Classical and weak solutions to local first-order mean field games through elliptic regularity

Sebastian Muñoz

Abstract. We study the regularity and well-posedness of the local, first-order forward–backward mean field games system, assuming a polynomially growing cost function and a Hamiltonian of quadratic growth. We consider systems and terminal data that are strictly monotone in density and study two different regimes depending on whether there exists a lower bound for the running cost function. The work relies on a transformation due to P.-L. Lions, which gives rise to an elliptic partial differential equation with oblique boundary conditions, that is strictly elliptic when the coupling is unbounded from below. In this case, we prove that the solution is smooth. When the problem is degenerate elliptic, we obtain existence and uniqueness of weak solutions analogous to those obtained by P. Cardaliaguet and P. J. Graber for the case of a terminal condition that is independent of the density. The weak solutions are shown to arise as viscous limits of classical solutions to strictly elliptic problems.

1. Introduction

The purpose of this paper is to study the well-posedness of the first-order mean field games system (MFG for short) with a local coupling:

$$\begin{cases} -u_t + H(x, D_x u) = f(x, m(x,t)), & (x,t) \in Q_T = \mathbb{T}^d \times (0,T), \\ m_t - \operatorname{div}(m D_p H(x, D_x u)) = 0, & (x,t) \in Q_T, \\ m(x,0) = m_0(x), \ u(x,T) = g(x, m(x,T)), & x \in \mathbb{T}^d, \end{cases} \quad \text{(MFG)}$$

where $H: \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is a strictly convex Hamiltonian of quadratic growth, $f, g: \mathbb{T}^d \times [0,\infty) \to [-\infty, \infty)$ are strictly increasing in their second variable m, f has polynomial growth in m, and m_0 is a strictly positive probability density. As is standard, we work on the flat d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ to avoid additional technicalities with spatial boundary conditions.

MFG were introduced by Lasry and Lions ([14, 18]), and at the same time, in a particular setting, by Huang, Malhamé, and Caines ([13]). They are non-cooperative differential games with infinitely many players, in which the players find an optimal strategy by

observing the distribution of the others. When the game is completely deterministic, such games are typically modeled by system (MFG), which has been successfully studied in the case where the function g is independent of the density m, and a complete theory of weak solutions has been obtained through variational methods by Cardaliaguet, Graber, Porretta, and Tonon ([2–4, 12]).

Our two main contributions are proving well-posedness when the terminal condition is strictly increasing with respect to m, and the attainment of classical solutions under the additional assumption that $f(\cdot, 0) \equiv -\infty$. When the latter blowup condition does not hold, we obtain weak solutions that are in line with the variational theory, and they are shown to enjoy higher regularity than in the case $g_m \equiv 0$, by virtue of the strict monotonicity of g.

The precise statements of our main results are as follows. We refer to Section 2 for the exact assumptions (M), (H), (F), (G), (SE), and (DE), to Section 5 for the definition of a weak solution, and to the notation subsection for the meaning of the function spaces mentioned in the theorems below.

Theorem 1.1. *Let $0 < \alpha < 1$, and assume that* (M), (H), (F), (G), *and* (SE) *hold. Then there exists a unique classical solution* $(u, m) \in C^{3,\alpha}(\overline{Q_T}) \times C^{2,\alpha}(\overline{Q_T})$ *to* (MFG).

Theorem 1.2. *Assume that* (M), (H), (F), (G), *and* (DE) *hold. Then the following is true:*

(i)   *There exists a weak solution* $(u, m) \in (\mathrm{BV}(Q_T) \cap L^\infty(Q_T)) \times (C([0, T], H^{-1}(\mathbb{T}^d)) \cap L^\infty(Q_T))$ *to* (MFG).

(ii)   *The solution (u, m) is the almost everywhere (a.e. for short) limit, as $\epsilon \to 0$, of solutions $(u^\epsilon, m^\epsilon) \in C^{3,\alpha}(\overline{Q_T}) \times C^{2,\alpha}(\overline{Q_T})$ to MFG systems satisfying* (SE). *Furthermore, $(u^\epsilon(\cdot, T), m^\epsilon(\cdot, T)) \to (u(\cdot, T), m(\cdot, T))$ a.e. in $\mathbb{T}^d$.*

(iii)   *If (u', m') is another weak solution to* (MFG), *then $m = m'$ a.e. in Q_T, and $u = u'$ a.e. in $\{m > 0\}$. Moreover, $m(\cdot, T) = m'(\cdot, T)$, $u(\cdot, T) = u'(\cdot, T)$, and $u(\cdot, 0) = u'(\cdot, 0)$ a.e. in $\mathbb{T}^d$.*

Despite the connections with the variational theory, we do not use variational methods. Instead, we follow the ideas of Lions and his work on the so-called planning problem, where the initial and terminal densities $m(\cdot, 0)$ and $m(\cdot, T)$ are prescribed ([14, 18]). It was first observed by Lions that if, for each fixed $x \in \mathbb{T}^d$, $f^{-1}(x, \cdot)$ is the inverse function of $f(x, \cdot)$, it is possible to formally eliminate the variable m from the system. This transforms the problem into a second-order quasilinear elliptic equation with a non-linear oblique boundary condition which, in the special case where $D_x H, D_x f, D_x g \equiv 0$, may be written as follows (see Section 2 for the general setting):

$$\begin{cases} -u_{tt} - \mathrm{Tr}((D_p H \otimes D_p H + \chi(-u_t + H)D_{pp}^2 H)D_{xx}^2 u) \\ \quad + 2D_p H \cdot D_x u_t = 0 & \text{in } Q_T, \\ -u_t + H - f(m_0) = 0 & \text{on } \mathbb{T}^d \times \{t = 0\}, \\ -g(f^{-1}(-u_t + H)) + u = 0 & \text{on } \mathbb{T}^d \times \{t = T\}, \end{cases} \quad (1.1)$$

where the function $\chi(w)$ is defined by

$$\chi(w) = f^{-1}(w)\, f_m(f^{-1}(w)).$$

We emphasize the fact that, while (1.1) is an elliptic second-order problem, the original system (MFG) is of first order and, in particular, it models a game with no diffusion.

Our approach to obtain classical solutions when (1.1) is strictly elliptic was developed by Lions, who applied, in his lectures at Collège de France, the following strategy for finding regular solutions to the planning problem. Viewed as a quasilinear elliptic equation with a non-linear boundary condition, the problem can be tackled with classical methods from the field of a priori estimates: specifically, maximum principle techniques and the Bernstein method to obtain bounds on the solution and its gradient, the application of classical estimates to bound the Hölder norm of the gradient up to the boundary, and soft functional analytic tools to attain the classical solutions.

In order to study the general MFG system, even when (1.1) happens to be degenerate elliptic and the solutions are expected to be discontinuous, our strategy is to first obtain smooth solutions in the strictly elliptic case, and to subsequently find the weak solution as a viscous limit of strictly elliptic problems. The success of this approach is based on the fact that, once smooth solutions are known to exist, every a priori estimate that is independent of the ellipticity constant can be used as a source of compactness and regularity for the limit. Our a priori estimates are supplemented by energy computations based on the Lasry–Lions monotonicity procedure, which is the canonical method for obtaining integral bounds and proving uniqueness in MFG systems.

To identify and motivate the condition that determines the strict or degenerate ellipticity of the system, we remark that the determinant corresponding to the elliptic equation in (1.1) becomes zero precisely as $\chi = m f_m \to 0$. This is in accordance with the heuristic principle that the regularity of u is lost in regions where there are few to no players (no information), as well as when the cost fails to be strictly monotone (concentration blowup). Because, as is standard, f is assumed to grow at least logarithmically as $m \to \infty$, this degeneracy can only happen as $m \to 0$. In the absence of diffusion, for the strict positivity of m to be preserved, we expect to have a very strong incentive for the players to navigate through regions of low density. With these considerations in place, we will classify system (MFG) as being strictly elliptic precisely when f has a singularity at $m = 0$, and as degenerate elliptic otherwise.

It should be noted that, for the stationary problem, classical solutions were obtained in [6] for the case where $f = \log m$, and in [10] for the case where $H(x, p) = \frac{1}{2}|p|^2 - V(x)$, under a small-oscillation assumption. For second-order systems with a (possibly) degenerate diffusion and a density-independent terminal condition, the variational theory was extended in [4], where it was shown (compare with Theorem 1.2) that the weak solutions to the first-order problem arise as viscous limits of weak solutions to second-order MFG systems. Finally, the most general result for weak solutions to the second-order problem is due to Porretta ([21]), and, unlike [4], it does not use variational methods.

The content and structure of the paper are described as follows. Section 2 explains the general setting and assumptions that will be used, followed by the statements of the preliminary results from the classical literature on quasilinear elliptic equations and oblique derivative problems that will be used to prove existence of classical solutions.

In Section 3 we obtain all the necessary a priori estimates for the strictly elliptic problem. The main results, which deal with the system in full generality, are summarized in Theorem 3.9. We also establish a minor variant, in the special case when the x dependence has a simple structure, that is, when $H(x, p) - f(x, m) \equiv H(p) - f(m) - V(x)$, in Theorem 3.10. This result states that, with this structural assumption, it is not necessary to require f to grow at most polynomially in m, allowing for examples such as $f(m) = e^m + \log m$. Section 3.1 contains the L^∞-bounds on the solution u, as well as two-sided bounds for the terminal density $m(\cdot, T)$, obtained through maximum principle methods. These methods exploit the fact that the strict monotonicity property of g is equivalent to the linearized version of problem (1.1) having an oblique boundary condition, which is of Robin type in the upper component of ∂Q_T. In Section 3.2, the gradient bound is obtained by means of the Bernstein method. To deal with the asymmetry between the space and time derivatives in (1.1), it is necessary to first get a precise bound for u_t in terms of the space gradient, utilizing the a priori lower bound on $m(\cdot, T)$ and the maximum principle. This, in turn, provides a "conditional" a priori lower bound for m, namely a lower bound that holds exclusively at points (x, t) where the function $H(x, D_x u)$ is close to its maximum value. The conditional nature of this bound, as well as the structure of (1.1) in its fully general form, requires a non-conventional choice of an auxiliary function of the space-time gradient.

Section 4 deals with the existence of classical solutions for the strictly elliptic problem, including the proof of Theorem 1.1. The corresponding variant for the case of a fast-growing f is presented in Theorem 4.3. It is first explained how a classical result from the theory of oblique derivative problems, due to G. M. Lieberman ([17]), immediately yields an a priori Hölder estimate for Du up to the boundary in terms of the L^∞-bounds on u and Du. Existence is then proved through an application of the non-linear method of continuity, the classical Schauder estimates for the linear oblique derivative problem, and a variant of a convergence theorem of R. Fiorenza ([7, 8, 16]).

In Section 5 we develop the weak theory for the degenerate elliptic problem, and obtain the proof of Theorem 1.2. It is first established that, for strictly elliptic problems, there exists an upper bound for the density that is independent of any lower bounds on $m(\cdot, T)$. After deriving some necessary energy estimates and defining an ϵ-perturbation of the coupling f that makes the problem strictly elliptic, the solution is obtained as the limit when $\epsilon \to 0$ of the corresponding smooth solutions. It is also proved, in Theorem 5.5, that when the data is independent of the space variable, the value function u and the terminal density $m(\cdot, T)$ are globally Lipschitz continuous.

Remark 1.3. We mention here some related work that was released after this paper. In [20], the author showed existence of classical solutions for the so-called extended MFG, a generalization of (MFG) introduced by Lions and Souganidis ([19]), having a fully general

continuity equation, and a non-separated Hamiltonian, namely $H = H(x, p, m)$, with arbitrary superlinear growth. In particular, classical solutions were obtained for first-order MFG with congestion. As for weak solutions to (MFG), the most general result to date was obtained by Cardaliaguet and Porretta ([5]), where the solution is obtained as a vanishing viscosity limit to the weak solutions from [21].

Notation

Let $n, k \in \mathbb{N}$. Given $x, y \in \mathbb{R}^n$, x and y will always be understood to be row vectors, and their scalar product xy^T will be denoted by $x \cdot y$. For any bounded set Ω, with $\Omega \subset Q_T$, $\Omega \subset \mathbb{T}^d$, or $\Omega \subset [0, T]$, and $0 \le \alpha < 1$, $C^{k,\alpha}(\Omega)$, refers to the space of k times differentiable real-valued functions with α-Hölder continuous kth order derivatives, and, for $u \in C^{0,\alpha}(\Omega)$, the Hölder semi-norm of u will be denoted by $[u]_{\alpha,\Omega}$. Similarly, if $H^{-1}(\mathbb{T}^d)$ denotes the dual space of the Sobolev space $H^1(\mathbb{T}^d)$, the space of $H^{-1}(\mathbb{T}^d)$-valued α-Hölder continuous functions $C^{0,\alpha}([0, T]; H^{-1}(\mathbb{T}^d))$ is equipped with the Hölder semi-norm $[\cdot]_{\alpha,[0,T],H^{-1}}$. For functions $\Phi(x, t, z, p, s) \in C^0(Q_T \times \mathbb{R} \times \mathbb{R}^{d+1})$, where typically $(x, t, z, p, s) = (x, t, u(x, t), D_x u, u_t)$, the conventions $\bar{x} \equiv (x, t)$ and $q \equiv (p, s)$ will always be in place. The notation Du, $D\Phi$ will always refer to the full gradient in all variables so that, for instance, $Du = D_{\bar{x}} u = (D_x u, u_t)$ and $D\Phi = (D_{\bar{x}}\Phi, \Phi_z, D_q\Phi)$. For $(x, t) \in \partial Q_T$, $\nu(x, t) = \pm(0, 0, \ldots, 1)$ denotes the outward-pointing unit normal vector. We write $C = C(K_1, K_2, \ldots, K_M)$ for a positive constant C depending monotonically on the non-negative quantities $K_1, \ldots, K_M$. We also define, for $K > 0$, and any set V, $V_K = \{(y, z, q) \in V \times \mathbb{R} \times \mathbb{R}^{d+1} : |z| + |q| \le K\}$. We write $C^k(\overline{Q_T})^*$ for the dual space of $C^k(\overline{Q_T})$. In particular, $C^0(\overline{Q_T})^*$ is the space of finite signed Borel measures on $\overline{Q_T}$, and $C^\infty(\overline{Q_T})^*$ is the space of distributions. Moreover, $\text{BV}(Q_T)$ is the space of functions of bounded variation, that is, the space of $L^1(Q_T)$ functions such that their distributional derivatives are elements of $C^0(\overline{Q_T})^*$, and $L^\infty_+(Q_T)$ consists of the functions $m \in L^\infty(Q_T)$ such that $m \ge 0$ almost everywhere (a.e. for short) in Q_T. Finally, for $m \in L^\infty_+(Q_T)$, $L^2_m(Q_T)$ consists of the functions v such that $|v|^2 m \in L^1(Q_T)$.

2. Assumptions and general setting

2.1. The MFG system as an elliptic problem

We now present the general elliptic formulation of the MFG system. As explained in the previous section, it is an equivalent problem satisfied by u, whenever the pair $(u, m) = (u, f^{-1}(\cdot, -u_t + H(\cdot, D_x u))) \in C^2(\overline{Q_T}) \times C^1(\overline{Q_T})$ is a classical solution to (MFG). It is obtained after eliminating m from the system, and it consists of a quasilinear elliptic equation with a non-linear oblique boundary condition,

$$\begin{cases} Qu = -\text{Tr}(A(x, Du)D^2u) + b(x, Du) = 0 & \text{in } Q_T, \\ Nu = B(x, t, u, Du) = 0 & \text{on } \partial Q_T, \end{cases} \tag{Q0}$$

where, for all $(x, t, z, p, s) \in \overline{Q_T} \times \mathbb{R} \times \mathbb{R}^{d+1}$,

$$A(x, p, s) = (D_p H, -1) \otimes (D_p H, -1)$$
$$+ \chi(x, -s + H(x, p)) \begin{pmatrix} D_{pp}^2 H(x, p) & 0 \\ 0 & 0 \end{pmatrix}, \tag{Q1}$$

$$b(x, p, s) = -D_x H(x, p) \cdot D_p H(x, p)$$
$$+ D_x f(x, f^{-1}(x, -s + H(x, p))) \cdot D_p H(x, p) \tag{Q2}$$
$$- \chi(x, -s + H(x, p)) \operatorname{Tr}(D_{xp}^2 H(x, p)),$$

$$B(x, 0, z, p, s) = -s + H(x, p) - f(x, m_0(x)),$$
$$B(x, T, z, p, s) = -g(x, f^{-1}(x, -s + H(x, p))) + z, \tag{B1}$$

with the function $\chi(x, w)$ being defined by

$$\chi(x, w) = f^{-1}(x, w) f_m(x, f^{-1}(x, w)).$$

We remark that the matrix A is clearly non-negative, and since $\det(A) = \chi^d \det D_{pp}^2 H$, the condition for degeneracy is $\chi = m f_m = 0$. For future use, we set

$$h(x, w) = \sqrt{\chi(x, w)}.$$

2.2. Assumptions

We now state the main assumptions (M), (H), (F), (G), and (E), that will be in place throughout the paper, except when explicitly stated. Assumption (E) on the ellipticity of the system contains the mutually exclusive possibilities (SE) and (DE), and it will always be made clear which of the two is in place. For the theory of weak solutions, the differentiability assumptions on the data can naturally be weakened through standard approximation arguments, but in the interest of clarity such matters will not be considered. Throughout the assumptions, the quantities $C_0 > 0$ and $0 \leq \tau < 1$ are fixed constants.

