vartheta) \partial_i u - f \vartheta \big) \, \mathrm{d}x \, \mathrm{d}t \leqslant \int_{\mathbb{T}^d} \vartheta(x,T) u_T(x) \, \mathrm{d}x.$$

Then, for any C^1 *map* $\vartheta \colon [0,T] \times \mathbb{T}^d \to \mathbb{R}$*, the function*

$$t \mapsto \int_{\mathbb{T}} \vartheta(x,t) u(x,t) dx$$

has a BV representative on $[0,T]$. *In particular, for any nonnegative* C^1 *map* $\vartheta\colon [0,T]\times\mathbb{T}^d \to \mathbb{R}$, *one has the integration by parts formula: for any* $0 \leqslant t_1 \leqslant t_2 \leqslant T$,

$$-\Big[\int_{\mathbb{T}^d} \vartheta u\Big]_{t_1}^{t_2} + \int_{t_1}^{t_2}\int_{\mathbb{T}^d} u\partial_t\vartheta + \langle D\vartheta, ADu\rangle + \vartheta\partial_i A_{ij}\partial_j u \leqslant \int_{t_1}^{t_2}\int_{\mathbb{T}^d} f\vartheta. \tag{2.32}$$

Proof. First, observe that $x \mapsto u(x,t)$ is a well-defined $L^1(\mathbb{T}^d)$ function for a.e. $t \in (0,T)$. Then by standard convolution smoothing arguments, one can check that (2.32) holds for a.e. $t_1, t_2 \in [0,T]$ with $t_1 \leqslant t_2$. Indeed, if ξ_ε is a convolution kernel, then $u_\varepsilon = \xi_\varepsilon * u$, $f_\varepsilon = \xi_\varepsilon * f$ can be shown to satisfy

$$-\partial_t u_\varepsilon - A_{ij}\partial_{ij}u_\varepsilon \leqslant f_\varepsilon + R_\varepsilon \tag{2.33}$$

where $R_\varepsilon \to 0$ in L^1 as $\varepsilon \to 0$. Then integration by parts implies, for $0 < t_1 \leqslant t_2 < T$ and ε small enough, that

$$-\Big[\int_{\mathbb{T}^d} \vartheta u_\varepsilon\Big]_{t_1}^{t_2} + \int_{t_1}^{t_2}\int_{\mathbb{T}^d} u_\varepsilon\partial_t\vartheta + \langle D\vartheta, ADu_\varepsilon\rangle + \vartheta\partial_i A_{ij}\partial_j u_\varepsilon \leqslant \int_{t_1}^{t_2}\int_{\mathbb{T}^d} (f_\varepsilon + R_\varepsilon)\vartheta. \tag{2.34}$$

Since $u_\varepsilon(\cdot,t) \to u(\cdot,t)$ in $L^1(\mathbb{T}^d)$ for a.e. t, and likewise $u_\varepsilon \to u, Du_\varepsilon \to Du$, and $f_\varepsilon \to f$ in L^1, so by letting $\varepsilon \to 0$ we deduce the (2.32) for a.e. $t_1, t_2 \in [0,T]$ with $t_1 \leqslant t_2$.

Now define, for a.e. $t \in [0,T]$, the functions

$$G(t) = \int_{\mathbb{T}^d} \vartheta(x,t)u(x,t)\,\mathrm{d}x + F(t), \quad F(t) = \int_t^T\int_{\mathbb{T}^d} u\partial_t\vartheta + \langle D\vartheta, ADu\rangle + \vartheta\partial_i A_{ij}\partial_j u - f\vartheta.$$

Now F is absolutely continuous, being the integral of an $L^1(0,T)$ function. By what we have shown $G(t)$ is increasing on its domain, and moreover $G(T) \leqslant \int_{\mathbb{T}^d} \vartheta(x,T)u_T(x)\,\mathrm{d}x$ by hypothesis. Thus $I := G - F$ is BV, and (2.32) indeed continues to hold for all $0 \leqslant t_1 \leqslant t_2 \leqslant T$, even if we replace $\int_{\mathbb{T}^d} \vartheta(x,t)u(x,t)\,\mathrm{d}x$ by any value between $I(t+)$ and $I(t-)$.

$\square$

We extend the functional D to triplets $(u, P, \gamma) \in \mathcal{K}$:

$$D(u, P, \gamma) = -\int_{\mathbb{T}^d} u(x, 0^+) m_0(x)\,\mathrm{d}x + \int_0^T \Phi^*\big(P(t)\big)\,\mathrm{d}t + \iint\limits_Q F^*\big(x, \gamma(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t.$$

We consider the following relaxation of problem (2.13):

$$\inf_{(u,P,\gamma)\in\mathcal{K}} D(u, P, \gamma). \tag{2.35}$$

2.6.8 Proposition. *We have*

$$\inf_{(u,P,\gamma)\in\mathcal{K}_0} D(u, P, \gamma) = \inf_{(u,P,\gamma)\in\mathcal{K}} D(u, P, \gamma).$$

The proof requires an integration by parts formula, established in the following lemma.

2.6.9 Lemma. *Let $(u, P, \gamma) \in \mathcal{K}$ and $(m, w) \in \mathcal{K}_1$ satisfy* (2.17)*.*

Assume that $mH^*(\cdot, -w/m) \in L^1(Q)$. Then

$$\gamma m \in L^1((0,T) \times \mathbb{T}^d), \quad \langle P(\cdot), \textstyle\int_{\mathbb{T}^d} G(x) w(x, \cdot)\,\mathrm{d}x\rangle \in L^1(0,T)$$

and for almost all $t \in (0,T)$ we have

$$\begin{aligned}&\int_{\mathbb{T}^d} (u(T)m_T - u(t)m(t))\,\mathrm{d}x + \\ &\qquad \int_t^T \int_{\mathbb{T}^d} \Big(m\gamma + mH^*\big(x, -\tfrac{w}{m}\big) + \langle P(t), G(x)w(x,t)\rangle\Big)\,\mathrm{d}x\,\mathrm{d}t \;\geqslant\; 0,\end{aligned} \tag{2.36}$$

and

$$\begin{aligned}&\int_{\mathbb{T}^d} (u(t)m(t) - u(0)m_0)\,\mathrm{d}x + \\ &\qquad \int_0^t \int_{\mathbb{T}^d} \Big(m\gamma + mH^*\left(x, -\tfrac{w}{m}\right) + \langle P(t), G(x,t)w(x,t)\rangle\Big)\,\mathrm{d}x\,\mathrm{d}t \;\geqslant\; 0.\end{aligned} \tag{2.37}$$

Moreover, if equality holds in the inequality (2.36) *for $t = 0$, then*

$$w = -mD_\xi H(\cdot, Du + G^\intercal P) \text{ a.e.}$$

Proof. In the interest of smoothing (m, w) by convolution, extend the pair to $[-1, T+1] \times \mathbb{T}^d$ by defining $m = m_0$ on $[-1, 0]$, $m = m(T)$ on $[T, T+1]$, and $w(s, x) = 0$ for $(s, x) \in (-1, 0) \cup (T, T+1) \times \mathbb{T}^d$. Let $\tilde{A}_{ij}$ be an extension of A_{ij} with $\tilde{A}_{ij} = A_{ij}$ if $t \in (0, T)$ and zero otherwise. Note that with these described extensions, (m, w) solves

$$\partial_t m - \partial_{ij}(\tilde{A}_{ij}(t, x)m) + \nabla \cdot w = 0 \quad \text{ on } (-1, T+1) \times \mathbb{T}^d.$$

Let $\xi^\epsilon = \xi^\epsilon(t, x) = \xi_1^\epsilon(t)\xi_2^\epsilon(x)$ be a smooth convolution kernel with support in a ball of radius ϵ. We smoothen the pair (m, w) with this kernel in a standard way into $(m_\epsilon, w_\epsilon) = (\xi^\epsilon * m, \xi^\epsilon * w)$. Then (m_ϵ, w_ϵ) solves

$$\partial_t m_\epsilon - \partial_{ij}(\tilde{A}_{ij} m_\epsilon) + \nabla \cdot w_\epsilon = \partial_j R_\epsilon \quad \text{ in } \left(-\frac{1}{2}, T + \frac{1}{2}\right) \tag{2.38}$$

in the sense of distributions, where

$$R_\epsilon := [\xi^\epsilon, \partial_j \tilde{A}_{ij}](m) + [\xi^\epsilon, \tilde{A}_{ij}\partial_j](m). \tag{2.39}$$

Here we use again the commutator notation [20]

$$[\xi^\epsilon, c](f) := \xi^\epsilon \star (cf) - c(\xi^\epsilon \star f). \tag{2.40}$$

By [20, Lemma II.1], we have that $R_\epsilon \to 0$ in L^q, since $m \in L^q$ and $\tilde{A}_{ij} \in W^{1,\infty}$. Fix time $t \in (0, T)$ at which $u(t^+) = u(t^-) = u(t)$ in $L^\kappa(\mathbb{T}^d)$ and $m_\epsilon(t)$ converges to $m(t)$. We have the following inequality based on the equality in (2.11),

$$-\partial_t u - A_{ij}\partial_{ij} u + H(x, Du + G^\intercal P) \leqslant \gamma. \tag{2.41}$$

Integrating this inequality against m_ϵ yields

$$\int_t^T \int_{\mathbb{T}^d} u\partial_t m_\epsilon + \partial_i u \partial_j(\tilde{A}_{ij} m_\epsilon) + m_\epsilon H(x, Du + G^\intercal P) + \int_{\mathbb{T}^d} m_\epsilon(t)u(t) - m_\epsilon(T)u_T \leqslant \int_t^T \int_{\mathbb{T}^d} \gamma m_\epsilon. \tag{2.42}$$

By (2.38), we have

$$\int_t^T \int_{\mathbb{T}^d} u\partial_t m_\epsilon + \partial_i u \partial_j(\tilde{A}_{ij} m_\epsilon) = \int_t^T \int_{\mathbb{T}^d} -\partial_i u R_\epsilon + \langle Du, w_\epsilon \rangle,$$

while the convexity of H in the last variable gives

$$\int_t^T \int_{\mathbb{T}^d} -m_\epsilon H^*\left(x, -\tfrac{w_\epsilon}{m_\epsilon}\right) \leqslant \int_t^T \int_{\mathbb{T}^d} \langle w_\epsilon, Du + G^\intercal P \rangle + m_\epsilon H(x, Du + G^\intercal P). \tag{2.43}$$

Combining these results yields

$$\int_{\mathbb{T}^d} m_\epsilon(t)u(t) \leqslant \int_{\mathbb{T}^d} m_\epsilon(T)u_T + \int_t^T \int_{\mathbb{T}^d} m_\epsilon \left(\gamma + H^*\left(x, -\tfrac{w_\epsilon}{m_\epsilon}\right)\right) + \langle G w_\epsilon, P \rangle + \partial_j u R_\epsilon.$$

Following now [11], we have that as $Du \in L^{\tilde{r}}$ (where we recall that $\tilde{r} \geqslant p$ or A is a constant matrix), and as m is continuous in time with respect to the topology on $P(\mathbb{T}^d)$, we have, as $\epsilon \to 0$,

$$\int_t^T \int_{\mathbb{T}^d} \partial_j u R_\epsilon \to 0,$$
$$\int_t^T \int_{\mathbb{T}^d} -m_\epsilon H^*\left(x, -\tfrac{w_\epsilon}{m_\epsilon}\right) \to \int_t^T \int_{\mathbb{T}^d} -m H^*\left(x, -\tfrac{w}{m}\right),$$
$$\int_{\mathbb{T}^d} m_\epsilon(T)u_T \to \int_{\mathbb{T}^d} m(T)u_T.$$

