Chin. Ann. Math.
34B(2), 2013, 161–192
DOI: 10.1007/s11401-013-0767-y

Chinese Annals of Mathematics, Series B

Control and Nash Games with Mean Field Effect*

Alain BENSOUSSAN[1] Jens FREHSE[2]

(*In honor of the scientific heritage of Jacques-Louis Lions*)

Abstract Mean field theory has raised a lot of interest in the recent years (see in particular the results of Lasry-Lions in 2006 and 2007, of Gueant-Lasry-Lions in 2011, of Huang-Caines-Malham in 2007 and many others). There are a lot of applications. In general, the applications concern approximating an infinite number of players with common behavior by a representative agent. This agent has to solve a control problem perturbed by a field equation, representing in some way the behavior of the average infinite number of agents. This approach does not lead easily to the problems of Nash equilibrium for a finite number of players, perturbed by field equations, unless one considers averaging within different groups, which has not been done in the literature, and seems quite challenging. In this paper, the authors approach similar problems with a different motivation which makes sense for control and also for differential games. Thus the systems of nonlinear partial differential equations with mean field terms, which have not been addressed in the literature so far, are considered here.

Keywords Mean field, Dynamic programming, Nash games, Equilibrium, Calculus of variations
2000 MR Subject Classification 49L20

1 Introduction

In this paper, we study the systems of nonlinear partial differential equations (or PDE for short) with mean field coupling. This extends the usual theory of a single PDE with mean field coupling. This extension has not been considered in the literature, probably because the motivation of mean field theory is precisely to eliminate the game aspect, by an averaging consideration. In fact, the starting point is a Nash equilibrium for an infinite number of players, with similar behavior. The averaging concept reduces this infinite number to a representative agent, who has a control problem to solve, with an external effect, representing the averaged impact of the infinite number of players. Of course, this framework relies on the assumption that the players behave in a similar way. Nevertheless, it eliminates the situation of a remaining Nash equilibrium for a finite number of players, with mean field terms. One may imagine groups with

Manuscript received August 12, 2012.
[1]International Center for Decision and Risk Analysis, School of Management, University of Texas-Dallas, France; School of Business, the Hong Kong Polytechnic University, Hong Kong, China; Graduate Department of Financial Engineering, Ajou University, Korea. E-mail: alain.bensoussan@utdallas.edu
[2]Institute for Applied Mathematics, University of Bonn, Germany.
*Project supported by the WCU World Class University program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (No. R31-20007) and the Research Grants Council of HKSAR (No. PolyU 5001/11P).

non-homogeneous behavior, in which case it is likely that one may recover systems of nonlinear PDE with mean field coupling. Although interesting, this extension has not been considered in the literature, and seems quite challenging. This is why we develop here a different motivation, which has interest in itself. It makes sense for control problems as well as for differential games. The mean field coupling term in our case has a different interpretation. Another interesting feature of our approach is that we do not need to consider an ergodic situation, as it is the case in the standard approach of mean field theory. In fact, considering strictly positive discounts is quite meaningful in our applications. This leads to systems of nonlinear PDE with mean field coupling terms, that we can study with a minimum set of assumptions. This is the objective of this paper. The ergodic case, when the discount vanishes, requires much stringent assumptions, as is already the case when there is no mean field terms. This case will be dealt with in a following article. We refer to [2, 5–7] for the situation without mean field term. Basically, our set of assumptions remains valid, and we have to incorporate additional assumptions to deal with the mean field terms.

2 Control Framework

2.1 Bellman equation

We consider a classical control problem here. We treat an infinite horizon problem, with stationary evolution of the state. In order to remain within a bounded domain, we assume that the state evolution, modelled as a diffusion, is reflected on the boundary of the domain. More precisely, we define a probability space Ω, with $\mathcal{A}$, P equipped with a filtration $\mathcal{F}^t$ and a standard n-dimensional $\mathcal{F}^t$ Wiener process $w(t)$. Let $\mathcal{O}$ be a smooth bounded domain of $\mathbb{R}^n$. We set $\Gamma = \partial\mathcal{O}$. We denote by $\nu = \nu(x)$ the outward unit normal on a point x of Γ. Let $g(x, v)$ be a continuously differentiable function from $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. The second argument represents the control. To simplify, we omit to consider constraints on the control. Let $v(t)$ be a stochastic process adapted to the filtration $\mathcal{F}^t$. A controlled diffusion reflected at the boundary Γ with initial state $x \in \mathcal{O}$ is a pair of processes $y(t)$, $\xi(t)$, such that $y(t)$ is continuous adapted, $y(t) \in \overline{\mathcal{O}}$, and $\xi(t)$ is continuous adapted scalar increasing

$$\begin{aligned} &\mathrm{d}y(t) = g(y(t), v(t))\mathrm{d}t + \sqrt{2}\mathrm{d}w(t) - \nu(y(t))\mathbb{1}_{y(t)\in\Gamma}\mathrm{d}\xi(t), \\ &y(0) = x. \end{aligned} \tag{2.1}$$

Next, let $f(x, v)$ be a scalar function on $\mathbb{R}^n \times \mathbb{R}^m$, which is continuous and continuously differentiable in v. We assume also that $f(x, v)$ is bounded below. We define the payoff

$$J_\alpha(x, v(\,\cdot\,)) = E\int_0^{+\infty} \exp(-\alpha t\, f(y(t), v(t)))\mathrm{d}t. \tag{2.2}$$