(M) (Assumptions on m_0) The initial density m_0 satisfies

$$m_0 \in C^4(\mathbb{T}^d), \quad m_0 > 0, \quad \text{and} \quad \int_{\mathbb{T}^d} m_0 = 1. \tag{M1}$$

(H) (Assumptions on H) The functions H, $D_p H$, $D_{pp}^2 H$ are four times continuously differentiable, and the following quadratic growth and uniform convexity conditions hold:

$$\frac{1}{C_0} I \leq D_{pp}^2 H(x, p) \leq C_0 I, \tag{H1}$$

$$D_p H(x, p) \cdot p \geq 2H(x, p) - C_0, \tag{H2}$$

$$|D_{ppp}^3 H(x, p)| \leq C_0(1 + |p|)^{-1}, \tag{H3}$$

for all $(x, p) \in \mathbb{T}^d \times \mathbb{R}^d$. The space oscillation of H is at most subquadratic in p, namely

$$|D^3_{xxp} H| \leq C(1 + |p|)^\tau, \quad |D^3_{xpp} H| \leq C_0(1 + |p|)^{\tau-1}. \tag{HX}$$

(F) (Assumptions on f) The continuous function $f: \mathbb{T}^d \times [0, \infty) \to [-\infty, \infty)$ is four times continuously differentiable on $\mathbb{T}^d \times (0, \infty)$ and strictly increasing in the second variable, with $f_m > 0$. The function f grows polynomially as $m \to \infty$, in the sense that its growth is at least of degree zero, namely

$$\liminf_{x \in \mathbb{T}^d, m \to \infty} m f_m(x, m) > 0, \tag{F1}$$

and its derivative f_m satisfies a polynomial bound $|m f_{mm}| \leq C_0 f_m$, which can be equivalently expressed in terms of $\chi(x, w)$ as

$$|\chi_w| \leq C_0. \tag{F2}$$

The space derivative of f satisfies the same polynomial bound,

$$|m(D_x f)_m| \leq C_0|D_x f|, \tag{FX1}$$

as well as the control

$$|D_x f|, \ |D^2_{xx} f| \leq C_0(1 + |f|^{\tau/2} + |m f_m|^{(1+\tau)/2}). \tag{FX2}$$

(G) (Assumptions on g) The continuous function $g: \mathbb{T}^d \times [0, \infty) \to [-\infty, \infty)$ is four times continuously differentiable on $\mathbb{T}^d \times (0, \infty)$ and strictly increasing in the second variable, with $g_m > 0$. The control required for its space oscillation is that, for each $x \in \mathbb{T}^d$,

$$\lim_{m \to \infty} g(x, m) = \sup_{\mathbb{T}^d \times [0, \infty)} g \quad \text{and} \quad g(x, 0) = \inf_{\mathbb{T}^d \times [0, \infty)} g. \tag{GX}$$

(E) (Ellipticity of the system) One of the following conditions holds:

system (MFG) is *strictly elliptic*, in the sense that $f(\cdot, 0) \equiv -\infty$, (SE)

or

system (MFG) is *degenerate elliptic*, in the sense that $f(\cdot, 0) > -\infty$. (DE)

In the case of (DE), since the density is not expected to be strictly positive, we assume that $g(\cdot, 0) > -\infty$.

A few comments should be made about the assumptions on the spatial oscillation. First, we remark that the subquadratic growth assumption (HX) can be interpreted as requiring that the purely quadratic part of H is independent of x. Condition (FX2), on the other hand, can be interpreted as being dual to (HX). Indeed, heuristically, since f is assumed to have

polynomial growth, $mf_m \approx f$, and $f = -u_t + H \approx |p|^2$, so both conditions impose the same polynomial growth bound in the variable $|p|$. We consider now the assumption (GX) on the x-oscillation of g. When g is bounded, the first (resp. second) condition in (GX) corresponds to a purely qualitative control on $|D_x g|$ that becomes stricter as $m \to \infty$ (resp. $m \to 0^+$). From the modeling point of view, it can be interpreted as the requirement that extremely crowded regions (resp. nearly empty regions) have roughly the same terminal value for the players.

Remark 2.1. For simplicity of the presentation, we observe that, up to increasing the value of C_0, the following inequalities are trivial consequences of (H1), (HX), and (SE), and they will be used freely when pertinent:

$$\frac{1}{C_0}|p|^2 - C_0 \le H(x, p) \le C_0|p|^2 + C_0, \quad |D_p H(x, p)| \le C_0(1 + |p|), \qquad (2.1)$$

$$|D_x H(x, p)| \le C_0(1 + |p|^{1+\tau}), \quad |D_{xx}^2 H(x, p)| \le C_0(1 + |p|^{1+\tau}), \qquad (2.2)$$

$$|D_{xp}^2 H(x, p)| \le C_0(1 + |p|)^\tau, \qquad (2.3)$$

$$\|\chi(\cdot, 0)\|_{C^0(\mathbb{T}^d)} + \|m_0\|_{C^1(\mathbb{T}^d)} + \|f\|_{C^2(\mathbb{T}^d \times [\min m_0, \max m_0])} \le C_0. \qquad (2.4)$$

2.3. Preliminary results

This subsection includes the classical results that will be required in Section 4 to obtain the higher regularity from a priori C^1 bounds. In this subsection only, it will not be assumed that problem (Q0) is explicitly given by (Q1), (Q2), and (B1), but instead (Q, N) will be a general pair of an elliptic quasilinear operator and a fully non-linear boundary operator. In particular, A and b will not necessarily be assumed to be independent of t and u. The first theorem is the classical interior Hölder gradient estimate for quasilinear equations, due to O. Ladyzhenskaya and N. Uraltseva ([17, Lem. 2.1]).

Theorem 2.2. Let $u \in C^2(Q_T)$ satisfy $Qu = 0$ in Q_T, with $A(x, t, z, q) \in C^1(Q_T \times \mathbb{R} \times \mathbb{R}^{d+1})$, $b(x, t, z, q) \in C^0(Q_T \times \mathbb{R} \times \mathbb{R}^{d+1})$. Suppose that $\|u\|_{C^1(Q_T)} \le K$, and that the constants λ_K, μ_K satisfy, in $Q_{T,K}$,

$$A \ge \lambda_K I \quad and \quad \mu_K \ge |A| + |DA| + |b|. \qquad (2.5)$$

Then, for any $V \subset\subset Q_T$, there exist constants $C = C(K, \mu_K/\lambda_K, \mathrm{dist}(V, \partial Q_T)^{-1})$ and $\gamma = \gamma(K, \mu_K/\lambda_K)$, such that

$$[Du]_{\gamma, V} \le C.$$

Next is the following local boundary Hölder estimate for the gradient in oblique problems, due to Lieberman ([17, Lem. 2.3]). In Theorem 2.3, the following definitions are in place:

$$B = \{(x, t) \in \mathbb{R}^{d+1} : |(x, t)| < 1\}, \quad B^+ = \{(x, t) \in B : t > 0\},$$
$$B^0 = \{(x, t) \in B : t = 0\}, \qquad\qquad B' = \{(x, t) \in B^+ : |(x, t)| < \tfrac{1}{3}\}.$$

Theorem 2.3. *Let* $u \in C^2(B^+ \cup B^0)$ *solve* $Qu = 0$ *in* B^+, $Nu = 0$ *on* B^0, *with* $A(x,t,z,q) \in C^1(\overline{B^+} \times \mathbb{R} \times \mathbb{R}^{d+1})$, $b(x,t,z,q) \in C^0(\overline{B^+} \times \mathbb{R} \times \mathbb{R}^{d+1})$, $B(x,t,z,q) \in C^1(B^0 \times \mathbb{R} \times \mathbb{R}^{d+1})$, $D_q B(x,t,z,q) \in C^1(B^0 \times \mathbb{R} \times \mathbb{R}^{d+1})$. *Assume furthermore that* (2.5) *holds in* $\overline{B_K^+}$, *as well as, on* B_K^0,

$$\lambda_K \leq -B_s,$$
$$\mu_K \geq |D_q B| + |D_z B| + |D_{\bar{x}} B| + |D_{qq}^2 B| + |D_{qz}^2 B| + |D_{q\bar{x}}^2 B|. \tag{2.6}$$

Then there are constants C *and* γ *depending only on* K *and* μ_K / λ_K *such that, if* $\|u\|_{C^1(B^+ \cup B^0)} \leq K$, *then*

$$[Du]_{\gamma, B'} \leq C.$$

For the next theorem, which is the basic Schauder estimate for linear oblique problems ([9, Thm. 6.30]), we recall that $\nu(x,t) = \pm(0,0,\ldots,1)$ denotes the outward-pointing normal vector at $(x,t) \in \partial Q_T$.

Theorem 2.4. *Assume that* $u \in C^2(\overline{Q_T})$ *solves the linear problem*

$$-\operatorname{Tr}(\tilde{A}(x,t)D^2 u) = \eta_1(x,t) \text{ in } Q_T, \quad \tilde{B}(x,t) \cdot Du = \eta_2(x,t) \text{ on } \partial Q_T,$$

where

$$\tilde{A}^{ij}, \eta_1 \in C^{0,\alpha}(\overline{Q_T}), \quad \tilde{B}, \eta_2 \in C^{1,\alpha}(\partial Q_T), \quad \tilde{A} \geq \lambda I, \quad \text{and} \quad \tilde{B} \cdot \nu \geq \lambda_0.$$

Then there exists $C = C(\frac{1}{\lambda}, \frac{1}{\lambda_0}, \|\tilde{A}^{ij}\|_{C^{0,\alpha}(\overline{Q_T})}, \|\tilde{B}\|_{C^{1,\alpha}(\partial Q_T)})$ *such that*

$$\|u\|_{C^{2,\alpha}(\overline{Q_T})} \leq C(\|u\|_{C^0(\overline{Q_T})} + \|\eta_1\|_{C^{0,\alpha}(\overline{Q_T})} + \|\eta_2\|_{C^{1,\alpha}(\partial Q_T)}).$$

The last result of this subsection is a variant of a convergence theorem of Fiorenza, which is a basic tool for using the method of continuity without the need of a priori second-derivative estimates ([16, Lem. 2, Cor. 1]).

Theorem 2.5. *Let* $0 < \alpha, \gamma < 1$. *For each* $n \in \mathbb{N}$, *let* $u_n \in C^{2,\alpha}(\overline{Q_T})$ *be a sequence of solutions to the quasilinear problems* $Q_n u = 0$, $N_n u = 0$, *where, for* C, K, γ, λ, λ_0 *independent of* n,

$$Q_n u = -\operatorname{Tr}(A_n(x,t,u,Du)D^2 u) + b_n(x,t,u,Du), \quad N_n u = B_n(x,t,u,Du),$$
$$\|A_n\|_{C^1(\overline{Q}_{T,K})} + \|b_n\|_{C^1(\overline{Q}_{T,K})} + \|B_n\|_{C^2(\overline{Q}_{T,K})} + \|D_q B_n\|_{C^2(\overline{Q}_{T,K})} \leq C,$$
$$A_n \geq \lambda I \text{ in } \overline{Q}_{T,K} \quad \text{and} \quad D_q B_n \cdot \nu \geq \lambda_0 \text{ in } \partial \overline{Q}_{T,K},$$
$$\|u_n\|_{C^{1+\gamma}(\overline{Q_T})} \leq K,$$

with $u_n \to u$ *uniformly, and* $(A_n, b_n, B_n) \to (A, b, B)$ *uniformly on* $\overline{Q}_{T,K}$. *Then* $u_n \to u$ *in* $C^{2,\alpha}(\overline{Q_T})$, *and* u *solves* (Q0).

3. A priori estimates

In this section we establish a priori estimates for the solution and the gradient, in the case where (MFG) is strictly elliptic. To account for the fact that the functions f and g depend on the space variable, we will make extensive use of the continuous, strictly increasing functions $f_0, g_0, f_1, g_1 \colon (0, \infty) \to \mathbb{R}$ defined by

$$f_0(m) = \min_{\mathbb{T}^d} f(\cdot, m), \quad g_0(m) = \min_{\mathbb{T}^d} g(\cdot, m),$$

$$f_1(m) = \max_{\mathbb{T}^d} f(\cdot, m), \quad g_1(m) = \max_{\mathbb{T}^d} g(\cdot, m).$$

3.1. Estimates for the solution and the terminal density

We first obtain a priori bounds for the C^0 norm of the solution u. As a corollary, positive, two-sided bounds for the terminal density are established.

Lemma 3.1. *Assume that* (SE) *holds. Then, there exists a constant* $C = C(C_0)$ *such that for any solution* $(u, m) \in C^2(\overline{Q_T}) \times C^1(\overline{Q_T})$ *of* (MFG)*, and every* $(x, t) \in \overline{Q_T}$,

$$g_0 f_1^{-1}(-C) - C(e^{CT} - e^{Ct}) \le u(x, t) \le g_1 f_0^{-1}(C) + C(e^{CT} - e^{Ct}). \tag{3.1}$$

Proof. The goal here is to modify u into a function that necessarily achieves its maximum at $\{t = T\}$, which is the region of the boundary where, by the strict monotonicity of g, the boundary condition of (Q0) provides information about u. This requires some estimates for the terms in (Q2). By (2.3) and (F2),

$$|\chi(x, f) \operatorname{Tr}(D_{xp}^2 H(x, D_x u))| \le C(1 + |f|)(1 + |D_x u|^\tau). \tag{3.2}$$

Moreover, by (FX2),

$$|D_x f(x, m(x, t)) \cdot D_p H(x, D_x u)| \le C(1 + |f|^{(1+\tau)/2})(1 + |D_x u|). \tag{3.3}$$

Now, given u, define the linear, uniformly elliptic operator Q_u by

$$Q_u v = -\operatorname{Tr}(A(x, Du) D^2 v).$$

Notice that $Q_u u = -b(x, Du)$. Let $\zeta \in C^2([0, T])$ be a function to be chosen later, and define

$$v = u + \zeta(t),$$

so that $v_t = u_t + \zeta'(t)$ and $D_x v = D_x u$. This yields, by (2.1), (2.2), (3.2), and (3.3),

$$\begin{aligned}
Q_u v = {}&-\zeta''(t) + D_x H(x, D_x v) \cdot D_p H(x, D_x v) - D_x f(x, m) \cdot D_p H(x, D_x v) \\
&+ \chi \operatorname{Tr}(D_{xp}^2 H(x, D_x v)) \\
\le {}&-\zeta''(t) + C(1 + |D_x v|^{1+\tau})(1 + |D_x v|) \\
&+ C(1 + |-v_t + H(x, D_x v) + \zeta'(t)|^{(1+\tau)/2})(1 + |D_x v|) \\
&+ C(1 + |-v_t + H(x, D_x v) + \zeta'(t)|)(1 + |D_x v|^\tau) \\
\le {}&-\zeta''(t) + C(1 + |D_x v|^3 + |v_t|^2) + C(1 + |D_x v|)|\zeta'(t)|,
\end{aligned}$$

where the constant C increases in each line. Now set $C_1 = 2C$ and fix C_1, still allowing C to increase at each step. We choose $\zeta(t) = \frac{k}{2C_1}(e^{2C_1 t} - e^{2C_1 T})$, where $k > 0$ is a parameter. Then,

$$\zeta'(t) = ke^{2C_1 t}, \quad \zeta''(t) = 2C_1|\zeta'(t)|,$$

and, consequently, at any interior maximum point (x, t) of v,

$$0 \leq Q_u v \leq -\zeta''(t) + C_1(1 + |\zeta'(t)|) \leq -C_1\zeta'(t) + C_1 = -C_1 ke^{2C_1 t} + C_1$$
$$\leq -C_1 k + C_1,$$

which can only hold if $k \leq 1$. Thus, if one chooses $k > 1$, v necessarily achieves its maximum value when $t = 0$ or $t = T$. If this happens at a point (x, t) where $t = 0$, then $u_t + \zeta' = v_t \leq 0$, $D_x u = D_x v = 0$. Therefore,

$$-\|H(\cdot, 0)\|_{C^0(\mathbb{T}^d)} \leq -v_t + H(x, 0)$$
$$= -u_t + H(x, D_x u) - \zeta'(0) = f(x, m_0(x, t)) - \zeta'(0),$$

implying that

$$k = \zeta'(0) \leq f(x, m_0(x, t)) + \|H(\cdot, 0)\|_{C^0(\mathbb{T}^d)}.$$

Hence, taking $k > \max_{x \in \mathbb{T}^d} f(x, m_0(x)) + \|H(\cdot, 0)\|_{C^0(\mathbb{T}^d)}$, it follows that v attains its maximum value at $t = T$. At this point, $u_t + \zeta'(t) = v_t \geq 0$, $D_x u = D_x v = 0$, and, as before,

$$\|H(x, 0)\|_{C^0} \geq -v_t + H(x, 0) = f(x, m(x, T)) - \zeta'(T),$$

which gives

$$f_0(m(x, T)) \leq f(x, m(x, T)) \leq \zeta'(T) + \|H(\cdot, 0)\|_{C^0(\mathbb{T}^d)}$$
$$\leq ke^{2C_1 T} + \|H(\cdot, 0)\|_{C^0(\mathbb{T}^d)} \leq C.$$