(For the second limit, cf. [10].) We also claim that

$$\int_t^T \int_{\mathbb{T}^d} \langle G w_\epsilon, P \rangle \to \int_t^T \int_{\mathbb{T}^d} \langle G w, P \rangle.$$

To see this, recall Equation (2.17) from Lemma 2.5.1. If $\frac{1}{s} + \frac{1}{pr} \geqslant 1$, we deduce that $w \in L^s$ and therefore $w_\epsilon \to w$ in L^s; from this the claim follows immediately. Otherwise, if $\frac{1}{s} + \frac{1}{pr} < 1$, then by Hypothesis (H1) we assume that G is constant. Therefore, we have

$$\int_t^T \int_{\mathbb{T}^d} \langle Gw_\epsilon, P\rangle = \int_t^T \left\langle \int_{\mathbb{T}^d} Gw_\epsilon, P \right\rangle = \int_t^T \left\langle \xi_1^\epsilon * \int_{\mathbb{T}^d} Gw, P \right\rangle \to \int_t^T \left\langle \int_{\mathbb{T}^d} Gw, P \right\rangle$$

because $t \mapsto \int_{\mathbb{T}^d} Gw$ is in L^s. Now since $u \in L^\kappa(Q)$, $m \in L^q(Q)$, and $\kappa \geqslant p$, $m_\epsilon u$ strongly converges to mu in $L^1(Q)$ and thus up to a subsequence, $\int_{\mathbb{T}^d} m_\epsilon(t)u(t) \to \int_{\mathbb{T}^d} m(t)u(t)$ a.e. We now have that

$$\int_{\mathbb{T}^d} m(t)u(t)\,\mathrm{d}x \leqslant \int_{\mathbb{T}^d} m(T)u_T\,\mathrm{d}x + \int_t^T \int_{\mathbb{T}^d} m\left(\gamma + H^*\left(x, -\tfrac{w}{m}\right)\right) + \langle P, Gw\rangle\,\mathrm{d}x\,\mathrm{d}t.$$

An analogous argument produces the other desired inequality, so now assume that equality holds in inequality (2.37) with $t = 0$. Then there is $t^* \in (0, T)$ where equality holds with $t = t^*$. Let

$$E_\sigma(t) := \Big\{(s, y) : s \in [t, T],\ m\left(H^*\left(y, -\tfrac{w}{m}\right) + H(x, Du + G^\intercal P\right) \geqslant$$

$$-\langle w, Du + G^\intercal P\rangle + \sigma\,\Big\}$$

If $|E_\sigma(t)| > 0$, then for $\epsilon > 0$ small enough, the set of s, y satisfying

$$m_\epsilon\left(H^*\left(y, -\tfrac{w_\epsilon}{m_\epsilon}\right) + H(x, Du + G^\intercal P)\right) \geqslant -\langle w_\epsilon, Du + G^\intercal P\rangle + \frac{\sigma}{2}$$

has measure larger than $\frac{|E_\sigma(t)|}{2}$. Then by (2.43), for the fixed choice of ϵ,

$$\int_{t^*}^T \int_{\mathbb{T}^d} -m_\epsilon H^*\left(x, -\tfrac{w_\epsilon}{m_\epsilon}\right) \leqslant \int_{t^*}^T \int_{\mathbb{T}^d} \langle w_\epsilon, Du + G^\intercal P\rangle + m_\epsilon H(x, Du + G^\intercal P) - |E_\sigma(t)|\sigma/4,$$

whereby we obtain strict inequality in (2.37) with $t = t^*$, a contradiction. Thus $|E_\sigma(t)| = 0$ for any σ and a.e. t,

$$\left\langle -w, Du + G^\intercal P \right\rangle = m\left(H(x, Du + G^\intercal P) + H^*(y, -\tfrac{w}{m})\right),$$

and hence

$$w = -mD_\xi H(\cdot, Du + G^\intercal P) \;\; a.e. \quad \text{in} \quad (0,T) \times \mathbb{T}^d.$$

□

Proof of Proposition 2.6.8. It is clear that the value of the relaxed problem is smaller than the value of problem (2.13). It remains to show the other inequality. For any $(m, w) \in \mathcal{K}_1$ with $mH^*(-w/m) \in L^1(Q)$, we have, by Fenchel-Young inequality and Lemma 2.6.9,

$$\begin{aligned} D(u, P, \gamma) \geqslant & -\int_{\mathbb{T}^d} u(0)m_0 + \int_0^T \left(\left\langle P(t), \textstyle\int_{\mathbb{T}^d} Gw\right\rangle - \Phi(\textstyle\int_{\mathbb{T}^d} Gw)\right) + \iint\limits_Q \left(\gamma m - F(m)\right) \\ \geqslant & -\int_{\mathbb{T}^d} u_T m(T) - \iint\limits_Q mH^*\left(-\frac{w}{m}\right) - \int_0^T \Phi(\textstyle\int_{\mathbb{T}^d} Gw) - \iint\limits_Q F(m) \\ = & -B(m, w). \end{aligned}$$

Maximizing the right-hand side with respect to (m, w), we obtain with Lemma 2.5.1 that

$$D(u, P, \gamma) \geqslant -\inf_{(m,w)\in\mathcal{K}_1} B(m, w) = \inf_{(u,P,\gamma)\in\mathcal{K}_0} D(u, P, \gamma),$$

which concludes the proof.

□

2.6.3 Existence of a Relaxed Solution

We establish now the existence of a relaxed solution.

2.6.10 Proposition. *The relaxed problem* (2.35) *has at least one solution* $(u, P, \gamma) \in \mathcal{K}$.

Proof. Let (u_n, P_n, γ_n) be a minimizing sequence for problem (2.13). By Proposition (2.6.8), it is also a minimizing sequence for the relaxed problem (2.35). We can, without loss of generality, assume that $\gamma_n \geqslant 0$, so long as we only require u_n to be a viscosity solution to the Hamilton-Jacobi equation. Let us replace γ_n with its positive part, i.e. $(\gamma_n)_+ := \max\{\gamma_n, 0\}$. Then we replace u_n with $\tilde{u}_n$, the continuous viscosity solution of

$$-\partial_t \tilde{u}_n - A_{ij}\partial_{ij}\tilde{u}_n + H(D\tilde{u}_n + G^\intercal P_n) = (\gamma_n)_+, \quad \tilde{u}_n(x,T) = u_T(x).$$

By [34], $\tilde{u}_n$ also satisfies this equation in the sense of distributions, and thus the new triple $(\tilde{u}_n, P_n, (\gamma_n)_+)$ is also a member of $\mathcal{K}$. We have $\tilde{u}_n \geqslant u_n$ and $F^*(\gamma_n) = F^*((\gamma_n)_+)$ for all $(x,t) \in Q$. Therefore, $D(\tilde{u}_n, P_n, (\gamma_n)_+) \leqslant D(u_n, P_n, \gamma_n)$, and thus the new sequence also minimizes D. The arguments below will then apply to $(\tilde{u}_n, P_n, (\gamma_n)_+)$ in place of (u_n, P_n, γ_n).

Step 1: [Bounds for (γ_n), (P_n), and (Du_n)]:

All constants C used in this part of the proof are independent of n. We integrate (2.20) against m_0 on Q and obtain

$$\int_{\mathbb{T}^d} u_n(0)m_0 + \iint\limits_Q \partial_j u_n \partial_i(A_{ij}m_0) + \iint\limits_Q H(Du_n + G^\intercal P_n)m_0 \leqslant \iint\limits_Q \gamma_n m_0 + \int_{\mathbb{T}^d} u_T m_0. \tag{2.44}$$

Let us recall that $m_0 \geqslant \frac{1}{C}$. The Hamiltonian term can be bounded from below as:

$$\iint\limits_{Q} H(Du_n + G^\intercal P_n)m_0 \geqslant \frac{1}{C}\|Du_n + G^\intercal P_n\|_r^r - C. \tag{2.45}$$

In light of the regularity assumptions on A and m_0, we also have that

$$\Big|\iint\limits_{Q} \partial_j u_n \partial_i (A_{ij} m_0)\Big| \leqslant C\|Du_n\|_1 \leqslant C\|Du_n\|_{\tilde{r}}. \tag{2.46}$$

Finally, the right-hand side of (2.44) is bounded by $C\|\gamma_n\|_p + C$. Combining this estimate with (2.45) and (2.46), we obtain that

$$\int_{\mathbb{T}^d} u_n(0)m_0 + \frac{1}{CB}\|Du_n + G^\intercal P_n\|_r^r - C\|Du_n\|_{\tilde{r}} \leqslant C\|\gamma_n\|_p + C \tag{2.47}$$

for any choice of $B \geqslant 1$. The constants C used are also independent of B. Now we use the fact that (u_n, P_n, γ_n) is a minimizing sequence and the growth assumptions on F^* and Φ^* to derive

$$-\int_{\mathbb{T}^d} u_n(0)m_0 + \frac{1}{C}\|P_n\|_{s'}^{s'} + \frac{1}{C}\|\gamma_n\|_p^p - C \leqslant D(u_n, P_n, \gamma_n) \leqslant C. \tag{2.48}$$

Summing up (2.47) and (2.48), we obtain

$$\frac{1}{CB}\|Du_n + G^\intercal P_n\|_r^r - C\|Du_n\|_{\tilde{r}} + \frac{1}{C}\|P_n\|_{s'}^{s'} + \frac{1}{C}\|\gamma_n\|_p^p \leqslant C\|\gamma_n\|_p + C. \tag{2.49}$$

Now by Hölder's inequality we have

$$\|Du\|_{\tilde{r}}^{\tilde{r}} \leqslant C\big(\|Du + G^\intercal P\|_{\tilde{r}}^{\tilde{r}} + \|G^\intercal P\|_{\tilde{r}}^{\tilde{r}}\big) \leqslant C\big(\|Du + G^\intercal P\|_r^r + \|P\|_{s'}^{s'} + 1\big)$$

and so

$$\frac{1}{CB}\|Du_n\|_{\tilde{r}}^{\tilde{r}} + \left[\frac{1}{C} - \frac{C}{B}\right]\|P_n\|_{s'}^{s'} - C\|Du_n\|_{\tilde{r}} + \frac{1}{C}\|\gamma_n\|_p^p \leqslant C\|\gamma_n\|_p + C. \tag{2.50}$$

We fix now $B = 2C^2$. The terms $\|Du_n\|_{\tilde{r}}$, $\|\gamma_n\|_p$ can be absorbed. For instance, the former can be absorbed by $\|Du\|_{\tilde{r}}^{\tilde{r}}$ insofar as for an arbitrarily small $\varepsilon > 0$, there exists

$C > 0$ (depending on ε) such that

$$\|Du_n\|_{\tilde{r}} \leqslant \varepsilon \|Du_n\|_{\tilde{r}}^{\tilde{r}} + C. \tag{2.51}$$

Taking ε small enough, we finally deduce from (2.50) the estimate

$$\|Du_n\|_{\tilde{r}}^{\tilde{r}} + \|P_n\|_{s'}^{s'} + \|\gamma_n\|_p^p \leqslant C, \tag{2.52}$$

so that $(\gamma_n)_{n\in\mathbb{N}}$ is bounded in $L^p(Q)$, $(P_n)_{n\in\mathbb{N}}$ is bounded in $L^{s'}((0,T);\mathbb{R}^k)$ and $(Du_n)_{n\in\mathbb{N}}$ is bounded in $L^{\tilde{r}}(Q)$. Inequality (2.49) further shows that $Du_n + G^\intercal P_n$ is bounded in $L^r(Q;\mathbb{R}^d)$. This implies that

$$\|H(Du_n + G^\intercal P_n)\|_1 \leqslant C.$$

Step 2 [Bound of u_n in $L^\kappa(Q)$]:

Now that we have estimates on P_n in $L^{s'}$, γ_n in L^p, Du_n in $L^{\tilde{r}}$, and $H(Du_n + G^\intercal P_n)$ in L^1, we can apply Corollary 2.6.3 in case $s' < r$ or Corollary 2.6.4 in case $s' \geqslant r$ and obtain $\|u_n\|_\kappa \leqslant C$, where κ is defined at the end of Section 2.6.1.