Thus, since $u(x, T) = v(x, T)$, taking into account the surjectivity of $f(x, \cdot)$,

$$\max v = v(x, T) = g(x, m(x, T)) \leq g(x, f_0^{-1}(C)) \leq g_1(f_0^{-1}(C)).$$

Finally, for arbitrary $(x, t) \in Q_T$,

$$u(x, t) = v(x, t) - \zeta(t) \leq g_1(f_0^{-1}(C)) + C(e^{CT} - e^{Ct}).$$

The lower estimate follows from a completely symmetrical argument. ∎

Corollary 3.2. *Assume* (SE)*, and let C be the constant from Lemma* 3.1*. Then, for every $x \in \mathbb{T}^d$,*

$$g_1^{-1} g_0 f_1^{-1}(-C) \leq m(x, T) \leq g_0^{-1} g_1 f_0^{-1}(C). \tag{3.4}$$

Proof. From the first inequality in (3.1), for each $x \in \mathbb{T}^d$,

$$g_0 f_1^{-1}(-C) \leq g(x, m(x, T)),$$

and thus, by definition of g_1,

$$g_0 f_1^{-1}(-C) \leq g_1(m(x,T)). \tag{3.5}$$

Observe that the application of g_1^{-1} on both sides of (3.5) is possible because, by (GX), the functions g_0 and g_1 have the same range. This yields the first inequality in (3.4). The second inequality is obtained through the same reasoning. ∎

Remark 3.3. A minor modification of the proof of Lemma 3.1 shows that, when H, f, and g are independent of x, the following sharper estimates hold:

$$g(\min m_0) + (f(\min m_0) - H(0))(T - t) \leq u(x,t)$$
$$\leq g(\max m_0) + (f(\max m_0) - H(0))(T - t),$$
$$\min m_0 \leq m(x,T) \leq \max m_0.$$

3.2. Estimates for the space-time gradient

Given the operator Q from (Q0), we recall that its linearization at $u \in C^2(\overline{Q_T})$ is the linear, uniformly elliptic operator

$$L_u(v) = -\operatorname{Tr}(A(x, Du)D^2 v) - D_q \operatorname{Tr}(A(x, Du)D^2 u) \cdot Dv + D_q b(x, Du) \cdot Dv. \tag{3.6}$$

The gradient estimate will be obtained through Bernstein's method. Specifically, we will bound $\|Du\|_{C^0(\overline{Q_T})}$ by evaluating the linearization $L_u(v)$ at appropriately chosen functions $v(x,t) = \Phi(x,t,u,Du)$, where $\Phi(x,t,z,q)$ is convex in q, exploiting the fact that, roughly speaking, convex functions of the gradient are expected to be subsolutions. For this purpose, we first obtain an explicit form for the terms in (3.6), as well as a general expression for the linearization applied to such functions v.

Lemma 3.4. *Let* $\Phi(x,t,p,s) \in C^2(\overline{Q_T} \times \mathbb{R}^{d+1})$, *assume that* $u \in C^3(\overline{Q_T})$ *solves* (Q0), *and set* $v(x,t) = \Phi(x,t,Du(x,t))$. *Then, for each* $q = (p,s) \in \mathbb{R}^{d+1}$, *and for each* $\bar{x} = (x,t) \in Q_T$,

$$-D_q \operatorname{Tr}(AD^2 u) \cdot q = 2(-D_p H D_{xx}^2 u + D_x u_t) D_{pp}^2 H \cdot p - \chi D_p (\operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)) \cdot p$$
$$+ \chi_w \operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)(s - D_p H \cdot p), \tag{3.7}$$

$$D_q b(x, Du) \cdot q = -(D_p H D_{xp}^2 H) \cdot p - (D_x H D_{pp}^2 H) \cdot p + (D_x f D_{pp}^2 H) \cdot p$$
$$+ \frac{1}{f_m} D_x f_m \cdot D_p H(-s + D_p H \cdot p) - \chi D_p \operatorname{Tr}(D_{xp}^2 H) \cdot p$$
$$+ \chi_w \operatorname{Tr}(D_{xp}^2 H)(s - D_p H \cdot p), \tag{3.8}$$

$$L_u v = -\operatorname{Tr}(D_{qq}^2 \Phi D^2 u A D^2 u) - \operatorname{Tr}(A D_{\bar{x}\bar{x}}^2 \Phi) - 2 \operatorname{Tr}(A D^2 u D_{\bar{x}q}^2 \Phi)$$
$$- D_p \Phi \cdot D_x b + D_{\bar{x}} \Phi \cdot D_q b$$
$$+ \sum_{i=1}^{d} \operatorname{Tr}(A_{x_i} D^2 u) \Phi_{p_i} - D_{\bar{x}} \Phi \cdot D_q \operatorname{Tr}(A D^2 u). \tag{3.9}$$

Proof. Using (Q1),

$$
\begin{aligned}
&-(D_q \operatorname{Tr}(AD^2 u)) \cdot q \\
&\quad = -D_q (\operatorname{Tr}((D_p H \otimes D_p H + \chi D_{pp}^2 H) D_{xx}^2 u) - 2 D_p H \cdot D_x u_t) \cdot q \\
&\quad = -2(D_p H D_{xx}^2 u D_{pp}^2 H) \cdot p - \chi D_p \operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u) \cdot p \\
&\qquad - \chi_w \operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)(-s + D_p H \cdot p) + 2(D_{pp}^2 H D_x u_t) \cdot p \\
&\quad = 2(-D_p H D_{xx}^2 u + D_x u_t) D_{pp}^2 H \cdot p - \chi D_p (\operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)) \cdot p \\
&\qquad + \chi_w \operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)(s - D_p H \cdot p),
\end{aligned}
$$

which shows (3.7). Equation (3.8) is an immediate consequence of (Q2). From the definition of v, it follows that

$$
Dv = D_{\bar{x}} \Phi + D_q \Phi D^2 u
$$

and

$$
D^2 v = D_{\bar{x}\bar{x}}^2 \Phi + (D^2 u D_{\bar{x}q}^2 \Phi + D_{q\bar{x}}^2 \Phi D^2 u) + D^2 u D_{qq}^2 \Phi D^2 u + D_q \Phi D^3 u.
$$

Thus, differentiating the equation $Qu = 0$ and taking the inner product with $D_q \Phi$ yields

$$
\begin{aligned}
0 &= D_q \Phi \cdot D_{\bar{x}}(-\operatorname{Tr}(A(x, Du(x,t)) D^2 u(x,t)) + b(x, Du(x,t))) \\
&= -\operatorname{Tr}(A D_q \Phi D^3 u) - \sum_{i=1}^{d+1} \operatorname{Tr}(A_{\bar{x}_i} D^2 u) \Phi_{q_i} - D_q \Phi D^2 u \cdot D_q \operatorname{Tr}(A D^2 u) \\
&\quad + D_q \Phi D^2 u \cdot D_q b + D_q \Phi \cdot D_{\bar{x}} b \\
&= -\operatorname{Tr}(A(D^2 v - (D_{\bar{x}\bar{x}}^2 \Phi + (D^2 u D_{\bar{x}q}^2 \Phi + D_{q\bar{x}}^2 \Phi) D^2 u + D^2 u D_{qq}^2 \Phi D^2 u)) \\
&\quad - D_q \operatorname{Tr}(A D^2 u) \cdot (Dv - D_{\bar{x}} \Phi) \\
&\quad + D_q b \cdot (Dv - D_{\bar{x}} \Phi) - \sum_{i=1}^{d+1} \operatorname{Tr}(A_{\bar{x}_i} D^2 u) \Phi_{q_i} + D_q \Phi D_{\bar{x}} b.
\end{aligned}
$$

Using the fact that A and b are independent of t, as well as (3.6), we obtain

$$
\begin{aligned}
0 &= L_u v + \operatorname{Tr}(A(D_{\bar{x}\bar{x}}^2 \Phi + (D^2 u D_{\bar{x}q}^2 \Phi + D_{q\bar{x}}^2 \Phi D^2 u) + D^2 u D_{qq}^2 \Phi D^2 u)) \\
&\quad + D_q \operatorname{Tr}(A D^2 u) \cdot D_{\bar{x}} \Phi \\
&\quad - D_q b \cdot D_{\bar{x}} \Phi - \sum_{i=1}^{d} \operatorname{Tr}(A_{x_i} D^2 u) \Phi_{p_i} + D_p \Phi \cdot D_x b \\
&= L_u v + \operatorname{Tr}(D_{qq}^2 \Phi D^2 u A D^2 u) + \operatorname{Tr}(A D_{\bar{x}\bar{x}}^2 \Phi) + 2 \operatorname{Tr}(A D^2 u D_{\bar{x}q}^2 \Phi) + D_p \Phi \cdot D_x b \\
&\quad - D_q b \cdot D_{\bar{x}} \Phi - \sum_{i=1}^{d} \operatorname{Tr}(A_{x_i} D^2 u) \Phi_{p_i} + D_{\bar{x}} \Phi \cdot D_q \operatorname{Tr}(A D^2 u),
\end{aligned}
$$

which proves (3.9). ∎

Corollary 3.5. *Let $(u, m) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ be a solution to (MFG), and set*

$$\eta_0 = \min\left(\min_{\mathbb{T}^d} m_0, \min_{\mathbb{T}^d} m(\cdot, T)\right), \quad \eta_1 = \max\left(\max_{\mathbb{T}^d} m_0, \max_{\mathbb{T}^d} m(\cdot, T)\right).$$

Then

$$L_u(u_t) = 0 \quad \text{and} \quad -C_0 - f_1(\eta_1) \le u_t \le \|H(\cdot, D_x u)\|_{C^0(\overline{Q_T})} - f_0(\eta_0). \quad (3.10)$$

Proof. Letting $\Phi(x, t, p, s) = s$ in Lemma 3.4, since $D_{\bar{x}}\Phi, D_p\Phi, D^2\Phi \equiv 0$, it follows that

$$L_u(u_t) = L_u(\Phi(x, t, Du)) = 0.$$

Hence, the maximum and minimum values of u_t are attained in ∂Q_T, and (3.10) then follows immediately from (2.1) and the Hamilton–Jacobi equation (HJ for short) in (MFG). ∎

By Corollary 3.2, this result reduces the problem to estimating $\|D_x u\|_{C^0}$, but it is also a key ingredient for obtaining that bound, particularly due to the fact that the term $\|H(\cdot, D_x u)\|_{C^0(\overline{Q_T})}$ has coefficient 1 in (3.10). We now begin to simplify the quantity (3.9) for the specific Φ that will be used in the proof of the gradient estimate, bounding one of the dominant signed terms by a simpler expression, using matrix algebra.

Lemma 3.6. *Assume that (SE) holds. In addition, for each $(x, t, p, s) \in \overline{Q_T} \times \mathbb{R}^{d+1}$, set $\widetilde{H}(x, t, p, s) = H(x, p)$, and define the matrix $\tilde{I} = (\delta^{ij}(1 - \delta^{i, d+1}))_{i,j=1}^{d+1}$. Then, for every $u \in C^2(Q_T)$,*

$$\text{Tr}(D_{qq}^2 \widetilde{H}(x, t, Du) D^2 u A(x, Du) D^2 u) \ge \frac{3}{4C_0}|-D_x u_t + D_p H(x, D_x u) D_{xx}^2 u|^2$$
$$+ \frac{1}{4C_0} \text{Tr}(\tilde{I} D^2 u A D^2 u)$$
$$+ \frac{3\chi}{4C_0^2}|D_{xx}^2 u|^2. \quad (3.11)$$

Proof. By (H1),

$$D_{qq}^2 \widetilde{H} \ge \frac{1}{C_0}\tilde{I};$$

thus, since the matrix $D^2 u A D^2 u$ is non-negative, multiplying both sides of the inequality by this matrix and taking the trace of both sides gives

$$\text{Tr}(D_{pp}^2 \widetilde{H} D^2 u A D^2 u) \ge \frac{1}{C_0} \text{Tr}(\tilde{I} D^2 u A D^2 u). \quad (3.12)$$

Now, by (Q1) and (H1),

$$\text{Tr}(\tilde{I} D^2 u A D^2 u) = \sum_{k=1}^{d} Du_{x_k} A \cdot Du_{x_k}$$
$$= \sum_{k=1}^{d} |(D_p H, -1) \cdot Du_{x_k}|^2 + \chi D_x u_{x_k} D_{pp}^2 H \cdot D_x u_{x_k}$$

$$\geq \sum_{k=1}^{d} |D_p H \cdot D_x u_{x_k} - u_{tx_k}|^2 + \frac{\chi}{C_0} |D_x u_{x_k}|^2$$

$$= |D_p H D_{xx}^2 u - D_x u_t|^2 + \frac{\chi}{C_0} |D_{xx}^2 u|^2. \tag{3.13}$$

Combining (3.12) and (3.13) yields, as desired,

$$\text{Tr}(D_{pp}^2 \tilde{H} D^2 u A D^2 u) = \frac{3}{4} \big(\text{Tr}(D_{pp}^2 \tilde{H} D^2 u A D^2 u) \big) + \frac{1}{4} \big(\text{Tr}(D_{pp}^2 \tilde{H} D^2 u A D^2 u) \big)$$

$$\geq \frac{3}{4C_0} |-D_x u_t + D_p H D_{xx}^2 u|^2 + \frac{3\chi}{4C_0^2} |D_{xx}^2 u|^2$$

$$+ \frac{1}{4C_0} \text{Tr}(\tilde{I} D^2 u A D^2 u). \qquad\blacksquare$$

The next lemma continues to simplify the linearizations. Since one of the dominant signed terms will later be shown to be of order $|D_x u|^4$, the goal will be to bound everything else by $(4 - \epsilon)$th powers of $|D_x u|$, $(2 - \epsilon)$th powers of u_t (dealing with these through Corollary 3.5), and second derivative terms that can be dealt with using the other dominant term (3.11). The usage of $\Phi(x, t, D_x u, u_t) = H(x, D_x u)$, as opposed to a more standard choice such as $|D_x u|^2$ or $|Du|^2$, is crucial in the next two results, in order to produce structural cancellation of terms that cannot be otherwise estimated, as well as to be able to use (3.10) without gaining any constant factors in the process.

Lemma 3.7. *Assume that* (SE) *holds. Let* $u \in C^3(\overline{Q_T})$ *be a solution to* (Q0)*, and let* $c, c' \in \mathbb{R}$. *Define*

$$\tilde{u} = u + c(T - t) + c', \quad v_1 = \frac{\tilde{u}^2}{2}, \quad v_2 = H(\cdot, D_x u).$$

Then, for each $(x, t) \in Q_T$, *there exists* $C(x, t) > 0$, *with*

$$C(x, t) = C\left(C_0, \|\tilde{u}\|_{C^0(Q_T)}, \frac{1}{\chi(x, f(x, m(x, t)))}, |h_w(x, f(x, m(x, t)))|, c\right),$$

such that

$$L_u(v_1) \leq -\frac{1}{2} |-u_t + D_p H(x, D_x u) D_x u|^2 - \frac{1}{C_0} \chi |D_x u|^2$$

$$+ C(x, t) \big(1 + |D_x u|^{2+\tau} + \chi |D_{xx}^2 u|^2 |f|^{\tau} |D_x u|^2 + |f|(1 + |D_x u|^{\tau})$$

$$+ |f|^{(1+\tau)/2} |D_x u| + \chi + |-D_x u_t + D_{xx}^2 u D_p H|^2 \big) \tag{3.14}$$

and

$$L_u(v_2) \leq \frac{-1}{2C_0} |-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2} |D_{xx}^2 u|^2$$

$$+ C(x, t) \big(1 + |D_x u|^{3+\tau} + \chi(1 + |D_x u|^{1+\tau}) + \chi^{(1+\tau)/2} |D_x u|^2$$

$$+ |f|(1 + |D_x u|^{1+\tau}) + |D_x u|^2 |f|^{(1+\tau)/2} \big). \tag{3.15}$$

Proof. Throughout this proof, the number $C = C(x, t)$ may increase at each step, with its size depending on (x, t) only monotonically through $\frac{1}{\chi}$ and $|h_w|$. For this reason, there is no loss of generality in assuming

$$\frac{1}{\chi} + |h_w| \le C. \tag{3.16}$$

Observe first that, since $D^2\tilde{u} = D^2 u$, one has $-\operatorname{Tr}(A(x, Du)D^2\tilde{u}) = -b(x, Du)$. Therefore,

$$- \operatorname{Tr}(A(x, Du)D^2 v_1) = -\tilde{u}b(x, Du) - D\tilde{u}A \cdot D\tilde{u},$$

$$- D_q \operatorname{Tr}(A(x, Du)D^2 u) \cdot Dv_1 + D_q b(x, Du) \cdot Dv_1$$
$$= -D_q \operatorname{Tr}(AD^2 u) \cdot \tilde{u}D\tilde{u} + D_q b \cdot \tilde{u}D\tilde{u}.$$

Consequently, by (Q1) and (H3),

$$L_u(v_1) = -\tilde{u}b - D\tilde{u}A \cdot D\tilde{u} + \tilde{u}(-D_q \operatorname{Tr}(AD^2 u) \cdot D\tilde{u} + D_q b \cdot D\tilde{u}) \tag{3.17}$$
$$= -\tilde{u}b - D\tilde{u}(D_p H, -1) \otimes (D_p H, -1) \cdot D\tilde{u} - \chi D_x\tilde{u}D_{pp}^2 H \cdot D_x\tilde{u}$$
$$\quad + \tilde{u}(-D_q \operatorname{Tr}(AD^2 u) \cdot D\tilde{u} + D_q b \cdot D\tilde{u})$$
$$\le -|-\tilde{u}_t + D_p H \cdot D_x u|^2 - \frac{1}{C_0}\chi|D_x\tilde{u}|^2$$
$$\quad + \tilde{u}(-b - D_q \operatorname{Tr}(ADu) \cdot D\tilde{u} + D_q b \cdot D\tilde{u})$$
$$= -|-\tilde{u}_t + D_p H \cdot D_x u|^2 - \frac{1}{C_0}\chi|D_x\tilde{u}|^2 + \tilde{u}(J_1 + J_2 + J_3).$$

The next task will be to estimate the terms J_i. In view of (Q2), (2.1), (2.2), (2.3), (F2), and (FX2),

$$|J_1| = |b(x, Du)| = |-D_x H \cdot D_p H + D_x f \cdot D_p H - \chi \operatorname{Tr}(D_{xp}^2 H)|$$
$$\le C(1 + |D_x u|^{2+\tau} + |f|^{(1+\tau)/2}(1 + |D_x u|) + |f|(1 + |D_x u|^\tau)). \tag{3.18}$$

As for J_2, (2.1), (H3), (3.7), and (3.16) imply that

$$|J_2| = |-D_q \operatorname{Tr}(AD^2 u) \cdot D\tilde{u}|$$
$$= |2(-D_p H D_{xx}^2 u + D_x u_t)D_{pp}^2 H \cdot D_x u - \chi D_p(\operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)) \cdot D_x u$$
$$\quad + 2hh_w \operatorname{Tr}(D_{pp}^2 H D_{xx}^2 u)(\tilde{u}_t - D_p H \cdot D_x u)|$$
$$\le C(1 + |-D_x u_t + D_p H D_{xx}^2 u|^2 + |D_x u|^2 + \chi + \chi|D_{xx}^2 u|^2)$$
$$\quad + \frac{1}{8(\|\tilde{u}\| + 1)}|-\tilde{u}_t + D_p H \cdot D_x u|^2. \tag{3.19}$$

By assumptions (HX), (F2), (FX1), and (FX2), together with (2.2), (3.8) and (3.16),

$$
\begin{aligned}
|J_3| &= |D_q b(x, Du) \cdot D\tilde{u}| \\
&= \Big| -(D_p H D^2_{xp} H) \cdot D_x u - (D_x H D^2_{pp} H) \cdot D_x u + (D_x f D^2_{pp} H) \cdot D_x u \\
&\quad + \Big(\frac{1}{f_m} D_x f_m \cdot D_p H \Big)(-\tilde{u}_t + D_p H \cdot D_x u) - \chi D_p(\mathrm{Tr}(D^2_{xp} H)) \cdot D_x u \\
&\quad + \chi_w \mathrm{Tr}(D^2_{xp} H)(\tilde{u}_t - D_p H \cdot D_x u) \Big| \\
&\le C \Big(1 + |D_x u|^{2+\tau} + (1 + |f|^{(1+\tau)/2})|D_x u| \\
&\quad + \Big(1 + \frac{1}{\chi} \Big)(1 + |f|^{\tau/2})(1 + |D_x u|)|-\tilde{u}_t + D_p H \cdot D_x u| \\
&\quad + (1 + |f|)(1 + |D_x u|^{\tau}) + (1 + |D_x u|^{\tau})|-\tilde{u}_t + D_p H \cdot D_x u| \Big) \\
&\le C(1 + |D_x u|^{2+\tau} + |f|^{(1+\tau)/2}|D_x u| + |f|^{\tau}|D_x u|^2 + |f||D_x u|^{\tau}) \\
&\quad + \frac{1}{8(\|\tilde{u}\| + 1)}|-\tilde{u}_t + D_p H \cdot D_x u|^2. \tag{3.20}
\end{aligned}
$$

Finally, using (3.18), (3.19), and (3.20) in (3.17) yields

$$
\begin{aligned}
L_u(v_1) &\le -\frac{3}{4}|-\tilde{u}_t + D_p H D_x u|^2 - \frac{1}{C_0}\chi|D_x u|^2 \\
&\quad + C(1 + |D_x u|^{2+\tau} + \chi|D^2_{xx} u|^2 + |f|^{\tau}|D_x u|^2 + |f|(1 + |D_x u|^{\tau}) \\
&\quad + |f|^{(1+\tau)/2}|D_x u| + \chi + |-D_x u_t + D^2_{xx} u D_p H|^2),
\end{aligned}
$$

which proves (3.14).

Next is the proof of (3.15). In view of Lemma 3.4, recalling that $\tilde{H}(x, t, Du) := H(x, Du)$,

$$
\begin{aligned}
L_u v_2 &= L_u(\tilde{H}(x, t, Du)) \\
&= -\mathrm{Tr}(D^2_{qq}\tilde{H} D^2 u A D^2 u) - \mathrm{Tr}(A D^2_{\bar{x}\bar{x}}\tilde{H}) - 2\mathrm{Tr}(A D^2 u D^2_{\bar{x}q}\tilde{H}) - D_p\tilde{H} \cdot D_x b \\
&\quad + D_{\bar{x}}\tilde{H} \cdot D_q b + \sum_{i=1}^{d} \mathrm{Tr}(A_{x_i} D^2 u)\tilde{H}_{p_i} - D_{\bar{x}}\tilde{H} \cdot D_q \mathrm{Tr}(A D^2 u),
\end{aligned}
$$

and Lemma 3.6 then implies

$$
\begin{aligned}
L_u v_2 &\le \frac{-3}{4C_0}|-D_x u_t + D_p H D^2_{xx} u|^2 - \frac{3\chi}{4C_0^2}|D^2_{xx} u|^2 - \frac{1}{4C_0}\mathrm{Tr}(\tilde{I} D^2 u A D^2 u) \tag{3.21} \\
&\quad - \mathrm{Tr}(A D^2_{\bar{x}\bar{x}}\tilde{H}) - 2\mathrm{Tr}(A D^2 u D^2_{\bar{x}q}\tilde{H}) - D_p H \cdot D_x b + D_x H \cdot D_p b \\
&\quad + \sum_{i=1}^{d} \mathrm{Tr}(A_{x_i} D^2 u) H_{p_i} - D_x H \cdot D_p \mathrm{Tr}(A D^2 u)
\end{aligned}
$$

$$= \frac{-3}{4C_0}|-D_xu_t + D_pHD_{xx}^2u|^2 - \frac{3\chi}{4C_0^2}|D_{xx}^2u|^2 - \frac{1}{4C_0}\operatorname{Tr}(\tilde{I}D^2uAD^2u)$$
$$+ K_1 + K_2 + K_3 + K_4 + K_5 + K_6.$$

As before, we proceed to estimate the K_i. Starting with K_1, we observe that, by (Q1), (H1), (2.1), and (2.3),

$$|K_1| = |\operatorname{Tr}(AD_{\bar{x}\bar{x}}^2\tilde{H})| \le C(1 + |D_xu|^{1+\tau})|A| \\ \le C(1 + |D_xu|^{1+\tau})(1 + |D_xu|^2 + \chi). \tag{3.22}$$

Similarly, using the Cauchy–Schwarz inequality,

$$|K_2| = |2\operatorname{Tr}(AD^2u(\tilde{I}D_{\bar{x}q}^2\tilde{H}))| = 2|\operatorname{Tr}(D_{\bar{x}q}^2\tilde{H}A(\tilde{I}D^2u)^T)|$$
$$\le \frac{1}{4C_0}\operatorname{Tr}((\tilde{I}D^2u)A(\tilde{I}D^2u)^T) + C\operatorname{Tr}(D_{\bar{x}q}^2\tilde{H}A(D_{\bar{x}q}^2\tilde{H})^T)$$
$$\le \frac{1}{4C_0}\operatorname{Tr}(\tilde{I}D^2uAD^2u) + C|D_{xp}^2H|^2(1 + |D_xu|^2 + \chi)$$
$$\le \frac{1}{4C_0}\operatorname{Tr}(\tilde{I}D^2uAD^2u) + C(1 + |D_xu|^{2(1+\tau)} + \chi(1 + |D_xu|^{2\tau})). \tag{3.23}$$

Next we will estimate $|K_3 + K_4|$. Differentiating equation (Q2) with respect to x, we obtain

$$D_xb(x,p,s) = -D_{xx}^2H \cdot D_pH - D_xHD_{xp}^2H + D_pHD_{xx}^2f$$
$$+ \frac{1}{f_m}D_pH(D_xf_m \otimes (-D_xf + D_xH))$$
$$+ D_xfD_{xp}^2H + \left(D_xf - mD_xf_m + \frac{mf_{mm}}{f_m}D_xf\right)\operatorname{Tr}(D_{xp}^2H)$$
$$- \chi_wD_xH\operatorname{Tr}(D_{xp}^2H) - \chi D_x\operatorname{Tr}(D_{xp}^2H). \tag{3.24}$$

Consequently, (2.1), (2.2), (2.3), (F2), (FX1), (FX2), and (3.16) yield

$$\left|-D_xb(x,p,s) + \frac{1}{f_m}D_pHD_xf_m \otimes D_xH\right|$$
$$\le C(1 + |p|^{2+\tau} + (1 + |p|)|f|^{(1+\tau)/2} + |f|(1 + |p|^\tau)),$$

so that, setting $z_1 = \frac{1}{f_m}(D_pH \cdot D_xf_m)(D_xH \cdot D_pH)$,

$$|K_3 + z_1| = |-D_xb(x,Du) \cdot D_pH + z_1|$$
$$\le C(1 + |D_xu|^{2+\tau} + (1 + |D_xu|)|f|^{(1+\tau)/2} + |f|(1 + |D_xu|^\tau))(1 + |D_xu|)$$
$$\le C(1 + |D_xu|^{3+\tau} + (1 + |D_xu|^2)|f|^{(1+\tau)/2} + |f|(1 + |D_xu|^{1+\tau})). \tag{3.25}$$

On the other hand, by (3.8),

$$|K_4 - z_1| = |D_pb \cdot D_xH - z_1|$$
$$= |-(D_pHD_{xp}^2H) \cdot D_xH - (D_xHD_{pp}^2H) \cdot D_xH + (D_xfD_{pp}^2H) \cdot D_xH$$
$$+ z_1 - \chi D_p\operatorname{Tr}(D_{xp}^2H) \cdot D_xH + \chi_w\operatorname{Tr}(D_{xp}^2H)(D_pH \cdot D_xH) - z_1|.$$

The terms z_1 and $-z_1$ then cancel out, and therefore (HX), (2.2), (2.3), and (F2) yield

$$|K_4 - z_1| \leq C(1 + |D_x u|^{2+2\tau} + |f|^{(1+\tau)/2}(1 + |D_x u|^{1+\tau}) + |f||D_x u|^{2\tau}). \quad (3.26)$$

The inequalities (3.25) and (3.26) thus imply

$$|K_3 + K_4| \leq C(1 + |D_x u|^{3+\tau} + (1 + |D_x u|^2)|f|^{(1+\tau)/2} + |f|(1 + |D_x u|^{1+\tau})). \quad (3.27)$$

The terms K_5 and K_6 will also be treated jointly. Let $(x, p, s) \in \mathbb{T}^d \times \mathbb{R}^{d+1}$, and set $w = -s + H(x, p)$. It follows from (Q1), (FX1), (HX), and (F2) that

$$A_{x_i}(x, p, s) = (D_p H_{x_i}, 0) \otimes (D_p H, -1) + (D_p H, -1) \otimes (D_p H_{x_i}, 0) + \chi_w H_{x_i} D_{qq}^2 \tilde{H}$$
$$+ O(1 + |w|^{\tau/2} + \chi^{(1+\tau)/2} + \chi(1 + |p|)^{-1+\tau}) \tilde{I}.$$

Therefore, using (3.7), and writing $z_2 = \chi_w \mathrm{Tr}(D_{pp}^2 H D_{xx}^2 u)(D_p H \cdot D_x H)$,

$$|K_5 + K_6| = \left| \sum_{i=1}^d \mathrm{Tr}(A_{x_i} D^2 u) H_{p_i} - D_p \mathrm{Tr}(A D^2 u) \cdot D_x H \right|$$
$$\leq |2(D_p H D_{xx}^2 u - D_x u_t) D_{px}^2 H \cdot D_p H + z_2$$
$$+ 2(-D_p H D_{xx}^2 u + D_x u_t) D_{pp}^2 H \cdot D_x H$$
$$- \chi D_p(\mathrm{Tr}(D_{pp}^2 H D_{xx}^2 u)) \cdot D_x H - z_2|$$
$$+ C(1 + |f|^{\tau/2} + \chi^{(1+\tau)/2} + \chi(1 + |D_x u|)^{-1+\tau})|D_{xx}^2 u|(1 + |D_x u|).$$

Once more, cancellation occurs and, consequently, (3.16), (H3), (2.1), (2.2), and (2.3) imply that

$$|K_5 + K_6| \leq \frac{1}{4C_0}|D_p H D_{xx}^2 u - D_x u_t|^2 + \frac{\chi}{4C_0^2}|D_{xx}^2 u|^2 \quad (3.28)$$
$$+ C(1 + |D_x u|^{2+2\tau} + (1 + |D_x u|^2)(1 + |f|^\tau + \chi^{(1+\tau)/2}) + \chi(1 + |D_x u|^{2\tau})).$$

Using (3.22), (3.23), (3.27), and (3.28) in (3.21) yields (3.15), completing the proof. ∎

We can now obtain the a priori gradient bound in terms of bounds for the solution u and the terminal density $m(\cdot, T)$, which were obtained in the previous subsection.

Lemma 3.8. *Assume that* (SE) *holds, and let* $(u, m) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ *be a solution to* (MFG). *For $K > 0$, set*

$$\beta_K = \|f\|_{C^1(\mathbb{T}^d \times [\frac{1}{K}, K])} + \|D_x g\|_{C^1(\mathbb{T}^d \times [\frac{1}{K}, K])} + \left\| \frac{1}{\chi} \right\|_{C^0(\mathbb{T}^d \times [-K, \infty))}$$
$$+ \|h_w\|_{C^0(\mathbb{T}^d \times [-K, \infty))}. \quad (3.29)$$

There exist constants C, C_1 with

$$C = C(C_1, \beta_{C_1}),$$
$$C_1 = C_1\left(C_0, T, \frac{1}{T}, \frac{1}{1-\tau}, \|u\|_{C^0(\overline{Q_T})}, \max m(T), \frac{1}{\min m(T)}, f_0\left(\min_{\mathbb{T}^d} m(T)\right)^-\right),$$

such that

$$\|Du\|_{C^0(\overline{Q_T})} \le C.$$

Proof. As was mentioned, the proof will proceed through Bernstein's method. By Corollary 3.5, it is sufficient to bound the space gradient. Since the estimate will be up to the boundary, as in [18], we linearize the HJ equation that holds at the extremal times:

$$T_u v = -v_t + D_p H(x, Du) D_x v.$$

We now normalize u to have a prescribed sign at the initial and terminal times. That is, we set

$$\tilde{u} = u + \|u\|_{C^0(\overline{Q_T})} + 1 - \frac{2(\|u\|_{C^0(\overline{Q_T})} + 1)}{T}(T - t),$$

so that

$$|\tilde{u}| \le C, \qquad \tilde{u}(\cdot, 0) \le -1, \quad \tilde{u}(\cdot, T) \ge 1, \tag{3.30}$$

and define

$$v(x, t) = H(x, D_x u) + \frac{c_1}{2}\tilde{u}^2,$$

where $0 < c_1 \le 1$ is a constant to be chosen later. Let $(x_0, t_0) \in \overline{Q_T}$ be a point where v achieves its maximum value. We will distinguish three cases:

Case 1. $t_0 = T$. Then $D_x v = 0$, $v_t \ge 0$. Therefore, (3.30), (H1) (H2), (2.2), and the HJ equation in (MFG), together with the fact that $m(\cdot, T) = g^{-1}(\cdot, u(\cdot, T))$, yield

$$\begin{aligned}
0 \ge T_u v &= T_u(H(x_0, D_x u)) + c_1 \tilde{u}(-\tilde{u}_t + D_p H(x_0, D_x u) \cdot D_x \tilde{u}) \\
&= D_x(f(x_0, m(x_0, T))) \cdot D_p H(x_0, D_x u) \\
&\quad + c_1 \tilde{u}(-u_t + D_p H(x_0, D_x u) \cdot D_x u - C) \\
&\ge \left(D_x f - \frac{f_m}{g_m} D_x g + \frac{f_m}{g_m} D_x u\right) \cdot D_p H + c_1 \tilde{u}(-u_t + 2H - C) \\
&\ge -C\left(1 + \frac{f_m}{g_m}\right)(1 + |D_x u|) + \frac{f_m}{g_m} D_p H \cdot D_x u + c_1 \tilde{u}(f + H - C) \\
&\ge -C\left(1 + \frac{f_m}{g_m}\right)(1 + |D_x u|) + 2\left(c_1 \tilde{u} + \frac{f_m}{g_m}\right)H.
\end{aligned}$$