Step 3 [Conclusion]:

The rest of the proof is very similar to the proof of [11, Proposition 5.4], we only give the main lines. By passing to a subsequence, we assume without loss of generality that

$$u_n \rightharpoonup \bar{u} \text{ in } L^\kappa(Q), \quad Du_n \rightharpoonup D\bar{u} \text{ in } L^{\tilde{r}}(Q), \quad Du_n + G^\intercal P_n \rightharpoonup D\bar{u} + G^\top \bar{P} \text{ in } L^r(Q;\mathbb{R}^d),$$
$$\gamma_n \rightharpoonup \bar{\gamma} \text{ in } L^p(Q), \quad P_n \rightharpoonup \bar{P} \text{ in } L^{s'}(0,T).$$

Since H is convex, $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$. By weak lower semicontinuity arguments, we have

$$\liminf_{n\to\infty} \iint_Q F^*(\gamma_n) + \int_0^T \Phi^*(P_n) \geqslant \iint_Q F^*(\bar{\gamma}) + \int_0^T \Phi^*(\bar{P}).$$

Using exactly the same arguments as in [11, Proposition 5.4 (Step 3)], one can prove that

$$\limsup_{n\to\infty} \int_{\mathbb{T}^d} u_n(0)m_0 \leqslant \int_{\mathbb{T}^d} \bar{u}(0)m_0,$$

which proves the optimality of $(\bar{u}, \bar{P}, \bar{\gamma})$.

$\square$

2.7 Existence and Uniqueness of a Solution for the MFG System

We prove in this section the existence and uniqueness of a *weak solution* to the MFG system (2.2).

2.7.1 Definition. *We say that a quadruplet* $(u, P, m, w) \in L^{\kappa}(Q) \times L^{s'}(0,T) \times L^{q}(Q) \times L^{\frac{r'q}{r'+q-1}}(Q)$ *is a weak solution if*

(i) The following integrability conditions hold: $Du \in L^{\tilde{r}}(Q)$ *and* $mH^*(\cdot, -m/w)) \in L^1(Q)$.

(ii) Equation (2.2)-*(i) holds in the sense of distributions,*

$$-\partial_t u - A_{ij}\partial_{ij}u + H(Du + G^{\intercal}P) \leqslant f(m), \quad u(T) \leqslant u_T$$

(iii) Equation (2.2)-*(ii) holds in the sense of distributions,*

$$\partial_t m - \partial_{ij}\big(A_{ij}m\big) - \nabla\cdot w = 0, \quad m(0) = m_0,$$

(iv) Equations (2.2)-*(iii)-(iv) hold almost everywhere,*

(v) The following equality holds:

$$\iint\limits_{Q} \Big(mf(m) + mH^*(-w/m) + \langle P, Gw\rangle\Big) + \int_{\mathbb{T}^d} m(T)u_T - \int_{\mathbb{T}^d} m_0 u(0) = 0. \quad (2.53)$$

2.7.2 Theorem. *There exists a weak solution (u, P, m, w) to the MFG system* (2.2)*. It is unique in the following sense: if (u, P, m, w) and (u', P', m', w') are two solutions, then $m = m'$, $w = w'$, $P = P'$ a.e. and $u = u'$ in $\{m > 0\}$.*

2.7.3 Theorem. *Let $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ be a minimizer of* (2.16) *and $(\bar{u}, \bar{P}, \bar{\gamma})$ be a minimizer of* (2.35)*. Then, $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ is a weak solution of the MFG system and $\bar{\gamma} = f(\bar{m})$. Conversely, any weak solution $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ of the MFG system is such that $(\bar{m}, \bar{w})$ is the solution to* (2.16) *and $(\bar{u}, \bar{P}, f(\bar{m}))$ is a solution to* (2.35)*.*

Proof. **Part 1.** Let $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ be the solution to (2.16) and $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$ be a solution to (2.35). Condition (iii) of Definition (2.7.1) is verified by the definition of $\mathcal{K}_1$. By Lemma 2.5.1 and Proposition 2.6.8, these two problems have the same value, thus

$$\begin{aligned}
0 &= D(\bar{u}, \bar{P}, \bar{\gamma}) + B(\bar{m}, \bar{w}) \\
&= \iint\limits_Q \big(F^*(\bar{\gamma}) + F(\bar{m})\big) + \int_0^T \Big(\Phi^*(\bar{P}) + \Phi\big(\textstyle\int_{\mathbb{T}^d} G\bar{w}\big)\Big) \\
&\qquad + \iint\limits_Q \bar{m}H^*(-\bar{w}/\bar{m}) + \int_{\mathbb{T}^d} u_T \bar{m}(T) - \int_{\mathbb{T}^d} \bar{u}(0) m_0.
\end{aligned}$$

By the Fenchel-Young inequality, we have

$$F^*(\bar{\gamma}) + F(\bar{m}) \geqslant \bar{\gamma}\bar{m} \quad \text{for a.e. } (x,t) \in Q, \tag{2.54}$$

$$\Phi^*(\bar{P}) + \Phi\big(\textstyle\int_{\mathbb{T}^d} G\bar{w}\big) \geqslant \big\langle \bar{P}, \textstyle\int_{\mathbb{T}^d} G\bar{w}\big\rangle \quad \text{for a.e. } t \in (0,T) \tag{2.55}$$

thus

$$0 \geqslant \iint\limits_Q \Big(\bar{m}\bar{\gamma} + \bar{m}H^*(-\bar{w}/\bar{m}) + \langle P, Gw\rangle\Big) + \int_{\mathbb{T}^d} \bar{m}(T) u_T - \int_{\mathbb{T}^d} m_0 \bar{u}(0). \tag{2.56}$$

This implies first that $\bar{m}H^*(-\bar{w}/\bar{m}) \in L^1(Q)$. Moreover, by Lemma 2.6.9, inequality (2.56) is in fact an equality and $\bar{w} = -\bar{m}D_\xi H(D\bar{u} + G^\intercal \bar{P})$. Moreover, the equality holds a.e. in (2.54) and (2.55) therefore,

$$\bar{\gamma} = F'(\bar{m}) = f(\bar{m}) \quad \text{for a.e. } (x,t) \in Q,$$

$$\bar{P} = D\Phi\Big(\textstyle\int_{\mathbb{T}^d} G\bar{w}\Big) = \Psi\Big(\textstyle\int_{\mathbb{T}^d} G\bar{w}\Big) \quad \text{for a.e. } t \in (0,T).$$

Since (2.56) is an equality and $\bar{\gamma} = f(\bar{m})$, (2.53) (condition (v)) holds true. Further, by the definition of $\mathcal{K}$ and $\bar{\gamma} = f(\bar{m})$, condition (ii) holds. We conclude then that $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ is a weak solution to the MFG system.

Part 2. Let $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ be a weak solution to (2.2). Let $\bar{\gamma} = f(\bar{m})$. The growth condition on f implies that $\bar{\gamma} \in L^p(Q)$. Therefore, $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ and $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$. It remains to show that $(\bar{u}, \bar{P}, \bar{\gamma})$ solves (2.35) and that $(\bar{m}, \bar{w})$ solves (2.35).

The argument is very similar to the one used in Proposition 2.6.8. It mainly consists in showing that $D(\bar{u}, \bar{P}, \bar{\gamma}) + B(\bar{m}, \bar{w}) = 0$. Since $\bar{\gamma} = f(\bar{m}) = F'(\bar{m})$ a.e., we have by convexity of F that

$$F(\bar{m}) + F^*(\bar{\gamma}) = \bar{\gamma}\bar{m}, \quad \text{for a.e. } (x,t) \in Q.$$

Similarly, since $\bar{P} = \Psi\Big(\int_{\mathbb{T}^d} Gw\Big) = D\Phi\Big(\int_{\mathbb{T}^d} Gw\Big)$, we have

$$\Phi\Big(\textstyle\int_{\mathbb{T}^d} Gw\Big) + \Phi^*(\bar{P}) = \Big\langle P, \textstyle\int_{\mathbb{T}^d} Gw\Big\rangle, \quad \text{for a.e. } t \in (0,T).$$

These two equalities and (2.53) yield:

$$
\begin{aligned}
D(\bar u,\bar P,\bar\gamma)+B(\bar m,\bar w) = & \iint\limits_Q \big(F^*(\bar\gamma)+F(\bar m)\big)+\int_0^T \Big(\Phi\big(\textstyle\int_{\mathbb{T}^d} G\bar w\big)+\Phi^*(\bar P)\Big)\\
& +\iint\limits_Q \bar m H^*(-\bar w/\bar m)+\int_{\mathbb{T}^d}\big(u_T\bar m(T)-\bar u(0)m_0\big)\\
= & \iint\limits_Q \bar m\bar\gamma+\iint\limits_Q\langle \bar P, G\bar w\rangle+\int_{\mathbb{T}^d} u_T\bar m(T)-\bar u(0)m_0+\iint\limits_Q \bar m H^*(-\bar w/\bar m)\\
= & \; 0.
\end{aligned}
$$

As a consequence, we obtain

$$
\inf_{(u,P,\gamma)\in\mathcal{K}} D(u,P,\gamma)\leqslant D(\bar u,\bar P,\bar\gamma)=-B(\bar m,\bar w)\leqslant -\min_{(m,w)\in\mathcal{K}_1} B(m,w).
$$

The first and the last term being equal, the two above inequalities are equalities, which shows the optimality of optimality of $(\bar u,\bar P,\bar\gamma)$ and $(\bar m,\bar w)$, respectively.

□

Proof of Theorem 2.7.2. By Lemma 2.5.1, problem (2.16) has a solution $(\bar m,\bar w)$ and by Proposition 2.6.10, problem (2.35) has a solution $(\bar u,\bar P,\bar\gamma)$. By Theorem 2.7.3, $(\bar u,\bar P,\bar m,\bar w)$ is a weak solution to the MFG system.

Now, let (u_1,P_1,m_1,w_1) and (u_2,P_2,m_2,w_2) be two weak solutions. By Theorem 2.7.3, (m_1,w_1) and (m_2,w_2) are solutions to problem (2.16), they are therefore equal. Relation (2.2)-(iii) implies that $P_1=P_2$. Let $(\bar m,\bar w,\bar P)=(m_1,w_1,P_1)$ denote the common values. Let $\bar\gamma=f(\bar m)$. Then $(u_1,\bar P,\bar\gamma)$ and $(u_2,\bar P,\bar\gamma)$ lie in $\mathcal{K}$ (by definition of weak solutions).