Thus, by (2.1),

$$|H(x_0, Du(x_0, t_0))| \le C.$$

Case 2. $t_0 = 0$. Similarly, we obtain $D_x v = 0$, $v_t \le 0$, and, since $\tilde{u}(\cdot, 0) \le -1$,

$$\begin{aligned}
0 \le T_u v &= D_x(f(x_0, m_0(x_0))) \cdot D_p H + c_1 \tilde{u}(-u_t + D_p H \cdot D_x u - C) \\
&\le C(1 + |D_x u|) + c_1 \tilde{u}(f(x_0, m_0) + H - C) \le C(1 + |D_x u|) + C + c_1 \tilde{u} H.
\end{aligned}$$

This implies $-c_1 \tilde{u}(H(x_0, D_x u)) \le C(1 + |D_x u|)$, and so we conclude once more that

$$|H(x_0, Du(x_0, t_0))| \le C.$$

Case 3. $0 < t_0 < T$. Then $Dv = 0$, $D^2v \le 0$, which yields

$$0 \le L_u v.$$

In order to make use of Lemma 3.7, it is necessary to eliminate the (x_0, t_0) dependence of the "constant" $C(x_0, t_0)$ from the lemma, which amounts to establishing an a priori upper bound on the quantities $1/\chi$ and $|h_w|$ at the point (x_0, t_0). By (F1) and (F2), $1/\chi$ and $|h_w| = |\chi_w/2\sqrt{\chi}|$ are both bounded above as $w \to \infty$, so it is enough to establish a lower bound for $w = f(x_0, m(x_0, t_0))$. By Corollary 3.5, there exists a point $(x_1, t_1) \in \partial Q_T$ where u_t achieves its maximum value. Then, since (x_0, t_0) is a maximum point for v, and the initial and terminal densities are both bounded below a priori,

$$
\begin{aligned}
f(x_0, m(x_0, t_0)) &= -u_t(x_0, t_0) + H(x_0, D_x u(x_0, t_0)) \\
&\ge -u_t(x_1, t_1) + H(x_1, D_x u(x_1, t_1)) - \frac{c_1}{2}\|\tilde{u}\|^2_{C^0(\overline{Q_T})} \\
&= f(x_1, m(x_1, t_1)) - \frac{c_1}{2}\|\tilde{u}\|^2_{C^0(\overline{Q_T})} \ge f_0(m(x_1, t_1)) - C \ge -C.
\end{aligned}
$$

This estimate, together with (H2) and (2.1), allows us to identify the dominant power of $|D_x u|$ in the linearization,

$$|-u_t + D_p H \cdot D_x u|^2 \ge (f + H)^2 - C \ge \frac{1}{2C_0^2}|D_x u|^4 - C. \tag{3.31}$$

Now, because of the form of the estimate in Lemma 3.7, it is also necessary to be able to compare powers of $|f|$ with powers of $|D_x u|$. By Corollary 3.5 and (2.1),

$$
\begin{aligned}
f(x_0, m(x_0, t_0)) &\le -u_t(x_0, t_0) + H(x_0, D_x u(x_0, t_0)) \\
&\le C + 2H(x_0, D_x u(x_0, t_0)) \le C(1 + |D_x u|^2).
\end{aligned} \tag{3.32}
$$

With these preliminaries done, we now apply Lemma 3.7, obtaining

$$
\begin{aligned}
0 \le L_u(v) = L_u \tilde{H} &+ c_1 L_u\left(\frac{\tilde{u}^2}{2}\right) \\
\le \frac{-1}{2C_0}&|-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2}|D_{xx}^2 u|^2 \\
&+ C(1 + |D_x u|^{3+\tau} + \chi(1 + |D_x u|^{1+\tau}) + \chi^{(1+\tau)/2}|D_x u|^2 \\
&\quad + |f|(1 + |D_x u|^{1+\tau}) + |D_x u|^2|f|^{(1+\tau)/2}) \\
&- \frac{c_1}{2}|-u_t + D_p H \cdot D_x u|^2 - \frac{c_1}{C_0}\chi|D_x u|^2 \\
&+ Cc_1(\chi|D_{xx}^2 u|^2 + |-D_x u_t + D_{xx}^2 u D_p H|^2).
\end{aligned} \tag{3.33}
$$

Applying (3.31) and (3.32) yields

$$
\begin{aligned}
0 \le -\frac{c_1}{4C_0^2}|D_x u|^4 &- \frac{c_1}{C_0}\chi|D_x u|^2 - \frac{1}{2C_0}|-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2}|D_{xx}^2 u|^2 \\
&+ C(1 + |D_x u|^{3+\tau} + \chi(1 + |D_x u|^{1+\tau}) + \chi^{(1+\tau)/2}|D_x u|^2) \\
&+ Cc_1(\chi|D_{xx}^2 u|^2 + |-D_x u_t + D_{xx}^2 u D_p H|^2)
\end{aligned}
$$

$$\leq -\frac{c_1}{4C_0^2}|D_x u|^4 - \frac{c_1}{C_0}\chi |D_x u|^2 - \frac{1}{2C_0}|-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2}|D_{xx}^2 u|^2$$

$$+ C(1 + |D_x u|^{3+\tau} + \chi(1 + |D_x u|^{1+\tau})) + \left(\frac{C}{c_1} + \frac{c_1}{2C_0}\chi\right)|D_x u|^2$$

$$+ Cc_1(\chi|D_{xx}^2 u|^2 + |-D_x u_t + D_{xx}^2 u D_p H|^2).$$

Now fix c_1 satisfying $c_1 < \frac{1}{4C_0 C(1+C_0)}$, where C is as in the previous line. This gives

$$0 \leq -\frac{c_1}{4C_0^2}|D_x u|^4 - \frac{1}{2C_0}\chi(c_1|D_x u|^2 - 2CC_0(1 + |D_x u|^{1+\tau}))$$

$$+ C(1 + |D_x u|^{3+\tau}) + \frac{C}{c_1}|D_x u|^2,$$

which may be rearranged as

$$\frac{c_1}{4C_0^2}|D_x u|^4 + \frac{1}{2C_0}\chi(c_1|D_x u|^2 - 2CC_0|D_x u|^{1+\tau} - 2CC_0)$$

$$\leq C(1 + |D_x u|^{3+\tau}) + \frac{C}{c_1}|D_x u|^2.$$

This finally implies that

$$c_1|D_x u|^2 - 2CC_0|D_x u|^{1+\tau} - 2CC_0 \leq 0$$

or

$$\frac{c_1}{4C_0^2}|D_x u|^4 \leq C(1 + |D_x u|^{3+\tau}) + \frac{C}{c_1}|D_x u|^2,$$

either of which yields

$$|H(x_0, D_x u(x_0, t_0))| \leq C. \qquad \blacksquare$$

We now summarize all of the a priori bounds obtained in this section.

Theorem 3.9. *Assume that* (SE) *holds, let* $(u, m) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ *be a solution to* (MFG), *and let* β *be defined by* (3.29). *Then there exist constants* L, L_1, K, K_1, *with*

$$L = \left(L_1, |g_1 f_0^{-1}(L_1)|, |g_0 f_1^{-1}(-L_1)|, g_0^{-1} g_1 f_0^{-1}(L_1), \frac{1}{g_1^{-1} g_0 f_1^{-1}(-L_1))}\right),$$

$$L_1 = L_1(C_0, T),$$

$$K = (K_1, \beta_{K_1}), \quad K_1 = K_1\left(L, \frac{1}{T}, \frac{1}{1-\tau}, f_0\left(\frac{1}{L}\right)^-\right),$$

such that

$$\|u\|_{C^0(\overline{Q_T})} + \|m(T)\|_{C^0(\mathbb{T}^d)} + \left\|\frac{1}{m(T)}\right\|_{C^0(\mathbb{T}^d)} \leq L \quad \text{and} \quad \|Du\|_{C^0(\overline{Q_T})} \leq K.$$

Proof. This result follows simply by the successive application of Lemma 3.1, Corollary 3.2, and Lemma 3.8. ∎

The following variation of Theorem 3.9 shows that, in the standard case where $H(x, p) = H(p) - V(x)$ and $f(x, m) = f(m)$, condition (F2) which requires f to grow at most polynomially may be significantly relaxed.

Theorem 3.10. *The conclusion of Theorem 3.9 still holds if condition (F2) is replaced by*

$$D^2_{xp}H, D^2_{xm}f \equiv 0 \quad and \quad \limsup_{x \in \mathbb{T}^d, w \to \infty} |h_w(x, w)| < \infty. \tag{HFX*}$$

Proof. We simply address all of the instances in which condition (F2) has been used so far. In the proofs of Lemma 3.1, Corollary 3.2, and Lemma 3.7, (F2) was exclusively used to estimate either space derivatives $D_x f$, $D_x H$, or terms that involve mixed derivatives $D^2_{xm}f$, $D^2_{xp}H$. With (HFX*) in place, such terms are, respectively, either bounded in C^1 norm or trivially zero. Condition (F2) was also used in the proof of Lemma 3.8 in order to obtain a bound for $|h_w|$ as $w \to \infty$, but this bound exists here by assumption. ∎

We note that the condition that (HFX*) imposes on h may be equivalently rewritten, in terms of f, as

$$\limsup_{x \in \mathbb{T}^d, m \to \infty} \frac{1}{m f_m} \left| m \frac{f_{mm}}{f_m} + 1 \right|^2 < \infty.$$

This condition, in particular, allows for f to be combinations of powers m^α, $-m^{-\beta}$, exponentials e^m, $-e^{1/m}$, and such typical examples, as long as one has the required blowup near $m = 0$ and as $m \to \infty$.

4. Classical solutions

To obtain classical solutions, it is necessary to have Hölder estimates for the gradient of the solution in terms of the C^1 norm. The following lemma, which is merely a restatement of Theorems 2.2 and 2.3 in the context of the MFG system, provides such an estimate.

Lemma 4.1. *Let $(u, m) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ be a solution to (Q0), and set $K = \|u\|_{C^1(\overline{Q_T})}$. Let $\mu_K, \lambda_K > 0$ be such that (2.5) holds in $\mathbb{T}^d_K$, and conditions (2.6) and*

$$\lambda_K \leq D_q B \cdot \nu \tag{4.1}$$

hold in $\partial Q_{T,K}$. There exist constants $C > 0, 0 < \gamma < 1$, with

$$C = C\left(K, \frac{\mu_K}{\lambda_K}\right), \quad \gamma = \gamma\left(K, \frac{\mu_K}{\lambda_K}\right),$$

such that

$$[Du]_{\gamma, \overline{Q_T}} \leq C.$$

Proof. The only thing to remark is that in order to apply Theorem 2.3, it is necessary to verify that λ_K can be chosen to satisfy (4.1), or, in other words, that N is indeed an oblique boundary operator. This follows directly from (B1), since

$$D_q B(x, 0, z, q) \cdot v(x, 0) = -B_s(x, 0, z, q) = 1 > 0,$$

$$D_q B(x, T, z, q) \cdot v(x, T) = B_s(x, T, z, q) = g_m f_w^{-1} = \frac{g_m}{f_m} > 0.$$

Therefore, the result follows by applying Theorems 2.2 and 2.3 locally, and extracting a finite subcover of $\overline{Q_T}$. The use of Theorem 2.3 is particularly straightforward since the boundary of Q_T is already flat. ∎

The strategy to prove existence will be to use the non-linear method of continuity, by constructing an explicit homotopy $(Q^\theta, N^\theta)_{\theta \in [0,1]}$ between (Q0) and an elliptic problem that comes from a much simpler MFG system, and trivially has a smooth solution. For each $\theta \in [0, 1]$ and each $(x, p, m) \in \mathbb{T}^d \times \mathbb{R}^d \times (0, \infty)$, define

$$H^\theta(x, p) = \theta H(x, p) + (1 - \theta)\left(\frac{1}{2}|p|^2 + f(x, 1)\right),$$

$$g^\theta(x, m) = \theta g(x, m) + (1 - \theta)m,$$

$$m_0^\theta(x) = \theta m_0(x) + (1 - \theta),$$

and consider the family of MFG systems

$$\begin{cases} -u_t + H^\theta(\cdot, D_x u) = f(\cdot, m), & u(\cdot, T) = g^\theta(\cdot, m(\cdot, T)), \\ m_t - \text{div}(m D_p H^\theta(\cdot, D_x u)) = 0, & m(\cdot, 0) = m_0^\theta(\cdot). \end{cases} \quad (\text{MFG}_\theta)$$

We observe that, when $\theta = 0$, the unique solution is $(u, m) \equiv (1, 1)$. Let $(Q^\theta u, N^\theta u)$ be the operators for the corresponding elliptic problem associated to (MFG$_\theta$), and let A^θ, b^θ, and B^θ be their coefficients. The following straightforward lemma is a version of Theorem 3.9, tailored to the family (MFG$_\theta$), that also includes the Hölder estimates of Lemma 4.1, and provides a priori bounds that hold uniformly in θ.

Lemma 4.2. *Assume that* (SE) *holds. For each* $\theta \in [0, 1]$, *let* $(u^\theta, m^\theta) \in C^{3,\alpha}(\overline{Q_T}) \times C^{2,\alpha}(\overline{Q_T})$ *be a solution to* (MFG$_\theta$). *Then there exist constants* $C > 0$ *and* $0 < \gamma < 1$, *independent of* θ, *such that*

$$\|u^\theta\|_{C^{1,\gamma}(\overline{Q_T})} \le C.$$

Proof. The strategy here is to apply Theorem 3.9 and Lemma 4.1 to the corresponding MFG system (MFG$_\theta$) that arises from the new data $H^\theta, g^\theta, m_0^\theta$, and to prove that those results lead to bounds that are uniform in θ. Let β be defined by (3.29), and, for each $\theta \in [0, 1]$, let $C_{0,\theta}$ and $0 \le \tau^\theta < 1$ be any two constants large enough that the inequalities (H1), (H2), (H3), (HX), (FX1), (FX2), (2.1), (2.2), (2.3), and (2.4) all hold when H, g,

m_0 are replaced by H^θ, g^θ, m_0^θ. Theorem 3.9 then yields constants L_θ, $L_{1,\theta}$, K_θ, $K_{1,\theta}$, with

$$L_\theta = \Big(L_{1,\theta}, |g_1^\theta f_0^{-1}(L_{1,\theta})|, |g_0^\theta f_1^{-1}(-L_{1,\theta})|, (g_0^\theta)^{-1} g_1^\theta f_0^{-1}(L_{1,\theta}),$$

$$\frac{1}{(g_1^\theta)^{-1} g_0^\theta f_1^{-1}(-L_{1,\theta})} \Big),$$

$$L_{1,\theta} = M(C_{0,\theta}, T),$$

$$K_\theta = K_\theta(K_{1,\theta}, \beta_{K_{1,\theta}}), \quad K_{1,\theta} = K_{1,\theta}\Big(L_\theta, \frac{1}{T}, \frac{1}{1-\tau^\theta}, f_0\Big(\frac{1}{L_\theta}\Big)^- \Big),$$

such that

$$\|u^\theta\|_{C^0(\overline{Q_T})} + \|m^\theta(T)\|_{C^0(\mathbb{T}^d)} + \Big\|\frac{1}{m^\theta(T)}\Big\|_{C^0(\mathbb{T}^d)} \le L_\theta, \quad \|Du^\theta\|_{C^0(\overline{Q_T})} \le K_\theta.$$

The goal is now to show that L_θ, K_θ may be chosen independently of θ. First we prove that this is true for $C_{0,\theta}$ and τ^θ. Conditions (FX1) and (FX2) trivially hold for the same C_0 and the new H^θ, g^θ, m_0^θ, because the functions H, g, m_0 do not appear in those inequalities. Since the map $H^0(p,x) = \frac{1}{2}|p|^2 + f(x,1)$ satisfies $D_p H^0 \equiv p$, it also satisfies (H1), (H3), (HX), and (2.3), with C_0 being replaced by a universal constant. Thus, since H^θ is a convex combination of H^0 and H, these inequalities still hold for H^θ, when C_0 is replaced by a convex combination of C_0 and a universal constant. By the same reasoning, conditions (H2), (2.1), and (2.2) hold for H^θ after replacing C_0 with a convex combination of C_0 and a constant depending only on C_0 and $\|f(\cdot,1)\|_{C^2(\mathbb{T}^d)} \le C_0$. Only condition (2.4) is left to consider, namely

$$\|\chi(\cdot,0)\|_{C^0(\mathbb{T}^d)} + \|m_0^\theta\|_{C^1(\mathbb{T}^d)} + \|f\|_{C^2(\mathbb{T}^d \times [\min m_0^\theta, \max m_0^\theta])} \le C_{0,\theta}. \tag{4.2}$$

The first term is already independent of θ, whereas, noticing that $\min m_0 \le 1 \le \max m_0$ and $|D_x m_0^\theta| = \theta |D_x m_0|$,

$$\|m_0^\theta\|_{C^1(\mathbb{T}^d)} + \|f\|_{C^2(\mathbb{T}^d \times [\min m_0^\theta, \max m_0^\theta])}$$

$$\le \|m_0\|_{C^1(\mathbb{T}^d)} + \|f\|_{C^2(\mathbb{T}^d \times [\min m_0, \max m_0])} \le C_0.$$

Thus, one may select

$$C_{0,\theta} = C_{0,\theta}(C_0), \quad \tau^\theta = \tau, \tag{4.3}$$

and consequently

$$L_{1,\theta} = L_{1,\theta}(C_{0,\theta}, T) = L_{1,\theta}(C_0, T) := L_1.$$

Now, by definition,

$$g_0^\theta(m) = \theta g_0 + (1-\theta)m, \quad g_1^\theta(m) = \theta g_1 + (1-\theta)m. \tag{4.4}$$