For $t \in (0,T)$, $(m,w) \in \mathcal{K}_1$, and $(u,P,\gamma) \in \mathcal{K}$, we introduce

$$B_t(m,w) = \int_t^T \int_{\mathbb{T}^d} \Big(mH^*(-w/m) + F(m)\Big) + \int_t^T \Phi\big(\textstyle\int_{\mathbb{T}^d} Gw\big) + \displaystyle\int_{\mathbb{T}^d} u_T m(T)$$
$$D_t(u,P,\gamma) = -\int_{\mathbb{T}^d} u(0)m_0 + \int_t^T \Phi^*(P) + \int_t^T \int_{\mathbb{T}^d} F^*(\gamma).$$

Proceeding as in the proof of Proposition 2.6.8, we obtain that

$$\inf_{(u,P,\gamma)\in\mathcal{K}} D_t(u,P,\gamma) \geqslant -B_t(\bar{m},\bar{w}).$$

By Lemma 2.6.9 and relation (2.53),

$$\int_t^T \int_{\mathbb{T}^d} \Big(\bar{m}f(\bar{m}) + \bar{m}H^*(-\bar{w}/\bar{m}) + \langle \bar{P}, G\bar{w}\rangle\Big) + \int_{\mathbb{T}^d} \bar{m}(T)u_T - \int_{\mathbb{T}^d} \bar{m}(t)u_i(t) = 0$$

for a.e. $t \in (0,T)$ and for $i = 1,2$. Proceeding as in the proof of Theorem 2.7.3, we obtain that $-B_t(\bar{m},\bar{w}) = D_t(u_i,\bar{P},\bar{\gamma})$. Thus $(u_1,\bar{P},\bar{\gamma})$ and $(u_2,\bar{P},\bar{\gamma})$ minimize D_t over $\mathcal{K}$.

Let $\bar{u} = u_1 \vee u_2$. Adapting the proof in [11, Theorem 6.2], we obtain that $(\bar{u},\bar{P},\bar{\gamma}) \in \mathcal{K}$. Since $D_t(\bar{u}) \leqslant D_t(u_i)$, we deduce that $(\bar{u},P,\gamma)$ also minimize D_t. It follows that

$$\int_{\mathbb{T}^d} u_1(t)\bar{m}(t) = \int_{\mathbb{T}^d} u_2(t)\bar{m}(t) = \int_{\mathbb{T}^d} \bar{u}(t)\bar{m}(t)$$

As $u_1 \leqslant \bar{u}$ and $u_2 \leqslant \bar{u}$, this implies that $u_1 = u_2 = \bar{u}$ a.e. in $\{\bar{m} > 0\}$ and concludes the proof.

□

2.8 Regularity Estimates

In this section we adapt the methods used in [28, 29] to show that weak solutions of (2.1) possess extra regularity–Sobolev estimates in both space and time–not required by Definition 2 .7.1. T hese e stimates h old u nder g eneral *s trong monotonicity* assumptions on the coupling $f(x,m)$ and *coercivity* on the Hamiltonian. We divide our results into "space regularity," i.e. estimates on derivatives with respect to x, and "time regularity," estimates on derivatives with respect to t.

2.8.1 Space Regularity

Before stating the result, let us enumerate a few additional assumptions.

2.8.1 Assumption. *A_{ij} is constant for every i,j.*

2.8.2 Assumption (Strong monotonicity). *We have a Lipschitz estimate on f of the form*

$$|f(x,m) - f(y,m)| \leqslant C(m^{q-1}+1)|x-y| \quad \forall x,y \in \mathbb{T}^d,\ m \geqslant 0. \tag{2.57}$$

We also assume that $f(x,m)$ is strongly monotone in m, that is, there exists $c_f > 0$ such that

$$\big(f(x,\tilde{m}) - f(x,m)\big)(\tilde{m}-m) \geqslant c_f \min\{\tilde{m}^{q-2}, m^{q-2}\}|\tilde{m}-m|^2 \ \forall \tilde{m}, m \geqslant 0,\ \tilde{m} \neq m. \tag{2.58}$$

If $q < 2$ one should interpret 0^{q-2} as $+\infty$ in (2.58)*. In this way, when $\tilde{m} = 0$, for instance,* (2.58) *reduces to $f(x,m)m \geqslant c_f m^q$, as in the more regular case $q \geqslant 2$.*

2.8.3 Assumption (Coercivity). *There exist $j_1, j_2 : \mathbb{R}^d \to \mathbb{R}^d$ and $c_H > 0$ such that*

$$H(x,\xi) + H^*(x,\zeta) - \xi \cdot \zeta \geqslant c_H |j_1(\xi) - j_2(\zeta)|^2. \tag{2.59}$$

In particular, and in light of our restriction on the growth of H, we specify that $j_1(\xi) \sim |\xi|^{r/2-1}\xi$ and $j_2(\zeta) \sim |\zeta|^{r'/2-1}\zeta$.

2.8.4 Assumption. *$m_0 \in W^{2,\infty}(\mathbb{T}^d), u_T \in W^{2,\infty}(\mathbb{T}^d), G \in W^{2,\infty}\left(\mathbb{T}^d; \mathcal{L}(\mathbb{R}^k, \mathbb{R}^d)\right)$, and H^* is twice continuously differentiable in x with*

$$|D^2_{xx} H^*(x,\zeta)| \leqslant C_H \left(|\zeta|^{r'} + 1\right). \tag{2.60}$$

Notice that Assumption 2.8.2 holds for the canonical case $f(x,m) = m^{q-1}$ or even if $f(x,m) = \tilde{f}(x)m^{q-1}$ for some strictly positive, Lipschitz continuous function $\tilde{f}$ on $\mathbb{T}^d$. Assumption 2.8.3 likewise holds for a canonical structure $H(x,\xi) = c(x)|\xi|^r$ for some strictly positive, $\mathcal{C}^2$ smooth function $c(x)$ on $\mathbb{T}^d$.

2.8.5 Proposition. *Let Assumptions 2.8.1, 2.8.2, 2.8.3, 2.8.4 hold. Then, if (u,m) is a weak solution of* (2.1)*,*

$$\|m^{\frac{q}{2}-1} Dm\|_{L^2(Q} \leqslant C \quad \textit{and} \quad \|m^{1/2} D(j_1(Du))\|_{L^2(Q} \leqslant C,$$

where C is a constant depending only on the data.

Proof. Throughout we use the notation $g^\delta(x) = g(x+\delta)$ for any function depending on $x \in \mathbb{T}^d$.

Take a smooth minimizing sequence $(u_n, P_n, \gamma_n) \in \mathcal{K}_0$ for the dual problem. Integrate (2.14) by parts against u_n and rearrange to get

$$\iint\limits_Q mH(x, Du_n + G^\intercal P_n)\,\mathrm{d}x\,\mathrm{d}t = \int_{\mathbb{T}^d} (u_T m(T) - u_n(0)m_0)\,\mathrm{d}x + \iint\limits_Q \gamma_n m - \langle Du_n, w\rangle\,\mathrm{d}x\,\mathrm{d}t. \tag{2.61}$$

Step 1. The following estimates show that (up to a subsequence) $Du_n \rightharpoonup Du$ in $L^{\tilde{r}}_m([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$ (see Section 2.6 for definition of $\tilde{r}$, and NB $\tilde{r} \leqslant \min\{r, s'\}$):

Using Young's inequality and Assumption (H2) we get

$$\frac{1}{C}\iint_Q |Du_n + G^\intercal P_n|^{\tilde r}\, m \,\mathrm{d}x\,\mathrm{d}t \leqslant \frac{1}{C}\iint_Q |Du_n + G^\intercal P_n|^{r}\, m \,\mathrm{d}x\,\mathrm{d}t + C$$

$$\leqslant \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 + \iint_Q \left((\gamma_n)_+ m + Cm\left|\frac{w}{m}\right|^{r'}\right) \mathrm{d}x\,\mathrm{d}t + C. \quad (2.62)$$

By possibly increasing C we get

$$\begin{aligned}
&\frac{1}{C}\iint_Q |Du_n|^{\tilde r}\, m \,\mathrm{d}x\,\mathrm{d}t \\
&\leqslant \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 + \iint_Q \left((\gamma_n)_+ m + Cm\left|\frac{w}{m}\right|^{r'} + C|G^\intercal P_n|^{\tilde r}\, m\right) \mathrm{d}x\,\mathrm{d}t + C \\
&\leqslant \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 + \iint_Q \left((\gamma_n)_+ m + Cm\left|\frac{w}{m}\right|^{r'}\right) \mathrm{d}x\,\mathrm{d}t + C\int_0^T |P_n|^{s'}\,\mathrm{d}t + C.
\end{aligned} \quad (2.63)$$

Since P_n is bounded in $L^{s'}$, we have that Du_n is bounded in $L^{\tilde r}_m$ where we recall that $\tilde r = \min(r, s')$. Thus, up to a subsequence, $Du_n \rightharpoonup \zeta$ for some $\zeta \in L^{\tilde r}_m$. The argument that $\zeta = Du$ $m-$a.e. follows as in [28]. We also have, up to a subsequence, that $P_n \rightharpoonup P$ in $L^{s'}(0,T)$, and thus also that $Du_n + G^\intercal P_n \rightharpoonup Du + G^\intercal P$ in $L^r_m([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$. Indeed, the upper bound given by (2.62) shows that $Du_n + G^\intercal P_n$ converges weakly in L^r_m, and its limit must be equal to $Du + G^\intercal P$ a.e. by taking the limit of each summand.

Step 2. Now use u_n^δ and $u_n^{-\delta}$ as test functions in (2.14) to get

$$\begin{aligned}
&\int_{\mathbb{T}^d} (u_T^\delta m(T) - u_n^\delta(0) m_0)\,\mathrm{d}x = \\
&\qquad \iint_Q \Big(H(x+\delta, Du_n^\delta + (G^\delta)^\intercal P_n) m - \gamma_n^\delta m + Du_n^\delta \cdot w\Big)\,\mathrm{d}x\,\mathrm{d}t
\end{aligned} \quad (2.64)$$

and

$$\int_{\mathbb{T}^d} (u_T^{-\delta} m(T) - u_n^{-\delta}(0) m_0)\, \mathrm{d}x = \iint\limits_Q \left(H(x-\delta, Du_n^{-\delta} + (G^{-\delta})^T P_n) m - \gamma_n^{-\delta} m + Du_n^{-\delta} \cdot w \right) \mathrm{d}x\, \mathrm{d}t \tag{2.65}$$

We have the optimality condition

$$\int_{\mathbb{T}^d} (u_T m(T) - u(0) m_0)\, \mathrm{d}x = -\iint\limits_Q \left(H^*\left(x, -\frac{w}{m}\right) m + P \cdot (Gw) + f(x,m) m \right) \mathrm{d}x\, \mathrm{d}t. \tag{2.66}$$

Take (2.64) + (2.65) − 2(2.66) to get

$$\begin{aligned}
&\int_{\mathbb{T}^d} \left(\left(u_T^{\delta} + u_T^{-\delta} - 2u_T \right) m(T) - u_n(0) \left(m_0^{\delta} + m_0^{-\delta} - 2m_0 \right) \right) \mathrm{d}x \\
&= \iint\limits_Q \Bigg(H(x+\delta, Du_n^{\delta} + (G^{\delta})^T P_n) m + H^*\left(x+\delta, -\frac{w}{m}\right) m + \\
&\qquad \left(Du_n^{\delta} + (G^{\delta})^T P_n \right) \cdot w \, \mathrm{d}x\, \mathrm{d}t \\
&+ \iint\limits_Q \Bigg(H(x-\delta, Du_n^{-\delta} + (G^{-\delta})^T P_n) m + H^*\left(x-\delta, -\frac{w}{m}\right) m + \\
&\qquad \left(Du_n^{-\delta} + (G^{-\delta})^T P_n \right) \cdot w \, \mathrm{d}x\, \mathrm{d}t \\
&+ \iint\limits_Q \Bigg(\left(2f(x,m) - \gamma_n^{\delta} - \gamma_n^{-\delta} \right) m + 2P \cdot (Gw) - \\
&\qquad P_n \cdot (G^{\delta} w + G^{-\delta} w)\, \mathrm{d}x\, \mathrm{d}t - I
\end{aligned} \tag{2.67}$$

where

$$I := \iint\limits_Q \left(H^*\left(x+\delta, -\frac{w}{m}\right) + H^*\left(x-\delta, -\frac{w}{m}\right) - 2H^*\left(x, -\frac{w}{m}\right) \right) m \tag{2.68}$$

and where we have used

$$\int_{\mathbb{T}^d} \left(u_n^{\delta}(0) + u_n^{-\delta}(0) - 2u(0) \right) m_0\, \mathrm{d}x = \int_{\mathbb{T}^d} u_n(0) \left(m_0^{\delta} + m_0^{-\delta} - 2m_0 \right) \mathrm{d}x. \tag{2.69}$$