Therefore,

$$|g_0^\theta f_1^{-1}(-L_1)| \leq \max(|g_0 f_1^{-1}(-L_1)|, f_1^{-1}(-L_1))$$
$$\leq \max(|g_0 f_1^{-1}(-L_1)|, f_0^{-1}(L_1)), \tag{4.5}$$

and similarly,

$$|g_1^\theta f_0^{-1}(L_1)| = |\theta g_1 f_0^{-1}(L_1) + (1-\theta) f_0^{-1}(L_1)|$$
$$\leq \max(|g_1 f_0^{-1}(L_1)|, f_0^{-1}(L_1)). \tag{4.6}$$

On the other hand, the following inequalities hold:

$$(g_0^\theta)^{-1} g_1^\theta \leq g_0^{-1} g_1, \quad g_1^{-1} g_0 \leq (g_1^\theta)^{-1} g_0^\theta. \tag{4.7}$$

Indeed, by (4.4),

$$g_0^\theta g_0^{-1} g_1 = \theta g_0 g_0^{-1} g_1 + (1-\theta) g_0^{-1} g_1 \geq \theta g_1 + (1-\theta) g_0^{-1} g_0 = g_1^\theta,$$

which shows the first inequality in (4.7), with the second one following in the same fashion. Now (4.7) yields

$$(g_0^\theta)^{-1} g_1^\theta f_0^{-1}(L_1) \leq g_0^{-1} g_1 f_0^{-1}(L_1),$$
$$\frac{1}{(g_1^\theta)^{-1} g_0^\theta f_1^{-1}(-L_1)} \leq \frac{1}{g_1^{-1} g_0 f_1^{-1}(-L_1)}. \tag{4.8}$$

Thus, (4.5), (4.6), and (4.8) yield

$$L_\theta = L_\theta \Big(M, |g_1 f_0^{-1}(L_1)|, |g_0 f_1^{-1}(L_1)|, g_0^{-1} g_1 f_0^{-1}(L_1),$$
$$\frac{1}{g_1^{-1} g_0 f_1^{-1}(-L_1)}, f_0^{-1}(L_1) \Big) := L,$$

and

$$K_{1,\theta} = K_{1,\theta} \Big(L, \frac{1}{T}, \frac{1}{1-\tau}, f_0 \Big(\frac{1}{L} \Big)^- \Big) := K_1, \quad K_\theta = K_\theta(K_1, \beta_{K_1}) := K.$$

Next we obtain the gradient Hölder estimate with the help of Lemma 4.1. We remark that the operator (Q^θ, N^θ) is clearly elliptic and oblique, because it comes from (MFG$_\theta$). Moreover, since A^θ, b^θ, and B^θ and their derivatives are, respectively, continuous functions of (x, t, z, p, s, θ) on the compact sets $\overline{Q}_{T,K} \times [0, 1]$ and $\partial Q_{T,K} \times [0, 1]$, it follows that there exist constants $\mu_{L+K} > 0$, $\lambda_{L+K} > 0$, independent of θ, satisfying (2.5) in $(\mathbb{T}^d)_{L+K}$, and (2.6), (4.1) in $\partial Q_{T,L+K}$, when the operators (Q, N) are replaced by (Q^θ, N^θ). Lemma 4.1 then yields constants $C > 0$, $0 < \gamma < 1$, independent of θ, such that

$$[Du^\theta]_{\gamma, \overline{Q_T}} \leq C. \qquad \blacksquare$$

With the help of this uniform estimate, the main theorem for the strictly elliptic problem may now be proved.

Proof of Theorem 1.1. The uniqueness part of the statement is an immediate consequence of the standard Lasry–Lions monotonicity method, and will be omitted. We define the Banach spaces

$$E = C^{3,\alpha}(\overline{Q_T}), \quad F = C^{1,\alpha}(\overline{Q_T}) \times C^{2,\alpha}(\partial Q_T),$$

and the continuously differentiable operator $S: E \times [0, 1] \to F$ by

$$S(u, \theta) = (Q^\theta u, N^\theta u), \quad (u, \theta) \in E \times [0, 1].$$

The partial Fréchet derivative of S with respect to the variable u at the point (u, θ) is the corresponding linearization, for fixed θ, of the differential operator (Q^θ, N^θ), namely $(L^1_{(u,\theta)}, L^2_{(u,\theta)})$, where

$$L^1_{(u,\theta)}(w) = -\operatorname{Tr}(A^\theta(x, Du)D^2 w) - D_q \operatorname{Tr}(A^\theta(x, Du)D^2 u) \cdot Dw$$
$$+ D_q b^\theta(x, Du) \cdot Dw,$$

$$L^2_{(u,\theta)}(w) = \begin{cases} -w_t + D_p H^\theta(x, D_x u) \cdot D_x w & \text{if } t = 0, \\ \frac{g^\theta_m}{f_m}(w_t - D_p H^\theta \cdot D_x w) + w & \text{if } t = T. \end{cases}$$

For fixed $(u, \theta) \in E \times [0, 1]$, the linear operator $L^1_{(u,\theta)}$ is uniformly elliptic and the linear boundary operator $L^2_{(u,\theta)}$ is oblique. Moreover, the homogeneous problem $(L^1_{(u,\theta)}w, L^2_{(u,\theta)}w) = (0, 0)$ has the form

$$-\operatorname{Tr}(\tilde{A}(x,t)D^2 w) + \tilde{b}(x,t) \cdot Dw = 0 \text{ in } Q_T, \quad \tilde{B}(x,t) \cdot Dw + \tilde{c}(x,t)w = 0 \text{ on } \partial Q_T,$$

where $\tilde{B} \cdot \nu > 0$, $\tilde{c} \geq 0$, and $\tilde{c} \not\equiv 0$, which implies that it has only the trivial solution in $C^{3,\alpha}(\overline{Q_T})$. Hence, by the standard Fredholm alternative for linear oblique problems (see [9]), the operator $(L^1_{(u,\theta)}, L^2_{(u,\theta)})$ is invertible in $C^{3,\alpha}(\overline{Q_T})$. The infinite-dimensional implicit function theorem then implies that the set

$$D = \{\theta \in [0, 1] : \text{the equation } S(u, \theta) = (0, 0) \text{ has a unique solution } u \in C^{3,\alpha}(\overline{Q_T})\}$$

is open in $[0, 1]$.

The next step is to show that D is also closed. Let $\{\theta_n\} \subset D$ be a sequence such that $\theta_n \to \theta \in [0, 1]$, and let $\{u_n\} \subset E$ be the corresponding sequence of solutions to $S(u_n, \theta_n) = (0, 0)$. By Lemma 4.2, there exist numbers $C > 0$, $0 < \gamma < 1$, independent of n, such that

$$\|u_n\|_{C^{1,\gamma}(\overline{Q_T})} \leq C.$$

The Arzelà–Ascoli theorem implies that, up to a subsequence, there exists $u \in C^{1,\gamma}(\overline{Q_T})$ such that $u_n \to u$ in $C^1(\overline{Q_T})$. By Theorem 2.5, it follows that $u \in C^{2,\alpha}(\overline{Q_T})$, $u_n \to u$

in $C^{2,\alpha}(\overline{Q_T})$, and $S(u,\theta) = 0$. In particular, the u_n are uniformly bounded in $C^{2,\alpha}(\overline{Q_T})$. Now, given $i \in \{1, \ldots, d\}$, differentiating the equation $(Q^{\theta_n}(u_n), N^{\theta_n}(u_n)) = (0,0)$ yields, for $w = D_{x_i} u_n$,

$$L^1_{(u_n,\theta_n)} w = \mathrm{Tr}(A^{\theta_n}_{x_i}(x, Du_n) D^2 u_n) - b^{\theta_n}_{x_i}(x, Du_n),$$

$$L^2_{(u_n,\theta_n)} w = \begin{cases} D_{x_i}(f(x, m^{\theta_n}_0(x))) - H^{\theta_n}_{x_i} & \text{if } t = 0, \\ \frac{g^{\theta_n}_m}{f_m}(H^{\theta_n}_{x_i} - f_{x_i}) + g^{\theta_n}_{x_i} & \text{if } t = T. \end{cases}$$

Therefore, by Theorem 2.4, there exists $C > 0$, independent of n, such that

$$\|w\|_{C^{2,\alpha}(\overline{Q_T})} \le C,$$

implying that $D_x u_n$ is bounded in $C^{2,\alpha}(\overline{Q_T})$. In particular, $u_n|_{\partial Q_T}$ is bounded in $C^{3,\alpha}(\partial Q_T)$, and by the standard Schauder theory for the Dirichlet problem, u_n is therefore bounded in $C^{3,\alpha}(\overline{Q_T})$. Consequently, $u \in C^{3,\alpha}(\overline{Q_T})$ and $\theta \in D$, proving that D is closed. Since $0 \in D$, it follows that $D = [0,1]$, which completes the proof. ∎

The next theorem is the corresponding variant of Theorem 1.2 for the case of a fast-growing f, which follows from the estimates in Theorem 3.10.

Theorem 4.3. *If condition* (F2) *is replaced by* (HFX*)*, the conclusion of Theorem* 1.1 *holds.*

Proof. All of the results in this section follow in this case by simply replacing the use of Theorem 3.9 by Theorem 3.10. ∎

5. Weak solutions

In this section we develop the theory of weak solutions, for the case where the strict ellipticity condition (SE) fails to hold. We begin by stating the definition of weak solution that will be used, which is in direct analogy with the one used in [2–4] to study the degenerate case in which $g_m \equiv 0$.

Definition 5.1 (Definition of weak solution). A pair $(u, m) \in \mathrm{BV}(Q_T) \times L^\infty_+(Q_T)$ is called a weak solution to (MFG) if the following conditions hold:

(i) $D_x u \in L^2(Q_T), u \in L^\infty(Q_T), m \in C^0([0,T]; H^{-1}(\mathbb{T}^d)), m(\cdot, T) \in L^\infty(\mathbb{T}^d)$.

(ii) u satisfies the HJ inequality

$$-u_t + H(\cdot, D_x u) \le f(\cdot, m) \text{ in } Q_T, \quad u(\cdot, T) = g(\cdot, m(\cdot, T)) \text{ in } \mathbb{T}^d,$$

in the distributional sense, with $u(\cdot, T) = g(\cdot, m(\cdot, T))$ in the sense of traces.

(iii) m satisfies the continuity equation

$$m_t - \mathrm{div}(m D_p H(\cdot, D_x u)) = 0 \text{ in } Q_T, \quad m(\cdot, 0) = m_0 \text{ in } \mathbb{T}^d \tag{5.1}$$

in the distributional sense, with $m(\cdot, 0) = m_0$ in $H^{-1}(\mathbb{T}^d)$.

(iv)   The following identity holds:

$$\iint_{Q_T} m(x,t)(H(x, D_x u) - D_p H(x, D_x u) \cdot D_x u - f(x,m)) \, dx \, dt$$

$$= \int_{\mathbb{T}^d} (m(x,T) g(x, m(x,T)) - m_0(x) u(x,0)) \, dx. \tag{5.2}$$

The solutions to the degenerate elliptic problem will be obtained as a "vanishing viscosity" limit of MFG systems satisfying (SE), in the following sense. Assuming that (DE) holds, we consider, for $\epsilon > 0$, the system

$$\begin{cases} -u_t^\epsilon + H(\cdot, D_x u^\epsilon) = f(\cdot, m^\epsilon) + \epsilon \log(m^\epsilon), \quad u(\cdot, T) = g(\cdot, m^\epsilon(\cdot, T)), \\ m_t^\epsilon - \operatorname{div}(m^\epsilon D_p H(\cdot, D_x u^\epsilon)) = 0, \quad m^\epsilon(0) = m_0. \end{cases} \tag{MFG$_\epsilon$}$$

Since (MFG$_\epsilon$) is strictly elliptic, by Theorem 1.1 it has a unique solution $(u^\epsilon, m^\epsilon) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$. The only missing ingredient necessary to obtain a solution as the limit when $\epsilon \to 0$ is the following minor modification of Lemma 3.8, which provides a global, a priori upper bound for the density that is independent of the size of $\frac{1}{\min(m(T))}$.

Lemma 5.2. *Assume that* (SE) *holds. If* $(u, m) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ *is a solution to* (MFG)*, then*

$$\max_{\overline{Q_T}} f(\cdot, m(\cdot, \cdot)) \leq C,$$

where

$$C = C\left(C_0, T, \frac{1}{T}, \frac{1}{1-\tau}, \|u\|_{C^0(\overline{Q_T})}, \max_{\mathbb{T}^d} m(T), f_1\left(\max_{\mathbb{T}^d} m(T)\right)^+, \right.$$

$$\left. \left\|\frac{1}{\chi}\right\|_{C^0(\mathbb{T}^d \times [1, \infty))}, \|h_w\|_{C^0(\mathbb{T}^d \times [1, \infty))}\right).$$

Proof. The argument is a simple variant of the proof of Lemma 3.8. Let v and $\tilde{u}$ have the same meaning as in said proof, with $0 < c_1 < 1$ once more being a free parameter, set $\tilde{v} = -u_t + v = f + \frac{c_1}{2}\tilde{u}^2$, and let (x_0, t_0) be a point where $\tilde{v}$ achieves its maximum value.

Case 1. If $t_0 \in \{0, T\}$, then

$$\tilde{v}(x_0, t_0) = f(x_0, m(x_0, t_0)) + \frac{c_1}{2}\tilde{u}^2 \leq \max(f_1(\max m(T)), f_1(\max m_0)) + C \leq C.$$

Case 2. Assume next that $0 < t_0 < T$. Without loss of generality, it may be assumed that $f = f(x_0, m(x_0, t_0)) \geq 1$, because otherwise there would be nothing to prove. Therefore, using (H3), since

$$|-u_t + D_p H \cdot D_x u| \geq -u_t + D_p H \cdot D_x u \geq f + H,$$

it follows, by (2.1), that

$$|-u_t + D_p H \cdot D_x u|^2 \geq \frac{1}{2}|f|^2 + \frac{1}{2C_0^2}|D_x u|^4 - C. \tag{5.3}$$

Now, by Corollary 3.5, $L_u(\tilde{v}) = L_u(v)$, so, as in (3.33), Lemma 3.7 yields

$$0 \leq L_u(\tilde{v}) \leq \frac{-1}{2C_0}|-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2}|D_{xx}^2 u|^2 - \frac{c_1}{2}|-u_t + D_p H \cdot D_x u|^2$$
$$- \frac{c_1}{C_0}\chi|D_x u|^2 + C(1 + |D_x u|^{3+\tau} + \chi(1 + |D_x u|^{1+\tau}) + \chi^{(1+\tau)/2}|D_x u|^2$$
$$+ |f|(1 + |D_x u|^{1+\tau}) + |D_x u|^2|f|^{(1+\tau)/2})$$
$$+ Cc_1(\chi|D_{xx}^2 u|^2 + |-D_x u_t + D_{xx}^2 u D_p H|^2).$$

Thus, by (F2), (5.3), and the fact that $f \geq 1$,

$$0 \leq \frac{-1}{2C_0}|-D_x u_t + D_p H D_{xx}^2 u|^2 - \frac{\chi}{2C_0^2}|D_{xx}^2 u|^2 - \frac{c_1}{4}\left(\frac{1}{C_0^2}|D_x u|^4 + f^2\right)$$
$$- \frac{c_1}{C_0}\chi|D_x u|^2 + C(1 + |D_x u|^{3+\tau} + (1+f)(1 + |D_x u|^{1+\tau})$$
$$+ (1+f)^{(1+\tau)/2}|D_x u|^2 + f(1+|D_x u|^{1+\tau}) + |D_x u|^2 f^{(1+\tau)/2})$$
$$+ Cc_1(\chi|D_{xx}^2 u|^2 + |-D_x u_t + D_{xx}^2 u D_p H|^2).$$

Once more, as in Lemma 3.8, fix c_1 such that $c_1 < \frac{1}{4C_0 C(1+C_0)}$, where the constant C is as in the previous line, obtaining

$$\frac{c_1}{4}\left(\frac{1}{C_0}|D_x u|^4 + f^2\right) \leq C(1 + |D_x u|^{3+\tau} + (1 + f)(1 + |D_x u|^{1+\tau})$$
$$+ (1 + f)^{(1+\tau)/2}|D_x u|^2 + f(1 + |D_x u|^{1+\tau})$$
$$+ |D_x u|^2 f^{(1+\tau)/2}).$$

The left- and right-hand sides have, respectively, degree 4 and degree $3 + \tau < 4$ in the non-negative variables $(|D_x u|, \sqrt{f})$, thus $|D_x u|^2 + f \leq C$, and, in particular, it follows that

$$\tilde{v}(x_0, t_0) \leq C. \qquad \blacksquare$$

We now obtain several a priori bounds for (u^ϵ, m^ϵ) that are independent of ϵ.