Since H is convex in the third argument, by the result of Step 1 and weak lower semicontinuity we have

$$\iint\limits_Q H(x\pm\delta, Du^{\pm\delta}+(G^{\pm\delta})^T P)m\,\mathrm{d}x\,\mathrm{d}t \leqslant \liminf \iint\limits_Q H(x\pm\delta, Du_n^{\pm\delta}+(G^{\pm\delta})^T P_n)m\,\mathrm{d}x\,\mathrm{d}t.$$

Letting $n\to\infty$ in (2.67) we obtain

$$\begin{aligned}
&\iint\limits_Q \Bigg(H(x+\delta, Du^\delta+(G^\delta)^T P)m + H^*\left(x+\delta, -\frac{w}{m}\right)m+ \\
&\qquad\qquad \Big(Du^\delta+(G^\delta)^T P\Big)\cdot w\,\mathrm{d}x\,\mathrm{d}t \\
&\quad+\iint\limits_Q \Bigg(H(x-\delta, Du^{-\delta}+(G^{-\delta})^T P)m + H^*\left(x-\delta, -\frac{w}{m}\right)m+ \\
&\qquad\qquad \Big(Du^{-\delta}+(G^{-\delta})^T P\Big)\cdot w\,\mathrm{d}x\,\mathrm{d}t \\
&\quad\leqslant \int_{\mathbb{T}^d}\Big(\Big(u_T^\delta+u_T^{-\delta}-2u_T\Big)m(T)-u(0)\Big(m_0^\delta+m_0^{-\delta}-2m_0\Big)\Big)\,\mathrm{d}x + I \\
&\qquad+\iint\limits_Q \Big(\Big(f(x+\delta, m^\delta)+f(x-\delta, m^{-\delta})-2f(x,m)\Big)m+ \\
&\qquad\qquad P\cdot\Big(\Big(G^\delta+G^{-\delta}-2G\Big)w\Big)\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{2.70}$$

By [28, computation (4.25)] we have

$$\begin{aligned}
&\int_{\mathbb{T}^d}\Big(f(x+\delta, m^\delta)+f(x-\delta, m^{-\delta})-2f(x,m)\Big)m\,\mathrm{d}x \\
&\quad\leqslant C|\delta|^2\left(1+\int_{\mathbb{T}^d}\min\{m^\delta, m\}^q\,\mathrm{d}x\right)-\frac{c_f}{2}\int_{\mathbb{T}^d}\min\{(m^\delta)^{q-2}, m^{q-2}\}|m^\delta-m|^2\,\mathrm{d}x.
\end{aligned} \tag{2.71}$$

Using estimate (2.59) on the left-hand side of (2.70), then using $|a+b|^2 \leqslant 2|a|^2 + 2|b|^2$, and combining with (2.71), then using Assumption 2.8.4 we deduce

$$\frac{c_H}{2} \iint\limits_Q \left| \left(Du^\delta + (G^\delta)^T P \right)^{r/2} - \left(Du^{-\delta} + (G^{-\delta})^T P \right)^{r/2} \right| m \, \mathrm{d}x \, \mathrm{d}t$$

$$+ \frac{c_f}{2} \int_{\mathbb{T}^d} \min\{(m^\delta)^{q-2}, m^{q-2}\} \left| m^\delta - m \right|^2 \mathrm{d}x$$

$$\leqslant |\delta|^2 \left(\|u_T\|_{W^{2,\infty}} + \|m_0\|_{W^{2,\infty}} \int_{\mathbb{T}^d} |u(0)| \, \mathrm{d}x + C_H \left(\left\| \frac{w}{m} \right\|_{L^r_m} + 1 \right) \right)$$

$$|\delta|^2 \left(C \left(1 + \int_{\mathbb{T}^d} \min\{m^\delta, m\}^q \, \mathrm{d}x \right) + \|G\|_{W^{2,\infty}} \iint\limits_Q |P \cdot w| \, \mathrm{d}x \, \mathrm{d}t \right). \quad (2.72)$$

□

2.8.2 Time Regularity

As in the previous subsection, we enumerate our assumptions before stating the main result.

2.8.6 Assumption. *We assume that* $A_{ij} = 0$.

We remark that Assumption 2.8.6 is much stronger than Assumption 2.8.1 but appears to be necessary, for technical reasons that appear in the estimates below.

2.8.7 Assumption (Strong monotonicity in time). *We assume that* (2.58) *holds.*

We assume that Ψ *is invertible, with inverse denoted by* $\left(\Psi^{-1}\right)(t, \cdot)$ *(for instance, it suffices to assume that its primitive* Φ *is strictly convex). We assume that, for some constant* $c_\Psi > 0$, $\left(\Psi^{-1}(\tilde{P}) - \Psi^{-1}(P)\right)$ *(- P)* $\geqslant c_\Psi \min\left\{ \left|\tilde{P}\right|^{s'-2}, |P|^{s'-2} \right\} |\tilde{P} - P|^2$

$\forall t, \tau \in [0, T]$, $P \in \mathbb{R}^k$.

2.8.8 Proposition. *Under Assumptions 2.8.6, 2.8.7, and* (2.60)*, for every* $\varepsilon > 0$*, there exists a constant* $C(\varepsilon)$*, depending only on* ε *and the data, such that*

$$\left\| \partial_t \left(m^{q/2} \right) \right\|_{L^2(Q_\varepsilon)} + \left\| \tfrac{\mathrm{d}}{\mathrm{d}t} \left(P^{s'/2} \right) \right\|_{L^2(\varepsilon, T-\varepsilon)} \leqslant C(\varepsilon) \tag{2.73}$$

where $Q_\varepsilon := \mathbb{T}^d \times (\varepsilon, T-\varepsilon)$*.*

2.8.9 Remark. *The proposition could also be proved for data depending on time, in particular with* $f(x,m)$ *and* $H(x,\xi)$ *replaced by* $f(t,x,m)$ *and* $H(t,x,\xi)$*, respectively. The only additional assumption needed would be a Lipschitz estimate in* t*, where the Lipschitz constant can depend on* x *(but not on* m *or* ξ*).*

Proof. Let $\varepsilon \in \mathbb{R}$ be small and $\eta : [0,T] \to [0,1]$ be smooth and compactly supported (with the support of η denoted $\mathrm{spt}(\eta)$) in $(0,T)$ such that

$$|\varepsilon| < \min \left\{ \mathrm{dist}(0, \mathrm{spt}(\eta)); \mathrm{dist}(T, \mathrm{spt}(\eta)) \right\}$$

and $\max_t |\varepsilon \eta'(t)| < 1$. If $\varepsilon > 0$ we set $\eta_\varepsilon(t) = t + \varepsilon \eta(t)$, which is a strictly increasing bijection from $[0,T]$ to itself. Then we set $\eta_{-\varepsilon} = \eta_\varepsilon^{-1}$, which is also smooth by the inverse function theorem. For competitors (u, P, γ) of the minimization problem for $\mathcal{A}$, let us define

$$u^\varepsilon(x,t) := u(x, \eta_\varepsilon(t)); \quad P^\varepsilon(t) = P(\eta_\varepsilon(t)); \quad \gamma^\varepsilon(x,t) := \eta_\varepsilon'(t) \gamma(x, \eta_\varepsilon(t)).$$

Notice that by construction, if $t \in \{0, T\}$ then $u(x,t) = u^\varepsilon(x,t)$ and $\gamma(x,t) = \gamma^\varepsilon(x,t)$, provided that $\gamma(x,t)$ is well-defined.

Similarly, for competitors (m, w) of minimization problem for $\mathcal{B}$, we define

$$m^\varepsilon(x,t) := m(x, \eta_\varepsilon(t)); \quad w^\varepsilon(x,t) := \eta_\varepsilon'(t) w(x, \eta_\varepsilon(t))$$

and here as well if $t \in \{0, T\}$ then $m(x,t) = m^\varepsilon(x,t)$ and $w(x,t) = w^\varepsilon(x,t)$.

We define moreover perturbations on the data as

$$\Phi^\varepsilon(t,v) := \eta'_\varepsilon(t)\Phi\left(v/\eta'_\varepsilon(t)\right),$$

$$f^\varepsilon(t,x,m) := \eta'_\varepsilon(t) f(x,m); \quad F^\varepsilon(t,x,m) := \eta'_\varepsilon(t) F(x,m),$$

from which the Legendre transforms w.r.t. the last variable satisfy

$$(\Phi^\varepsilon)^*(t,P) = \eta'_\varepsilon(t)\Phi(P), \ (F^\varepsilon)^*(t,x,\gamma) := \eta'_\varepsilon(t) F^*(x,\gamma/\eta'_\varepsilon(t)).$$

Finally, we define

$$H^\varepsilon(t,x,\xi) := \eta'_\varepsilon(t) H(x,\xi), \quad \text{thus } (H^\varepsilon)^*(x,\zeta) := \eta'_\varepsilon(t) H^*(x,\zeta/\eta'_\varepsilon(t)).$$

Step 1. Take a smooth minimizing sequence (u_n, P_n, γ_n) in $\mathcal{K}_0$. Use $u_n^{\pm\varepsilon}$ as a test function in $\partial_t m + \nabla \cdot w = 0$ to get

$$\int_{\mathbb{T}^d} \big(u_T m(T) - u_n(0) m_0\big)\, \mathrm{d}x \geqslant \iint_Q \big(H^\varepsilon(t,x,Du_n^\varepsilon + G^\intercal P^\varepsilon(t))m \\ - \gamma_n^\varepsilon m + Du_n^\varepsilon \cdot w \, \mathrm{d}x\, \mathrm{d}t \tag{2.74}$$

and

$$\int_{\mathbb{T}^d} \big(u_T m(T) - u_n(0) m_0\big)\, \mathrm{d}x \geqslant \iint_Q \big(H^{-\varepsilon}(t,x,Du_n^{-\varepsilon} + G^\intercal P^{-\varepsilon}(t))m \\ - \gamma_n^{-\varepsilon} m + Du_n^{-\varepsilon} \cdot w \, \mathrm{d}x\, \mathrm{d}t. \tag{2.75}$$

Take (2.74) + (2.75) − 2(2.66) to get

$$
\begin{aligned}
&\int_{\mathbb{T}^d} 2\left(u(0)-u_n(0)\right) m_0 \,\mathrm{d}x \\
&\quad \geqslant \iint\limits_Q \left(H^\varepsilon(x, Du_n^\varepsilon + G^\intercal P^\varepsilon) + H^*\left(x, -\frac{w}{m}\right) + (Du_n^\varepsilon + G^\intercal P^\varepsilon)\cdot \frac{w}{m}\right) m \,\mathrm{d}x\,\mathrm{d}t \\
&+ \iint\limits_Q \left(H^{-\varepsilon}(x, Du_n^{-\varepsilon} + G^\intercal P^{-\varepsilon}) + H^*\left(x, -\frac{w}{m}\right) + \left(Du_n^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right)\cdot \frac{w}{m}\right) m \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \iint\limits_Q \left(2f(x,t,m) - \gamma_n^\varepsilon - \gamma_n^{-\varepsilon}\right) m \,\mathrm{d}x\,\mathrm{d}t + \\
&\qquad\qquad \iint\limits_Q \left(2P\cdot(Gw) - P^\varepsilon\cdot(Gw) - P^{-\varepsilon}\cdot(Gw)\right)\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{2.76}
$$

Letting $n\to\infty$ we get

$$
\begin{aligned}
&\iint\limits_Q \left(H^\varepsilon(x, Du^\varepsilon + G^\intercal P^\varepsilon) + (H^*)^\varepsilon\left(x, -\frac{w}{m}\right) + (Du^\varepsilon + G^\intercal P^\varepsilon)\cdot \frac{w}{m}\right) m \,\mathrm{d}x\,\mathrm{d}t \\
&+ \iint\limits_Q \left(H^{-\varepsilon}(x, Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}) + (H^*)^{-\varepsilon}\left(x, -\frac{w}{m}\right) + \left(Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right)\cdot \frac{w}{m}\right) m \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \iint\limits_Q \left(2f(x,m) - f^\varepsilon(t,x,m^\varepsilon) - f^{-\varepsilon}(t,x,m^{-\varepsilon})\right) m \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad\qquad + \iint\limits_Q \left(2P\cdot(Gw) - P^\varepsilon\cdot(Gw) - P^{-\varepsilon}\cdot(Gw)\right)\mathrm{d}x\,\mathrm{d}t \leqslant R(\varepsilon)
\end{aligned} \tag{2.77}
$$

where

$$
R(\varepsilon) := \iint\limits_Q \left((H^*)^\varepsilon\left(x, -\frac{w}{m}\right) + (H^*)^{-\varepsilon}\left(x, -\frac{w}{m}\right) - 2H^*\left(x, -\frac{w}{m}\right)\right)\mathrm{d}x\,\mathrm{d}t. \tag{2.78}
$$

Arguing as in [29, Proposition 3.3, Step 1] and using the estimate on $D^2_{xx}H^*$, we have $R(\varepsilon) = O(\varepsilon^2)$.

Next we perform changes of variables and the relation $P = D_v\Phi(t, \int Gw)$ to rewrite

$$\iint\limits_Q \big(2P\cdot(Gw) - P^\varepsilon\cdot(Gw) - P^{-\varepsilon}\cdot(Gw)\big)\,\mathrm{d}x\,\mathrm{d}t = \iint\limits_Q (P^\varepsilon - P)\cdot(Gw^\varepsilon - Gw)\,\mathrm{d}x\,\mathrm{d}t$$

$$= \int_0^T (P^\varepsilon - P)\cdot\Big(\left(\Psi^{-1}\right)^\varepsilon(t, P^\varepsilon) - \Psi^{-1}(t,P)\Big)\,\mathrm{d}t. \quad (2.79)$$

Using the same argument as in [29, Proposition 3.3, Step 4], Assumption 2.8.7 implies

$$\int_0^T (P^\varepsilon - P)\cdot\Big(\left(\Psi^{-1}\right)^\varepsilon(t, P^\varepsilon) - \left(\Psi^{-1}\right)(t,P)\Big)\,\mathrm{d}t$$
$$\geqslant \frac{c_\Psi}{2}\int_0^T \min\big\{|P^\varepsilon(t)|, |P(t)|\big\}^{s'-2}|P^\varepsilon(t) - P(t)|^2\,\mathrm{d}t - C|\varepsilon|^2\int_0^T |P(t)|^{s'}\,\mathrm{d}t. \quad (2.80)$$

We use an analogous argument (or see [29, Proposition 3.3, Step 4]) we deduce

$$\iint\limits_Q \big(2f(x,m) - f^\varepsilon(t,x,m^\varepsilon) - f^{-\varepsilon}(t,x,m^{-\varepsilon})\big)\,m\,\mathrm{d}x\,\mathrm{d}t$$

$$= \iint\limits_Q \big(f^\varepsilon(t,x,m^\varepsilon) - f(x,m)\big)\,(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t$$

$$\geqslant \frac{c_f}{2}\iint\limits_Q \min\{m^\varepsilon, m\}^{q-2}|m^\varepsilon - m|^2\,\mathrm{d}x\,\mathrm{d}t - C|\varepsilon|^2\iint\limits_Q m^q\,\mathrm{d}x\,\mathrm{d}t. \quad (2.81)$$

Finally, using Assumption 2.8.3 we get

$$\iint_Q \left(H^\varepsilon(x, Du^\varepsilon + G^\intercal P^\varepsilon) + (H^*)^\varepsilon \left(x, -\frac{w}{m}\right) + (Du^\varepsilon + G^\intercal P^\varepsilon) \cdot \frac{w}{m} \right) m \, \mathrm{d}x \, \mathrm{d}t$$

$$+ \iint_Q \left(H^{-\varepsilon}(x, Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}) + (H^*)^{-\varepsilon} \left(x, -\frac{w}{m}\right) + \left(Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right) \cdot \frac{w}{m} \right) m \, \mathrm{d}x \, \mathrm{d}t$$

$$\geqslant c_H \iint_Q \left| j_1 \left(Du^\varepsilon + G^\intercal P^\varepsilon\right) - j_2 \left(\frac{w}{m}\right) \right|^2 m \, \mathrm{d}x \, \mathrm{d}t$$

$$+ c_H \iint_Q \left| j_1 \left(Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right) - j_2 \left(\frac{w}{m}\right) \right|^2 m \, \mathrm{d}x \, \mathrm{d}t$$

$$\geqslant \frac{c_H}{2} \iint_Q \left| j_1 \left(Du^\varepsilon + G^\intercal P^\varepsilon\right) - j_1 \left(Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right) \right|^2 m \, \mathrm{d}x \, \mathrm{d}t \quad (2.82)$$

Combining (2.78), (2.80), (2.81), and (2.82) with (2.77), we get

$$\frac{c_H}{2} \iint_Q \left| j_1 \left(Du^\varepsilon + G^\intercal P^\varepsilon\right) - j_1 \left(Du^{-\varepsilon} + G^\intercal P^{-\varepsilon}\right) \right|^2 m \, \mathrm{d}x \, \mathrm{d}t$$

$$+ \frac{c_f}{2} \iint_Q \min\left\{m^\varepsilon, m\right\}^{q-2} |m^\varepsilon - m|^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$+ \frac{c_\Psi}{2} \int_0^T \min\left\{|P^\varepsilon(t)|, |P(t)|\right\}^{s'-2} |P^\varepsilon(t) - P(t)|^2 \, \mathrm{d}t \leqslant C|\varepsilon|^2, \quad (2.83)$$

where we have used the estimates on $\iint_Q m^q \, \mathrm{d}x \, \mathrm{d}t$ and $\int_0^T |P(t)|^{s'} \, \mathrm{d}t$. The conclusion follows.

□

CHAPTER THREE
Classical Solution for a Mean Field Game of Moderate Interactions

3.1 Abstract

We analyze a system of partial differential equations that model a potential mean field game of moderate interactions. Such a game models agents considering the positions of nearby opponents to be of higher import, and so decisions made concentrate on nearby information. The existence of classical solutions on any finite time horizon is provided under generic assumptions. The augmentation here from traditional Mean Field Games is the introduction of a new local coupling term that can be viewed as part of the Hamiltonian.

3.2 Introduction

We provide the existence of classical solutions to the following coupled PDE system:

$$\begin{cases} -u_t - \sigma\Delta u + H(\nabla u) - b(x,m)\cdot\nabla u = f(x,m) \qquad (x,t)\in Q \\ m_t - \sigma\Delta m + \nabla\cdot[m(b(x,m) - D_pH(\nabla u))] = 0 \qquad (x,t)\in Q \\ u(x,T) = g(x); \qquad m(x,0) = m_0(x), \qquad x\in\mathbb{T}^d \end{cases} \tag{3.1}$$

where $x \in \mathbb{T}^d$, the d-dimensional torus and $Q := \mathbb{T}^d \times [0,T]$ for some $T > 0$ fixed. Solutions to the system (3.1) are Nash Equilibria for a class of Mean Field Game, a competition amongst a high population of agents each attempting to optimize a personal value function that in turn depends on the state of all players. Mean Field Games were introduced in the works of Lasry and Lions [39] and Huang et al [33]. The term Mean Field, inspired by physics, relates to each player viewing the remaining players as one large entity, rather than attempting to view competitors individually. Applications in a variety of fields interested in social phenomena abound: Economics, logistics, biology, physics. In a typical Mean Field Game, the cost functional to be optimized by each player typically incorporates a monotone interaction term $f(m)$, where m denotes the distribution of player states. Such a coupling term often bakes in a penalty for being too close, for instance in a crowd of individuals moving to a new location, individuals still value personal space for safety and comfort, despite wanting to quickly reach their destination. The model considered presently introduces a new coupling in the function $b(x,m) \cdot \nabla u$, which involves again the distribution of player states as well as the solution u itself. In [21], the existence (but not uniqueness) of weak solutions with explicitly quadratic Hamiltonian was proved, with uniqueness on a small enough time horizon. We extend this by providing the existence of classical solutions for any finite time horizon, and a more general Hamiltonian.

With the eventual aim of employing the Leray-Schauder fixed point theorem, the parameterized system $(\mathrm{MFG})_\tau$ will be given by, with $\tau \in [0,1]$,

$$\begin{cases} -u_t - \sigma\Delta u + \tau\left[H(\nabla u) - b(x,m)\cdot\nabla u\right] = \tau f(x,m) \quad (x,t) \in Q \\ m_t - \sigma\Delta m + \tau\nabla\cdot[m(b(x,m) - D_pH(\nabla u))] = 0 \quad (x,t) \in Q \\ u(x,T) = \tau g(x); \qquad m(x,0) = m_0(x) \quad x \in \mathbb{T}^d. \end{cases} \tag{3.2}$$

For the convenience of the reader we state the Leray-Schauder Theorem and clarify the space X to be used for our purposes.

3.2.1 Theorem (Leray-Schauder Fixed Point). *Let X be a Banach space and let $T : X \times [0,1] \to X$ be a continuous and compact mapping. Let $x_0 \in X$. Assume that $T(x,0) = x_0$ 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 $T(x,1) = x$.*

In application of this theorem we will let $X := W^{2,1,p}(Q) \times W^{2,1,p}(Q)$, and at that point the mapping T will be constructed.

3.3 Assumptions

Suppose that

$$f : \mathcal{P}(\mathbb{T}^d) \to \mathcal{C}^2(\mathbb{T}^d), \quad g : \mathbb{T}^d \to \mathbb{R} \quad \text{with} \quad \|f\|_{\mathcal{C}^2} + \|g\|_{\mathcal{C}^2} + \|b\|_{\mathcal{C}^2} \leqslant C.$$

Suppose further that $g, b \in \mathcal{C}^{2+\alpha}$.

We remark that $f(m)(x)$, as a function on $\mathbb{T}^d$ for each $m \in \mathcal{P}(\mathbb{T}^d)$ will be occasionally presented as $f(x,m)$ as it appears above in the system $(\mathrm{MFG})_\tau$, with the same remark applying to $b(m)(x)$.