Lemma 5.3. *Assume* (DE)*, and let* $(u^\epsilon, m^\epsilon) \in C^{3,\alpha}(\overline{Q_T}) \times C^{2,\alpha}(\overline{Q_T})$ *be the solution to* (MFG$_\epsilon$)*. Then there exist constants* $L, L_1, C,$ *and* $C_1,$ *with*

$$L = L(L_1, |g_1 f_0^{-1}(L_1)|, g_0^{-1} g_1 f_0^{-1}(L_1)),$$
$$L_1 = \left(C_0, T, \left|\min_{\mathbb{T}^d} f(\cdot, 0)\right|, \left|\min_{\mathbb{T}^d} g(\cdot, 0)\right|\right),$$
$$C = C(C_1, f_1(C_1)^+),$$
$$C_1 = C_1\left(L, \frac{1}{T}, \frac{1}{1-\tau}, \left\|\frac{1}{\chi}\right\|_{C^0(\mathbb{T}^d \times [1+\max_{\mathbb{T}^d} f^+(\cdot, 0), \infty))},$$
$$\|h_w\|_{C^0(\mathbb{T}^d \times [1+\max_{\mathbb{T}^d} f^+(\cdot, 0), \infty))}\right),$$

such that, for every $\epsilon \leq \frac{1}{C}$,

$$\|u^\epsilon\|_{C^0(\overline{Q_T})} \leq L, \quad \|m^\epsilon\|_{C^0(\overline{Q_T})} + \|u_t^{-,\epsilon}\|_{C^0(\overline{Q_T})} \leq C, \tag{5.4}$$

$$\|u_t^\epsilon\|_{L^1(Q_T)} + \|D_x u^\epsilon\|_{L^2(Q_T)} + \epsilon\|\log m^\epsilon\|_{L^1(Q_T)} \leq C. \tag{5.5}$$

Proof. By replacing f, H with $f - C$, $H - C$, for C depending only on C_0, there is no loss of generality in assuming $f(\cdot, 0) \leq -1$. It is readily seen that conditions (F2), (FX1), (FX2), and (2.4) hold for $f^\epsilon = f + \epsilon \log(\cdot)$, uniformly in ϵ, up to increasing C_0 by a finite value. By Lemma 3.1, there exists L_1 such that

$$g_0(f_1^\epsilon)^{-1}(-L_1) - L_1(e^{L_1 T} - e^{L_1 t}) \leq u^\epsilon \leq g_1(f_0^\epsilon)^{-1}(L_1) + L_1(e^{L_1 T} - e^{L_1 t}).$$

Now, (DE) implies that

$$-L - L_1(e^{L_1 T} - e^{L_1 t}) \leq u^\epsilon.$$

On the other hand, if $L_1 > f_0(1)$, then $f_0^{-1}(L_1) > 1$, hence

$$f_0^\epsilon(f_0^{-1}(L_1)) = L_1 + \epsilon \log(f_0^{-1}(L_1)) \geq L_1,$$

which implies

$$g_1 f_0^{-1}(L_1) \geq g_1(f_0^\epsilon)^{-1}(L_1). \tag{5.6}$$

Consequently,

$$u^\epsilon \leq g_1 f_0^{-1}(L_1) + L_1(e^{L_1 T} - e^{L_1 t}),$$

which proves the first inequality in (5.4). Now, by Corollary 3.2 and (5.6),

$$m^\epsilon(\cdot, T) \leq L.$$

Thus, Lemma 5.2 implies that

$$f(\cdot, m^\epsilon) + \epsilon \log m^\epsilon(\cdot, \cdot) = f^\epsilon(\cdot, m^\epsilon) \leq C, \tag{5.7}$$

and (F1) yields, for C depending only on $\|\frac{1}{\chi}\|_{C^0(\mathbb{T}^d \times [1,\infty))}$ and $|\min f(\cdot, 0)|$,

$$\frac{1}{C}\log(m^\epsilon(\cdot, \cdot)) - C \leq f(\cdot, m^\epsilon(\cdot, \cdot)).$$

Therefore, by (5.7) we conclude that $\|m^\epsilon\|_{C^0(\overline{Q_T})} \leq C$ for large enough C and small enough ϵ. The lower bound for u_t^ϵ is simply a consequence of (5.7), the relation $-u_t^\epsilon + H = f^\epsilon$, and the fact that H is bounded below. This completes the proof of (5.4).

Now, integrating the HJ equation for u^ϵ yields

$$\iint_{Q_T} H(\cdot, D_x u^\epsilon) + \iint_{Q_T} \epsilon(\log m^\epsilon)^-$$

$$= \iint_{Q_T} (f(\cdot, m^\epsilon) + \epsilon(\log m^\epsilon)^+) + \int_{\mathbb{T}^d} (u^\epsilon(T) - u^\epsilon(0)). \tag{5.8}$$

It then follows from (5.4) and (2.1) that $\|D_x u^\epsilon\|_{L^2(Q_T)}$ and $\epsilon\|\log m^\epsilon\|_{L^1(Q_T)}$ are bounded. Finally,

$$\left|\iint_{Q_T} u_t^\epsilon\right| = \left|\int_{\mathbb{T}^d} u^\epsilon(T) - u^\epsilon(0)\right| \le C,$$

and since u_t^ϵ is bounded below, this proves (5.5). ∎

After extracting a subsequence, Lemma 5.3 implies the existence of $(u, m) \in$ $BV(Q_T) \times L_+^\infty(Q_T)$ such that, as $\epsilon \to 0$,

$$u^\epsilon \to u \text{ in } L^1(Q_T) \text{ and pointwise a.e.,} \quad m^\epsilon \overset{*}{\rightharpoonup} m \text{ in } L^\infty(Q_T),$$
$$u_t^\epsilon \overset{*}{\rightharpoonup} u_t \text{ in } C^0(\overline{Q_T})^*, \quad D_x u^\epsilon \rightharpoonup D_x u \text{ in } L^2(Q_T). \tag{5.9}$$

We now show that, up to a further subsequence, this convergence can be strengthened.

Lemma 5.4. *Assume that* (DE) *holds, let* $(u^\epsilon, m^\epsilon) \in C^3(\overline{Q_T}) \times C^2(\overline{Q_T})$ *be the solution to* (MFG$_\epsilon$), *and let* $(u, m) \in BV(Q_T) \times L_+^\infty(Q_T)$ *be a subsequential limit. Then, up to extracting a subsequence,*

$$m^\epsilon \to m \text{ in } C^0([0, T], H^{-1}(\mathbb{T}^d)) \text{ and a.e. in } Q_T, \tag{5.10}$$
$$m^\epsilon(\cdot, T) \to m(\cdot, T), \quad u^\epsilon(\cdot, T) \to u(\cdot, T) \text{ a.e. in } \mathbb{T}^d, \tag{5.11}$$
$$D_x u^\epsilon \to D_x u \text{ in } L_m^2(Q_T) \text{ and a.e. in } \{m > 0\},$$
$$m^\epsilon|D_x u^\epsilon|^2 \to m|D_x u|^2 \text{ in } L^1(Q_T) \text{ and a.e. in } Q_T, \tag{5.12}$$
$$\epsilon \log m^\epsilon \to 0 \text{ in } L_m^1(Q_T). \tag{5.13}$$

Moreover, $u(\cdot, T) = g(\cdot, m(\cdot, T))$ *and* $m \in C^{0,\frac{1}{2}}([0, T], H^{-1}(\mathbb{T}^d))$, *with*

$$[m]_{\frac{1}{2},[0,T],H^{-1}} \le \|m D_p H(\cdot, D_x u)\|_{L^2(Q_T)},$$

and, for almost every $s \in [0, T]$, *including* $s = 0$,

$$\int_s^T \int_{\mathbb{T}^d} m(x, t)(H(x, D_x u) - D_p H(x, D_x u) \cdot D_x u - f(x, m)) \, dx \, dt$$
$$= \int_{\mathbb{T}^d} (m(x, T)g(x, m(x, T)) - m(x, s)u(x, s)) \, dx. \tag{5.14}$$

Proof. Let $\epsilon, \epsilon' > 0$. We employ the standard Lasry–Lions method with the pairs (u^ϵ, m^ϵ) and $(u^{\epsilon'}, m^{\epsilon'})$. Namely, subtracting the HJ and continuity equations from each other, respectively, yields a new system for $u^\epsilon - u^{\epsilon'}$ and $m^\epsilon - m^{\epsilon'}$. Multiplying the equations satisfied by $u^\epsilon - u^{\epsilon'}$ and $m^\epsilon - m^{\epsilon'}$, respectively, by $m^\epsilon - m^{\epsilon'}$ and $u^\epsilon - u^{\epsilon'}$, and then integrating over Q_T, leads to the identity

$$M_{\epsilon,\epsilon'} + M_{\epsilon',\epsilon} + M_g + M_f + K_{\epsilon,\epsilon'}^- + K_{\epsilon',\epsilon}^- = K + K_{\epsilon,\epsilon'}^+ + K_{\epsilon',\epsilon}^+,$$

where

$$M_{\epsilon,\epsilon'} = \iint_{Q_T} m^\epsilon (H^{\epsilon'} - H^\epsilon - D_p H^\epsilon \cdot (D_x u^{\epsilon'} - D_x u^\epsilon)),$$

$$M_g = \int_{\mathbb{T}^d} (g(\cdot, m^\epsilon(\cdot, T)) - g(\cdot, m^{\epsilon'}(\cdot, T)))(m^\epsilon(\cdot, T) - m^{\epsilon'}(\cdot, T)),$$

$$M_f = \iint_{Q_T} (f(\cdot, m^\epsilon) - f(\cdot, m^{\epsilon'}))(m^\epsilon - m^{\epsilon'}),$$

$$K^{\pm}_{\epsilon,\epsilon'} = \iint_{Q_T} \epsilon' (\log m^{\epsilon'})^{\pm} m^\epsilon, \quad K = \iint_{Q_T} -(\epsilon' \log m^\epsilon m^{\epsilon'} + \epsilon \log m^\epsilon m^\epsilon),$$

and $H^\epsilon = H(\cdot, D_x u^\epsilon)$, $D_p H^\epsilon = D_p H(\cdot, D_x u^\epsilon)$. By (5.4), and the fact that the map $m \mapsto m \log m$ is bounded below, $K \to 0$ as $\epsilon, \epsilon' \to 0$. Similarly, (5.4) implies that the quantity $\log m^{\epsilon'} m^\epsilon$ is bounded above, and thus $K^+_{\epsilon,\epsilon'} \to 0$ as $\epsilon, \epsilon' \to 0$. Since, by monotonicity and convexity, each term on the left-hand side of the equation is non-negative, all of them converge to zero. Every claim of a.e. convergence in what follows is tacitly meant to hold after extracting a subsequence. Since $M_f \to 0$, we have $m^\epsilon \to m$ a.e. in Q_T. Similarly, since $M_g \to 0$, there exists a function $m_T \in L^\infty(\mathbb{T}^d)$ such that $m^\epsilon(T) \to m_T$ a.e. in $\mathbb{T}^d$. Now, using (H1), one obtains

$$\frac{1}{2C_0} \iint_{Q_T} (m^\epsilon + m^{\epsilon'})|D_x u^\epsilon - D_x u^{\epsilon'}|^2 \leq M_{\epsilon,\epsilon'} + M_{\epsilon',\epsilon},$$

which implies (5.12). Finally, sending $\epsilon' \to 0$ first and then $\epsilon \to 0$, using the fact that $K^-_{\epsilon',\epsilon} \to 0$, yields (5.13).

Next we show the continuity properties of m, which in particular give a meaning to the expression $m(\cdot, T)$. Integrating the equation for $m^\epsilon - m^{\epsilon'}$ over a cylinder $\mathbb{T}^d \times [0, t]$ against a test function $\phi \in C^\infty(\mathbb{T}^d)$ yields

$$\int_{\mathbb{T}^d} (m^\epsilon - m^{\epsilon'})\phi = -\int_0^t \int_{\mathbb{T}^d} (m^\epsilon D_p H(\cdot, D_x u^\epsilon) - m^{\epsilon'} D_p H(\cdot, D_x u^{\epsilon'})) \cdot D\phi,$$

which implies

$$\sup_{t \in [0,T]} \|m^\epsilon(\cdot, t) - m^{\epsilon'}(\cdot, t)\|_{H^{-1}(\mathbb{T}^d)}$$

$$\leq \sqrt{T} \|m^\epsilon D_p H(\cdot, D_x u^\epsilon) - m^{\epsilon'} D_p H(\cdot, D_x u^{\epsilon'})\|_{L^2(Q_T)}.$$

Consequently, (2.1) and (5.12) together imply that $m^\epsilon \to m$ in $C^0([0, T], H^{-1}(\mathbb{T}^d))$. Similarly, testing ϕ against the continuity equation for m^ϵ in a cylinder $\mathbb{T}^d \times [s, t]$ yields

$$\int_{\mathbb{T}^d} (m^\epsilon(x, t) - m^\epsilon(x, s))\phi(x)\, dx = -\int_s^t \int_{\mathbb{T}^d} m^\epsilon D_p H(x, D_x u^\epsilon) \cdot D\phi(x)\, dx\, dt,$$

thus

$$\|m^\epsilon(\cdot, t) - m^\epsilon(\cdot, s)\|_{H^{-1}(\mathbb{T}^d)} \leq \sqrt{t - s} \|m^\epsilon D_p H(\cdot, D_x u^\epsilon)\|_{L^2(Q_T)},$$

and sending $\epsilon \to 0$ produces the desired Hölder estimate.

Since $m^\epsilon(\cdot, T) \to m_T$ a.e. in $\mathbb{T}^d$, and $m^\epsilon(\cdot, T) \to m(\cdot, T)$ in $H^{-1}(\mathbb{T}^d)$, it follows that $m_T = m(\cdot, T)$, and that $u^\epsilon(\cdot, T) = g(\cdot, m^\epsilon(\cdot, T)) \to g(\cdot, m(\cdot, T))$ almost everywhere. Moreover, (5.9) implies that $u^\epsilon(\cdot, T) \overset{*}{\rightharpoonup} u(\cdot, T)$ in $C^\infty(\mathbb{T}^d)^*$, and therefore $u(\cdot, T) = g(\cdot, m(\cdot, T))$.

It remains to show (5.14). For this purpose, performing once more the Lasry–Lions computation, this time on the system for (u^ϵ, m^ϵ), and integrating on $\mathbb{T}^d \times [s, T]$ yields

$$\int_s^T \int_{\mathbb{T}^d} m^\epsilon(x, t)(H^\epsilon - D_p H^\epsilon \cdot D_x u^\epsilon - f(x, m^\epsilon) - \epsilon \log m^\epsilon)\, dx\, dt$$
$$= \int_{\mathbb{T}^d} (m^\epsilon(x, T)g(x, m^\epsilon(x, T)) - m^\epsilon(x, s)u^\epsilon(x, s))\, dx.$$

By Fubini's theorem, for a.e. $s \in [0, T]$, $u^\epsilon(\cdot, s)$ and $m^\epsilon(\cdot, s)$ converge, respectively, to $u(\cdot, s) \in L^\infty(\mathbb{T}^d)$ and $m(\cdot, s) \in L^\infty(\mathbb{T}^d)$ a.e. in $\mathbb{T}^d$. Thus, for such s, using (5.12), one obtains (5.14) after letting $\epsilon \to 0$. When $s = 0$, $m^\epsilon(s) = m_0$, and thus the $C^\infty(\mathbb{T}^d)^*$-convergence of $u^\epsilon(0)$ is sufficient to conclude. ∎

We now prove the main result for the degenerate elliptic problem, Theorem 1.2.

Proof of Theorem 1.2. First we will establish that (u, m) is indeed a weak solution. By Lemmas 5.3 and 5.4, (u, m) satisfies condition (i) of Definition 5.1. Next, by the HJ equation for u^ϵ,

$$-u_t^\epsilon + H^\epsilon = f(\cdot, m^\epsilon) + \epsilon \log m^\epsilon \le f(\cdot, m^\epsilon) + \epsilon(\log m^\epsilon)^+.$$

Integration against a non-negative function $\phi \in C^\infty(\overline{Q_T})$ then yields

$$\int_{\mathbb{T}^d} (-u^\epsilon(\cdot, T)\phi(\cdot, T) + u^\epsilon(\cdot, 0)\phi(\cdot, 0)) + \iint_{Q_T} u^\epsilon \phi_t + \iint_{Q_T} H^\epsilon \phi$$
$$\le \iint_{Q_T} f(\cdot, m^\epsilon)\phi + \epsilon(\log m^\epsilon)^+ \phi. \tag{5.15}$$

Using the convexity of H,

$$\iint_{Q_T} H(\cdot, D_x u)\phi + D_p H(\cdot, D_x u) \cdot (D_x u^\epsilon - D_x u)\phi \le \iint_{Q_T} H(\cdot, D_x u^\epsilon)\phi,$$

and therefore, since $D_x u^\epsilon \rightharpoonup D_x u$ in $L^2(Q_T)$,

$$\iint_{Q_T} H\phi \le \liminf_{\epsilon \to 0^+} \iint_{Q_T} H^\epsilon \phi.$$

Letting $\epsilon \to 0$ in (5.15), Lemmas 5.3 and 5.4 and (5.9) thus yield

$$\int_{\mathbb{T}^d} (-u(\cdot, T)\phi(\cdot, T) + u(\cdot, 0)\phi(\cdot, 0)) + \iint_{Q_T} u\phi_t + \iint_{Q_T} H\phi \le \iint_{Q_T} f\phi.$$

This completes the proof of condition (ii) in Definition 5.1. The third condition follows immediately by testing the continuity equation of m^ϵ against an arbitrary $\phi \in C^\infty(\overline{Q_T})$, and appealing to Lemma 5.4 to let $\epsilon \to 0$ in the equality

$$\int_{\mathbb{T}^d} (m^\epsilon(\cdot, T)\phi(\cdot, T) - m_0\phi(\cdot, 0)) - \iint_{Q_T} m^\epsilon \phi_t + \iint_{Q_T} m^\epsilon D_p H(\cdot, D_x u^\epsilon) \cdot D_x \phi = 0.$$

Finally, condition (iv) of Definition 5.1 has already been established in Lemma 5.4.