Assumptions on the Hamiltonian. For $H : Q \times \mathbb{R}^d \to \mathbb{R}$, it is assumed to be strictly convex in the third variable p, and that for $(x,t) \in Q$,

$$\frac{1}{C}|p|^2 - C \leqslant H(x,t,p) \leqslant C|p|^2 + C. \tag{3.3}$$

The convex conjugate with respect to p, given as $H^*(x,t,p) := L(x,t,-p)$, then adheres as well to

$$\frac{1}{C}|p|^2 - C \leqslant L(x,t,p) \leqslant C|p|^2 + C. \tag{3.4}$$

We note also that the convex conjugate of $\tilde{H}(p) := H(p) - b(x,m) \cdot p$ is

$$\tilde{H}^*(v) := \sup_p \left\{ \langle v + b(x,m), p \rangle - H(p) \right\} = H^*(x,t,v+b(x,m)) = L(x,t,-v-b(x,m)).$$

That is,

$$\tilde{L}(x,t,v) = L(x,t,v-b(x,m)).$$

We further assume that

$$|L(x,t,p) - L(y,t,p)| \leqslant C|x-y| \left(1 + |p|^2\right),$$

and that

$$|L(x,t,p) - L(x,t,v)| \leqslant C|p-v| \left(1 + |p|^2 + |v|^2\right)$$

uniformly in x.

3.4 Existence of Classical Solution

3.4.1 A priori Estimates on Fixed Points

We first present a set of results from [5], [37] that will be often cited in what follows.

3.4.1 Theorem. *Let $p > d+2$. 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*

$$u_t - \sigma \Delta u = h, \qquad u(x,0) = u_0(x), \qquad (x,t) \in Q$$

satisfies

$$\|u\|_{W^{2,1,p}(Q)} \leqslant C(\|u_0\|_{W^{2-2/p,p}(\mathbb{T}^d)} + \|h\|_{L^p(Q)}).$$

Further, there exists $\delta \in (0,1)$ and $\tilde{C} > 0$ such that

$$\|u\|_{C^\delta(Q)} + \|\nabla u\|_{C^\delta(Q,\mathbb{R}^d)} \leqslant \tilde{C}\|u\|_{W^{2,1,p}(Q)}.$$

Proof. For the first assertion see [5] Theorem 6, for the second see [37] Lemma II.3.3, page 80 and Corollary, page 342.

□

3.4.2 Theorem. *Let $p > d+2$. For all $\alpha \in (0,1)$, for all $R > 0$, there exist $\beta \in (0,1)$ and $C > 0$ such that for all $u_0 \in \mathcal{C}^{2+\alpha}(\mathbb{T}^d)$, $b \in \mathcal{C}^{\alpha,\alpha/2}(Q,\mathbb{R}^d)$, $c \in \mathcal{C}^{\alpha,\alpha/2}(Q)$, and $h \in \mathcal{C}^{\alpha,\alpha/2}(Q)$ solving*

$$\partial_t u - \sigma\Delta u + \langle b, \nabla u\rangle + cu = h, \quad u(x,0) = u_0$$

and satisfying that each norm in the respective space of u_0, b, c, h be no more than R, the solution u lies in $\mathcal{C}^{2+\beta,1+\beta/2}(Q)$ and satisfies $\|u\|_{\mathcal{C}^{2+\beta,1+\beta/2}}(Q) \leqslant C$.

Proof. See [5].

□

3.4.3 Proposition. *Let u_τ, m_τ solve (MFG)$_\tau$. Then, u_τ and m_τ are in $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ for some Hölder coefficient α. Further, $\nabla u_\tau \in \mathcal{C}^\alpha(Q)$.*

To prove this, we first establish a series of lemmas, and remind the reader that in what follows, C is a constant that may change from line to line, but may only depend on the data. Important as well, C is independent of the Leray-Schauder parameter τ.

3.4.4 Lemma. $\|u_\tau\|_\infty + \|\nabla u_\tau\|_\infty \leqslant C$

Proof. Since u_τ solves $(\text{MFG})_\tau$, it is the value function associated with the stochastic optimal control problem given by

$$u_\tau(x,t) = \tau \left(\inf_{\alpha \in L^2_{\mathbb{F}}(t,T;\mathbb{R}^d)} \mathbb{E}\left[\int_t^T \big(L(X_s, s, \alpha_s + b(m)(X_s)) + f(m)(X_s) \big) \, \mathrm{d}s + g(X_T) \right] \right) \tag{3.5}$$

where $(X_s)_{s\in[t,T]}$ is the solution to the stochastic dynamic

$$\mathrm{d}X_s = \tau \alpha_s \, \mathrm{d}s + \sqrt{2\sigma} \, \mathrm{d}B_s, \quad X_t = x, \tag{3.6}$$

the infimum taken over the set of stochastic processes on (t,T), with values in $\mathbb{R}^d$, adapted to the filtration $\mathbb{F}$ generated by the Brownian motion $(B_s)_{s\in[0,T]}$ with finite second moment: $\mathbb{E}\left[\int_t^T |\alpha(s)|^2 \, \mathrm{d}s\right] < \infty$. It follows then that u_τ is bounded above by choosing $\alpha = 0$, and appealing to the fact that $\|b\|_\infty + \|f\|_\infty + \|g\|_\infty \leqslant C$. We also bound u_τ from below via assumptions (3.3): For any choice of α,

$$u_\tau(x,t) \geqslant \tau \mathbb{E}\left[\int_t^T \frac{1}{C} |\alpha + b(m)(X_s)|^2 - C + f(m)(X_s) \, \mathrm{d}s + g(X_T) \right] \geqslant -\tau C \geqslant -C.$$

where once more C depends only on $\|f\|_\infty, \|b\|_\infty, \|g\|_\infty$ and that $\mathbb{E}\left[\int_t^T |\alpha|^2 \, \mathrm{d}s\right] \leqslant C$.

To bound ∇u_τ, choose $\varepsilon \in (0,1)$. For any $(x,t) \in Q$, take an ε-optimal control $\tilde{\alpha}_s$ for (3.5). Set

$$dX_s = \tau \tilde{\alpha}_s + \sqrt{2\sigma} \, \mathrm{d}B_s, \quad X_t = x, \quad Y_s = X_s - x + y,$$

from which it follows that $Y_s - X_s = y - x$ and $Y_t = y$. We then have that

$$u_\tau(x,t) + \varepsilon \geqslant \tau \mathbb{E}\left[\int_t^T \big(L(X_s, s, \tilde{\alpha}_s + b(m)(X_s)) + f(m)(X_s) \big) \, \mathrm{d}s + g(X_T) \right],$$

$$u_\tau(y,t) \leqslant \tau \mathbb{E}\left[\int_t^T \big(L(Y_s,s,\tilde{\alpha}_s + b(m)(Y_s)) + f(m)(Y_s)\big)\,\mathrm{d}s + g(Y_T)\right].$$

Thus,

$$u_\tau(y,t) - u_\tau(x,t) \leqslant \varepsilon + A + B$$

where

$$A = \mathbb{E}\left[\int_t^T L(Y_s,s,\tilde{\alpha}_s + b(m)(Y_s)) - L(X_s,s,\tilde{\alpha}_s + b(m)(X_s))\,\mathrm{d}s\right]$$

$$B = \mathbb{E}\left[\int_t^T f(m)(Y_s) - f(m)(X_s)\,\mathrm{d}s + g(Y_T) - g(X_T)\right].$$

By the assumptions (3.3), we have that

$$|B| \leqslant \mathbb{E}\left[\int_t^T C|Y_s - X_s|\,\mathrm{d}s + C|Y_T - X_T|\right] = \mathbb{E}\left[\int_t^T C|y-x|\,\mathrm{d}s + C|y-x|\right] = C|y-x|,$$

where the final C depends on the Lipschitz constants of f, g, and the quantity $T-t$.

By the assumptions (3.3), it follows that

$$|A| \leqslant \mathbb{E}\Bigg[C|y-x|\left(1 + \int_t^T |\tilde{\alpha}_s + b(m)(Y_s)|^2\,\mathrm{d}s + \int_t^T |\tilde{\alpha}_s + b(m)(Y_s)|^2\,\mathrm{d}s\right)$$
$$+ C\int_t^T |b(m)(Y_s) - b(m)(X_s)|\left(1 + |\tilde{\alpha}_s + b(m)(Y_s)|^2\right)\Bigg]$$

As $\|b\|_\infty \leqslant C$ and Lipschitz, and $\mathbb{E}\left[\int_t^T |\tilde{\alpha}_s|^2\,\mathrm{d}s\right] < \infty$, we have that $|A| \leqslant C|y-x|$ as well. With $x, y \in \mathbb{T}^d$ it follows that $\|\nabla u_\tau\|_\infty \leqslant C$, with bound independent of τ, and u_τ Lipschitz.

□

3.4.5 Lemma. $\|u_\tau\|_{W^{2,1,p}} + \|u_\tau\|_{\mathcal{C}^\alpha} + \|\nabla u_\tau\|_{\mathcal{C}^\alpha} \leqslant C$

Proof. The previous result gives $h := \tau(f - H(\nabla u_\tau) - b\cdot\nabla u_\tau)$ bounded, and thus in $L^p(Q)$ with $\|h\|_{L^p(Q)} \leqslant C$. Both Sobolev and Hölder space estimates of u_τ then follow immediately from Theorem (3.4.1), as well as the Hölder estimate of ∇u_τ, with

$$\|u_\tau\|_{W^{2,1,p}} \leqslant C\left(\|g\|_{W^{2-2/p,p}(\mathbb{T}^d)} + \|h\|_{L^p(Q)}\right) \leqslant C,$$

$$\|u_\tau\|_{\mathcal{C}^\alpha} + \|\nabla u_\tau\|_{\mathcal{C}^\alpha} \leqslant C\|u\|_{W^{2,1,p}} \leqslant C.$$

□

3.4.6 Lemma. $\|m_\tau\|_{\mathcal{C}^\alpha} \leqslant C$

Proof. This result follows from the second assertion of Theorem (3.4.1), as the Fokker-Planck equation can be viewed as a parabolic equation with coefficients in $L^p(Q)$, by the boundedness of m_τ.

□

3.4.7 Lemma. $\|u_\tau\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}} \leqslant C$

Proof. Since $\|\nabla u_\tau\|_{\mathcal{C}^\alpha} \leqslant C$ and $\|H(x,t,\cdot)\|_{\mathcal{C}^\alpha} \leqslant C$ on bounded sets, it holds that

$$\|H(\nabla u_\tau) - b \cdot \nabla u_\tau\|_{\mathcal{C}^\alpha} \leqslant C.$$

This fact, along with assumptions (3.3) yield the result by Theorem (3.4.2).

□

3.4.8 Lemma. $\|m_\tau\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}} \leqslant C$

Proof. Since $\|u_\tau\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}} \leqslant C$ has been shown in the previous step, we have that m_τ satisfies

$$(m_\tau)_t - \sigma\Delta(m_\tau) = \tau\nabla \cdot [m_\tau(\nabla(u_\tau) - b(x, m_\tau))], \quad m_\tau(x,0) = m_0(x),$$

a parabolic equation with Hölder coefficients and the result follows once more from Theorem (3.4.2).

□

3.4.2 Application of the Leray-Schauder Theorem

We now state and prove the main result.