Next is the proof of uniqueness. Let (u', m') be another weak solution to (MFG). By the fact that every non-negative distribution can be identified with a non-negative measure, the HJ inequality for u' may be written as

$$-u'_t + H(\cdot, D_x u') + \mu = f(\cdot, m'),$$

where $\mu \in C(\overline{Q_T})^*$ is a non-negative, finite measure on $\overline{Q_T}$. We carefully apply the Lasry–Lions procedure to the (u^ϵ, m^ϵ) and (u', m') systems. Set $v^\epsilon = u^\epsilon - u'$, $w^\epsilon = m^\epsilon - m'$. Subtracting the two corresponding HJ equations for u', u^ϵ, and integrating against m^ϵ, we obtain

$$\iint_{Q_T} (v^\epsilon_t m^\epsilon + (H' - H^\epsilon)m^\epsilon) + \iint_{Q_T} m^\epsilon d\mu = \iint_{Q_T} m^\epsilon(f' - f^\epsilon), \qquad (5.16)$$

where $H' = H(\cdot, D_x u')$, $f' = f(\cdot, m')$. Integrating the HJ equation for u^ϵ against m' gives

$$\iint_{Q_T} (-u^\epsilon_t m' + H^\epsilon m') = \iint_{Q_T} m' f^\epsilon.$$

Now subtracting the continuity equations for m^ϵ, m' and testing against u^ϵ yields

$$\iint_{Q_T} (-u^\epsilon_t m^\epsilon + u^\epsilon_t m') + (m^\epsilon D_p H^\epsilon - m' D_p H') \cdot D_x u^\epsilon = -\int_{\mathbb{T}^d} v^\epsilon(T) m'(T), \quad (5.17)$$

and integrating the continuity equation for m^ϵ against $-u'$ we get

$$\iint_{Q_T} (m^\epsilon u'_t - m^\epsilon D_p H^\epsilon D_x u') = \int_{\mathbb{T}^d} (m^\epsilon(T)u'(T) - m_0 u'(0)). \qquad (5.18)$$

Finally, by condition (iv) in Definition 5.1 and the fact that (u', m') is a weak solution,

$$\iint_{Q_T} -m'(H' - D_p H' \cdot D_x u') = -\iint_{Q_T} m' f' - \int_{\mathbb{T}^d} (m'(T)u'(T) - m_0 u'(0)). \quad (5.19)$$

Adding (5.16), (5.17), (5.18), and (5.19) yields

$$M_{\epsilon,1} + M_{\epsilon,2} + M_{g,\epsilon} + M_{f,\epsilon} + K^-_\epsilon + K_{1,\epsilon} = K_{2,\epsilon} + K^+_\epsilon, \qquad (5.20)$$

where

$$M_{\epsilon,1} = \iint_{Q_T} m^\epsilon (H' - H^\epsilon - D_p H' \cdot (D_x u^\epsilon - D_x u')),$$

$$M_{\epsilon,2} = \iint_{Q_T} m' (H^\epsilon - H' - D_p H^\epsilon \cdot (D_x u' - D_x u^\epsilon)),$$

$$M_{g,\epsilon} = \int_{\mathbb{T}^d} (m^\epsilon(\cdot, T) - m'(\cdot, T))(g(\cdot, m^\epsilon(\cdot, T)) - g(\cdot, m'(\cdot, T))),$$

$$M_{f,\epsilon} = \iint_{Q_T} (f(\cdot, m^\epsilon) - f(\cdot, m'))(m^\epsilon - m'),$$

$$K_\epsilon^\pm = \iint_{Q_T} \epsilon (\log m^\epsilon)^\pm m', \quad K_{1,\epsilon} = \iint_{Q_T} m^\epsilon d\mu, \quad K_{2,\epsilon} = -\iint_{Q_T} \epsilon \log m^\epsilon m^\epsilon.$$

The only new term relative to the proof of Lemma 5.4 is $K_{1,\epsilon}$, which is clearly non-negative. Thus, as before, each individual term on the left-hand side of (5.20) converges to zero as $\epsilon \to 0$. In particular, since $M_{f,\epsilon} \to 0$, it follows that $m = m'$ a.e. A posteriori, since $M_{\epsilon,2} \to 0$, (5.12) and the strict convexity of H imply that $D_x u = D_x u'$ a.e. in $\{m > 0\}$. Moreover, $M_{g,\epsilon} \to 0$ implies $m(\cdot, T) = m'(\cdot, T)$ a.e., and thus $u(\cdot, T) = g(\cdot, m(\cdot, T)) = g(\cdot, m'(\cdot, T)) = u'(\cdot, T)$ a.e. in $\mathbb{T}^d$.

It remains to show that $u(s) = u'(s)$ a.e. in $\{m(s) > 0\}$, for a.e. $s \in [0, T]$ including the case $s = 0$. The function $\bar{u} = \max(u, u')$ is in $\mathrm{BV}(Q_T) \cap L^\infty(Q_T)$, and, since $D_x u, D_x u' \in L^2(Q_T)$, the chain rule yields $D_x \bar{u} = D_x u$ a.e. in $\{u \geq u'\}$ and $D_x \bar{u} = D_x u'$ a.e. in $\{u \leq u'\}$. In particular, $D_x \bar{u} = D_x u$ a.e. in $\{m > 0\}$. Following [4, Thm. 5.2], through a mollification procedure, the theory of viscosity solutions implies that $\bar{u}$ is a distributional subsolution to the HJ equation. Therefore, testing the HJ inequality of $\bar{u}$ against m^ϵ in an interval $[s, T]$,

$$\int \bar{u}(s) m^\epsilon(s)$$

$$\leq \int_{\mathbb{T}^d} \bar{u}(T) m^\epsilon(T) + \int_s^T \int_{\mathbb{T}^d} (-m_t^\epsilon \bar{u} - m^\epsilon H(x, D_x \bar{u})) + m^\epsilon f(x, m) \, dx \, dt$$

$$= \int_{\mathbb{T}^d} u(T) m^\epsilon(T)$$

$$+ \int_s^T \int_{\mathbb{T}^d} (-\operatorname{div}(m^\epsilon D_p H^\epsilon) \bar{u} - m^\epsilon H(x, D_x \bar{u}) + m^\epsilon f(x, m)) \, dx \, dt$$

$$= \int_{\mathbb{T}^d} u(T) m^\epsilon(T) + \int_s^T \int_{\mathbb{T}^d} m^\epsilon (D_p H^\epsilon \cdot D_x \bar{u} - H(x, D_x \bar{u}) + f(x, m)) \, dx \, dt.$$

Using Lemma 5.4 and the dominated convergence theorem to let $\epsilon \to 0$,

$$\int_{\mathbb{T}^d} \bar{u}(s) m(s)$$

$$\leq \int_{\mathbb{T}^d} u(T) m(T) + \int_s^T \int_{\mathbb{T}^d} m(D_p H \cdot D_x \bar{u} - H(x, D_x \bar{u}) + f(x, m)) \, dx \, dt$$

$$= \int_{\mathbb{T}^d} u(T)m(T) + \int_s^T \int_{\mathbb{T}^d} m(D_p H \cdot D_x u - H(x, D_x u) + f(x, m)) \, dx \, dt$$

$$= \int_{\mathbb{T}^d} u(s)m(s)$$

for a.e. $s \in [0, T]$, including $s = 0$, where (5.14) was used in the last equality. Given that $\bar{u} \geq u$, this implies $\bar{u}(s) = u(s)$ a.e. in $\{m(s) > 0\}$ for such s. At last, applying the Lasry–Lions method to the system for $(u^\epsilon - u', m^\epsilon)$, over an interval $[s, T]$, yields

$$\int_s^T \int_{\mathbb{T}^d} ((H' - H^\epsilon)m^\epsilon + m^\epsilon D_p H^\epsilon \cdot (D_x u^\epsilon - D_x u')) + \int_s^T \int_{\mathbb{T}^d} m^\epsilon \, d\mu$$

$$+ \int_s^T \int_{\mathbb{T}^d} m^\epsilon (f^\epsilon - f') + \int_{\mathbb{T}^d} (m^\epsilon(T)v^\epsilon(T) - m^\epsilon(s)v^\epsilon(s)) = 0.$$

Since $\int_s^T \int_{\mathbb{T}^d} m^\epsilon \, d\mu \leq \iint_{Q_T} m^\epsilon \, d\mu = K_1 \to 0$, sending $\epsilon \to 0$ results in

$$\int_{\mathbb{T}^d} m(T)(u(T) - u(T)) - m(s)(u(s) - u'(s)) = 0,$$

that is,

$$\int_{\mathbb{T}^d} m(s)u(s) = \int_{\mathbb{T}^d} m(s)u'(s).$$

By the fact that $u(s) = \bar{u}(s) \geq u'(s)$ a.e. in $\{m(s) > 0\}$, one then concludes that $u = u'$ a.e. in $\{m > 0\}$ and, since $m_0 > 0$, we have $u(0) = u'(0)$ a.e. in $\mathbb{T}^d$. ∎

Next we prove that when the data is independent of the space variable, the solution u is Lipschitz continuous.

Theorem 5.5. *Assume that* (DE) *holds, and let* H, f, *and g be independent of* x. *Then the sequence* u^ϵ *is uniformly bounded in* $C^1(\overline{Q_T})$ *as* $\epsilon \to 0$. *In particular, the weak solution* $u = \lim_{\epsilon \to 0} u^\epsilon$ *and the terminal condition* $m(\cdot, T)$ *are globally Lipschitz continuous.*

Proof. There is no loss of generality in assuming $f(0) < -1$. By Remark 3.3,

$$\min m_0 \leq m^\epsilon(\cdot, T) \leq \max m_0. \tag{5.21}$$

We set, for $K > 0$,

$$\chi^\epsilon(w) = (f^\epsilon)^{-1}(w) \cdot f^\epsilon_m((f^\epsilon)^{-1}(w)), \quad h^\epsilon = \sqrt{\chi^\epsilon},$$

$$\beta_K^\epsilon = \|f^\epsilon\|_{C^1([\frac{1}{K}, K])} + \|g\|_{C^1([\frac{1}{K}, K])} + \left\|\frac{1}{\chi^\epsilon}\right\|_{C^0([-K, \infty))} + \|h_w^\epsilon\|_{C^0([-K, \infty))}.$$

Then, in view of (DE), (5.21), and Lemma 3.8, there exist constants C and C_1, with $C = C(C_1, \beta_{C_1}^\epsilon)$ and $C_1 = C_1(C_0, T, T^{-1}, (1 - \tau)^{-1}, \|u^\epsilon\|_{C^0(\overline{Q_T})})$, such that

$$\|Du^\epsilon\|_{C^0(\overline{Q_T})} \leq C.$$

The only issue here is that the quantities

$$K_{1,\epsilon}(-C_1) = \left\| \frac{1}{\chi^\epsilon} \right\|_{C^0([-C_1,\infty))} \quad \text{and} \quad K_{2,\epsilon}(-C_1) = \|h_w^\epsilon\|_{C^0([-C_1,\infty))}$$

may not be bounded independently of ϵ. However, the proof of Lemma 3.8 shows that, defining $\tilde{u}^\epsilon$ as in said proof, the gradient bound depends only on $K_{1,\epsilon}(a)$ and $K_{2,\epsilon}(a)$, where $a \in \mathbb{R}$ is any number satisfying the following condition: for all small enough $0 \le c_1 < 1$, at any maximum point (x_0, t_0) of $H(D_x u^\epsilon) + \frac{c_1}{2}(\tilde{u}^\epsilon)^2$, the inequality $f^\epsilon(m(x_0, t_0)) \ge a$ holds. At such a point (x_0, t_0), for small enough ϵ and c_1, Corollary 3.5 yields

$$f^\epsilon(m(x_0, t_0)) = -u_t^\epsilon + H(D_x u^\epsilon) \ge f^\epsilon(\min m_0) - \|H(D_x u^\epsilon)\|_{C^0} + H(D_x u^\epsilon)$$

$$\ge f^\epsilon(\min m_0) - \frac{c_1}{2}\|\tilde{u}^\epsilon\|_{C^0}^2 = f(\min m_0) + \epsilon \log \min m_0 - c_1\|\tilde{u}^\epsilon\|_{C^0}^2$$

$$> f\left(\frac{1}{2}\min m_0\right).$$

Thus, the condition is satisfied by $a = f(\frac{1}{2}\min m_0)$. Since $a > f(0)$, it follows from (F1), (F2) that $K_{1,\epsilon}(a)$ and $K_{2,\epsilon}(a)$ are bounded uniformly as $\epsilon \to 0$. The Arzelà–Ascoli theorem implies the result. ∎

We finally note that, in the case $d = 1$, since there exists an a priori lower bound for the density m in terms of its boundary values (obtained in [1, 11, 15]), the solutions are seen to be smooth. However, when $d > 1$, even in the special case of Theorem 5.5, where an a priori bound for the gradient was obtained, we do not know whether the solution to the degenerate elliptic problem enjoys higher regularity.

References

[1] T. Bakaryan, R. Ferreira, and D. Gomes, Some estimates for the planning problem with potential. *NoDEA Nonlinear Differential Equations Appl.* **28** (2021), no. 2, Paper No. 20, 23 Zbl 1461.91035   MR 4232900

[2] P. Cardaliaguet, Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications*, pp. 111–158, Springer INdAM Ser. 11, Springer, Cham, 2015   Zbl 1329.49068   MR 3408214

[3] P. Cardaliaguet and P. J. Graber, Mean field games systems of first order. *ESAIM Control Optim. Calc. Var.* **21** (2015), no. 3, 690–722   Zbl 1319.35273   MR 3358627

[4] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon, Second order mean field games with degenerate diffusion and local coupling. *NoDEA Nonlinear Differential Equations Appl.* **22** (2015), no. 5, 1287–1317 Zbl 1344.49061 MR 3399179.

[5] P. Cardaliaguet and A. Porretta, An introduction to mean field game theory. In *Mean field games*, pp. 1–158, Lecture Notes in Math. 2281, Springer, Cham, 2020 Zbl 1457.91058 MR 4214774

[6] L. C. Evans, Some new PDE methods for weak KAM theory. *Calc. Var. Partial Differential Equations* **17** (2003), no. 2, 159–177 Zbl 1032.37048 MR 1986317

[7] R. Fiorenza, Sui problemi di derivata obliqua per le equazioni ellittiche. *Ricerche Mat.* **8** (1959), 83–110 Zbl 0090.31404 MR 114995

[8] R. Fiorenza, Sulla hölderianitá delle solutzioni dei problemi di derivata obliqua regolare del secondo ordine. *Ricerche Mat.* **14** (1965), 102–123 Zbl 0134.09201 MR 201796

[9] D. Gilbarg and N. S. Trudinger, Classical solutions; the Schauder approach. In *Elliptic partial differential equations of second order*, pp. 120–130, Classics in Mathematics, Springer, Berlin, 2001 Zbl 1042.35002 MR 1814364

[10] D. A. Gomes, H. Mitake, and K. Terai, The selection problem for some first-order stationary mean-field games. *Netw. Heterog. Media* **15** (2020), no. 4, 681–710 Zbl 1455.35261 MR 4175148

[11] D. A. Gomes and T. Seneci, Displacement convexity for first-order mean-field games. *Minimax Theory Appl.* **3** (2018), no. 2, 261–284 Zbl 1419.91102 MR 3882532

[12] P. J. Graber, Optimal control of first-order Hamilton-Jacobi equations with linearly bounded Hamiltonian. *Appl. Math. Optim.* **70** (2014), no. 2, 185–224 Zbl 1305.49035 MR 3255706

[13] M. Huang, R. P. Malhamé, and P. E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006), no. 3, 221–251 Zbl 1136.91349 MR 2346927

[14] J.-M. Lasry and P.-L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007), no. 1, 229–260 Zbl 1156.91321 MR 2295621

[15] H. Lavenant and F. Santambrogio, Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational mean field games. *Comm. Partial Differential Equations* **43** (2018), no. 12, 1761–1802 Zbl 1414.49026 MR 3927372

[16] G. M. Lieberman, Solvability of quasilinear elliptic equations with nonlinear boundary conditions. *Trans. Amer. Math. Soc.* **273** (1982), no. 2, 753–765 Zbl 0498.35039 MR 667172

[17] G. M. Lieberman, The nonlinear oblique derivative problem for quasilinear elliptic equations. *Nonlinear Anal.* **8** (1984), no. 1, 49–65 Zbl 0541.35032 MR 732415

[18] P.-L. Lions, Courses at the Collège de France, www.college-de-france.fr.

[19] P.-L. Lions and P. E. Souganidis, Extended mean-field games. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **31** (2020), no. 3, 611–625 Zbl 1460.35020 MR 4170639

[20] S. Munoz, Classical solutions to local first order extended mean field games. 2021, arXiv:2102.13093.

[21] A. Porretta, Weak solutions to Fokker-Planck equations and mean field games. *Arch. Ration. Mech. Anal.* **216** (2015), no. 1, 1–62 Zbl 1312.35168 MR 3305653

Sebastian Muñoz
Department of Mathematics, University of Chicago, 5801 South Ellis Avenue, Chicago, IL 60637, USA; sbstn@math.uchicago.edu