3.4.9 Theorem. *Under the given assumptions of Chapter Three, the system* (3.1) *has a classical solution for any finite time horizon.*

Proof. Define $X := [\mathcal{C}^{1+\alpha,\alpha/2}]^2$. For a given $(u, m, \tau) \in X \times [0,1]$, the pair $(\tilde{u}, \tilde{m}) = \mathbf{T}(u, m, \tau)$ is defined as follows: $\tilde{u}$ is the solution to

$$-\tilde{u}_t - \sigma\Delta\tilde{u} + \tau\left[H(\nabla u) - b(x,m)\cdot\nabla u\right] = \tau f(x,m); \quad \tilde{u}(x,T) = \tau g(x) \tag{3.7}$$

and $\tilde{m}$ is the solution to

$$\tilde{m}_t - \sigma\Delta\tilde{m} + \tau\nabla\cdot[m(-\nabla u + b(x,m))] = 0; \quad \tilde{m}(x,0) = m_0(x) \tag{3.8}$$

$\mathbf{T}$ *is constant for* $\tau = 0$ It follows from construction that $\mathbf{T}(u, m, 0)$ is constant for all choices of u, m.

Fixed points of $\mathbf{T}$ *are a priori bounded in* X. Suppose now that $\mathbf{T}(u, m, \tau) = (u, m)$. Then u, m solve MFG_τ, and by Proposition (3.4.3) there exists a constant C independent of τ, u, and m such that

$$\|(u,m)\|_X < C.$$

$\mathbf{T}$ *is continuous.* By Theorem 6 BHP (RPT), the solution to

$$-\tilde{u}_t - \sigma\Delta\tilde{u} = \tau f(x,m) - \tau H(\nabla u) + \tau b(x,m)\cdot\nabla u$$

in X is a continuous mapping of the right hand side due to the results of Proposition (3.4.3), and thus continuous with respect to (u, m, τ) by composition. With $\tilde{u}$ continuously depending on the input, the same can then be said for $\tilde{m}$ with right hand side

$$-\tau\nabla \cdot [m(-\nabla\tilde{u} + b(x,m))].$$

T *is compact.* By Theorem (3.4.2), a sequence (u_k, m_k) such that, with constant C independent of u_k, m_k, τ,

$$\|u_k\|_{\mathcal{C}^{1+\alpha,\alpha/2}} + \|m_k\|_{\mathcal{C}^{1+\alpha,\alpha/2}} \leqslant C$$

yields a sequence $(\tilde{u}_k, \tilde{m}_k) \in \left[\mathcal{C}^{2+\alpha,1+\alpha/2}\right]^2$, which is compactly embedded in X by the Arzela-Ascoli theorem, thus, by possibly passing to a subsequence, a fixed limit point $(u, m) \in X$ is obtained. We can now apply the Leray-Schauder fixed point theorem to conclude that $\mathbf{T}(u, m, 1) = (u, m)$ for some (u, m), which is therefore a solution to the PDE system. $\square$

CHAPTER FOUR

Further Remarks

4.1 Concluding Remarks and Extensions

In both cases, boundary conditions were ignored by working on the flat d-dimensional torus $\mathbb{T}^d$. This is evidently restrictive, and so there is a rich landscape to investigate for those interested in looking at analytical or topological considerations boundaries could provide. In a congestion game model, certainly an impassable wall boundary would affect player strategy, although to some extent one can include extreme penalty costs for unwanted actions such as passing through a restricted area. M. Cirant in [19] looked at a non-time dependant MFG with Neumann conditions at the boundary of a C^2 domain. Also of interest are matters of aborption at a boundary, to model players leaving a physical area, or exiting a market due to lack of resources. An example of a probabalistic treatment can be found in [7], while results for a Cournot Mean Field Game of Controls with absorption are given by Graber and Sircar in a 2021 preprint [25].

Another avenue of possible extension is slotting in local or non-local phenomena in various capacities. The two models in this document were augmentations of this nature, adding a non-local and local coupling feature respectively. Local and non-local changes need not involve the coupling: an interesting non-local change would be to non-localize the spatial diffusion operator with a fractional Laplacian or the generator of a stable Lévy process.

Finally, exploited heavily in Chapter Two is the use of the potential structure of the game: that the Nash Equilibria can be considered as a critical point of some functional. Not all MFG are potential, however, and so a different toolkit entirely would need to be employed in this case, an example of which being a 2020 article by Z. Kobeissi [35] where classical well posedness is proved using some novel assumptions on the behavior of the Hamiltonian that cause a useful contraction mapping with regard to the fixed point condition. Compare this with Bonnans, Hadikhanloo, and Pfeiffer [5], where the potentiality of the game provided satisfaction of the fixed point issue with minimization arguments.

4.2 Further Extensions to Chapter Three

The theorem from [37] that forms the main tool for the a priori estimates in the technique of Chapter Three requires significant assumptions on the regularity of the vector field $b(x, m)$ to obtain immediate Hölder regularity for the value function u. However, one can weaken substantially these assumptions and still produce a number of regularity results about u and m. We suggest an adaptation of the methods previously applied to a traditional mean field game by Gomes, Pimentel, and Sánchez-Morgado in [23].

The suggested relaxation we will make is that $b(x, m)$ is now only Lipschitz and bounded. However, there will be additional slight restrictions made to reach various conclusions, used to extend the techniques of [23] to suit our purposes. Many of the assumptions (convexity, coercivity of the Hamiltonian, various growth conditions) are standard to the field. Under the assumptions of [23], with the added features of

$b(x,m)$, the same methods could likely be applied to obtain a classical solution to the model of Chapter Three.

To allude to the ideas, this theorem could be proved by “bootstrapping” regularity results by alternating between improvements in the Hamilton Jacobi equation (3.1) (i) and the Fokker-Planck equation (3.1) (ii). The target regularity being that both u and m be Hölder continuous of certain parameters, uniformly in the mollification parameter (ε), as do the derivatives of the sequence functions $(u^\varepsilon, m^\varepsilon)$. As the reader can see in [23], unlike the unified potentiality technique of Chapter Two, each bootstrapping step is its own problem, requiring disparate approaches at each state. The basis for applying the techniques of the aforementioned reference are justified, as the mollified system has a solution by Chapter Three, and the limit as $\varepsilon \to 0$ as in [23] could then be investigated. Uniqueness of the solutions proven to exist in Chapter Three could also be investigated via some stronger monotonicity assumptions as in [5].

BIBLIOGRAPHY

[1] Robert J Aumann. Markets with a continuum of traders. *Econometrica: Journal of the Econometric Society*, pages 39–50, 1964.

[2] Alexander Aurell and Boualem Djehiche. Mean-field type modeling of nonlocal crowd aversion in pedestrian crowd dynamics. *SIAM Journal on Control and Optimization*, 56(1):434–455, 2018.

[3] 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.

[4] Alain Bensoussan, Jens Frehse, and Phillip Yam. *Mean field games and mean field type control theory.* Springer, 2013.

[5] J Frédéric Bonnans, Saeed Hadikhanloo, and Laurent Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Applied Mathematics & Optimization*, pages 1–34, 2019.

[6] Joseph Frédéric Bonnans, Justina Gianatti, and Laurent Pfeiffer. A lagrangian approach for aggregative mean field games of controls with mixed and final constraints, 2021.

[7] Luciano Campi and Markus Fischer. N-player games and mean-field games with absorption. *The Annals of Applied Probability*, 28(4):2188 – 2242, 2018.

[8] P. Cardaliaguet, F. Delarue, J.M. Lasry, and P.L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games: (AMS-201).* Annals of Mathematics Studies. Princeton University Press, 2019.

[9] Pierre Cardaliaguet. Weak solutions for first order mean field games with local coupling. In *Analysis and Geometry in Control Theory and its Applications*, volume 11 of *Springer INdAM Series*, pages 111–158. Springer, 2015.

[10] Pierre Cardaliaguet and P. Jameson Graber. Mean field games systems of first order. *ESAIM: COCV*, 21(3):690–722, 2015.

[11] Pierre Cardaliaguet, P. Jameson Graber, Alessio Porretta, and Daniela Tonon. Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1287–1317, 2015.

[12] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12:335–363, 2018.

[13] Pierre Cardaliaguet and Alessio Porretta. An introduction to mean field game theory. In *Mean Field Games*, pages 1–158. Springer, 2020.

[14] R Carmona and F Delarue. Probabilistic Theory of Mean Field Games: vol. I, Mean Field FBSDEs, Control, and Games. *Stochastic Analysis and Applications, Springer Verlag*, 2017.

[15] R Carmona and F Delarue. Probabilistic Theory of Mean Field Games: vol. II, Mean Field Games with Common Noise and Master Equations. *Stochastic Analysis and Applications. Springer Verlag*, 2017.

[16] 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.

[17] Patrick Chan and Ronnie Sircar. Bertrand and Cournot mean field games. *Applied Mathematics & Optimization*, 71(3):533–569, 2015.

[18] Patrick Chan and Ronnie Sircar. Fracking, renewables, and mean field games. *SIAM Review*, 59(3):588–615, 2017.

[19] Marco Cirant. Multi-population mean field games systems with neumann boundary conditions. *Journal de Mathématiques Pures et Appliquées*, 103(5):1294–1315, 2015.

[20] Ronald J DiPerna and Pierre-Louis Lions. Ordinary differential equations, transport theory and sobolev spaces. *Inventiones mathematicae*, 98(3):511–547, 1989.

[21] Franco Flandoli, Maddalena Ghio, and Giulia Livieri. n-player games and mean field games of moderate interactions, 2021.

[22] Diogo A Gomes, Stefania Patrizi, and Vardan Voskanyan. On the existence of classical solutions for stationary extended mean field games. *Nonlinear Analysis: Theory, Methods & Applications*, 99:49–79, 2014.

[23] Diogo A. Gomes, Edgard A. Pimentel, and Héctor Sánchez-Morgado. Time-dependent mean-field games in the subquadratic case. *Communications in Partial Differential Equations*, 40(1):40–76, 2015.

[24] Diogo A Gomes and Vardan K Voskanyan. Extended deterministic mean-field games. *SIAM Journal on Control and Optimization*, 54(2):1030–1055, 2016.

[25] P. Graber and Ronnie Sircar. Master equation for cournot mean field games of control with absorption. 11 2021.

[26] P Jameson Graber and Alain Bensoussan. Existence and uniqueness of solutions for Bertrand and Cournot mean field games. *Applied Mathematics & Optimization*, 77(1):47–71, 2018.

[27] P. Jameson Graber, Vincenzo Ignazio, and Ariel Neufeld. Nonlocal Bertrand and Cournot Mean Field Games with General Nonlinear Demand Schedule. *arXiv preprint*, 2020.

[28] P Jameson Graber and Alpár R Mészáros. Sobolev regularity for first order mean field games. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 35(6):1557–1576, September 2018.

[29] 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):115, 2019.

[30] P Jameson Graber and Charafeddine Mouzouni. Variational mean field games for market competition. In *PDE Models for Multi-Agent Phenomena*, volume 28 of *Springer INdAM Series*, 2018.

[31] Graber, Philip Jameson and Mouzouni, Charafeddine. On mean field games models for exhaustible commodities trade. *ESAIM: COCV*, 26:11, 2020.

[32] 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.

[33] 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.

[34] Hitoshi Ishii. On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions. *Funkcial. Ekvac*, 38(1):101–120, 1995.

[35] Z Kobeissi. On classical solutions to the mean field game system of controls. *arXiv preprint*, 2019.

[36] Z Kobeissi. Mean field games with monotonous interactions through the law of states and controls of the agents. *arXiv preprint*, 2020.

[37] Ol'ga Aleksandrovna Ladyženskaja, Vsevolod Alekseevich Solonnikov, and Nina N. Ural'ceva. *Linear and Quasi-linear Equations of Parabolic Type*, volume 23 of *Translations of mathematical monographs.* American Mathematical Society, 1968.

[38] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[39] Pierre-Louis Lions. Théorie des jeux de champ moyen et applications (mean field games). *Cours du College de France*, 2007-2012.