We define the value function

$$u_\alpha(x) = \inf_{v(\,\cdot\,)} J_\alpha(x, v(\,\cdot\,)). \tag{2.3}$$

It is a fundamental result of dynamic programming that the value function is the solution to a partial differential equation, the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} & -\triangle u_\alpha(x) + \alpha u_\alpha(x) = H(x, Du_\alpha(x)), \quad x \in \mathcal{O}, \\ & \frac{\partial u_\alpha}{\partial \nu}\Big|_\Gamma = 0 \end{aligned} \tag{2.4}$$

with the following notations:

$$\begin{aligned} & H(x,q) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \\ & H(x,q) = \inf_v L(x,v,q), \\ & L(x,v,q) = f(x,v) + q \cdot g(x,v). \end{aligned} \tag{2.5}$$

An essential question becomes solving the PDE (2.4), and finding a sufficiently smooth solution. In the interpretation, which we shall discuss, we assume the regularity allowing to perform the calculations that we describe (in particular, taking derivatives).

The function H is called the Hamiltonian, and the function L is called the Lagrangian. Since the function L is continuously differentiable in v, and the infimum is attained at points, such that

$$\frac{\partial L}{\partial v}(x,v,q) = 0. \tag{2.6}$$

We shall assume that we can find a measurable map $\widehat{v}(x,q)$, which satisfies (2.6) and achieves the infimum in (2.5). We then have

$$H(x,q) = f(x, \widehat{v}(x,q)) + q \cdot g(x, \widehat{v}(x,q)). \tag{2.7}$$

It is also convenient to write

$$G(x,q) = g(x, \widehat{v}(x,q)). \tag{2.8}$$

With this notation, we can write Bellman equation as follows:

$$\begin{aligned} & -\triangle u_\alpha(x) - G(x, Du_\alpha(x)) + \alpha u_\alpha(x) = f(x, \widehat{v}(x, Du_\alpha(x))), \\ & \frac{\partial u_\alpha}{\partial \nu}\Big|_\Gamma = 0. \end{aligned} \tag{2.9}$$

The main result of dynamic programming is that the infimum in (2.3) is attained for the control

$$\widehat{v}(t) = \widehat{v}(\widehat{y}(t), Du_\alpha(\widehat{y}(t))), \tag{2.10}$$

where the process $\widehat{y}(t)$, i.e., the optimal trajectory, together with an increasing process $\widehat{\xi}(t)$, is the solution to

$$\begin{aligned} & \mathrm{d}\widehat{y}(t) = G(\widehat{y}(t), Du_\alpha(\widehat{y}(t)))\mathrm{d}t + \sqrt{2}\mathrm{d}w(t) - \nu(\widehat{y}(t))\mathbb{1}_{\widehat{y}(t)\in\Gamma}\mathrm{d}\widehat{\xi}(t), \\ & \widehat{y}(0) = x, \quad \widehat{y}(t) \in \overline{\mathcal{O}}. \end{aligned} \tag{2.11}$$

The main feature is that the optimal control is obtained through a feedback $\widehat{v}(x, Du_\alpha(x))$. We note $\widehat{v}_\alpha(x) = \widehat{v}(x, Du_\alpha(x))$.

2.2 Revisiting Bellman equation

The fundamental result of Dynamic Programming motivates the following approach. Suppose that we restrict ourselves to the controls defined through feedbacks. A feedback is simply a measurable map $v(x)$. In fact, x can be restricted to $\overline{\mathcal{O}}$. To each feedback, we associate the function $u_{v(\,\cdot\,),\alpha}(x)$ as a solution to

$$\begin{aligned}
&-\triangle u_{v(\,\cdot\,),\alpha}(x) - g(x,v(x))\cdot Du_{v(\,\cdot\,),\alpha}(x) + \alpha u_{v(\,\cdot\,),\alpha}(x) = f(x,v(x)),\\
&\frac{\partial u_{v(\,\cdot\,),\alpha}}{\partial \nu}\Big|_{\Gamma} = 0.
\end{aligned} \tag{2.12}$$

In fact, a feedback defines a particular case of control. We define the trajectory related to the feedback $v(\,\cdot\,)$ by considering the reflected diffusion

$$\begin{aligned}
&\mathrm{d}y(t) = g(y(t),v(y(t)))\mathrm{d}t + \sqrt{2}\mathrm{d}w(t) - \nu(y(t))1\!\!1_{y(t)\in\Gamma}\mathrm{d}\xi(t),\\
&y(0) = x.
\end{aligned} \tag{2.13}$$

To save notation, we omit to write that the trajectory depends on the feedback. The control corresponding to $v(\,\cdot\,)$ is $v(y(t))$. The corresponding payoff (see (2.2)) is thus

$$E\int_0^{+\infty} \exp(-\alpha t\, f(y(t),v(y(t))))\mathrm{d}t. \tag{2.14}$$

We shall also write it as $J_\alpha(x,v(\,\cdot\,))$ to avoid redundant notation. However, here $v(\,\cdot\,)$ refers to the feedback. It is easy to check that

$$u_{v(\,\cdot\,),\alpha}(x) = J_\alpha(x,v(\,\cdot\,)). \tag{2.15}$$

If we take $v(\,\cdot\,) = \widehat{v}_\alpha(\,\cdot\,)$, then $u_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x) = u_\alpha(x)$, $\forall x$, where $u_\alpha(x)$ is the solution to Bellman equations (2.4) and (2.9). From maximum principle considerations, we can assert that

$$u_\alpha(x) \le u_{v(\,\cdot\,),\alpha}(x), \quad \forall v(\,\cdot\,),\ \forall x\in\overline{\mathcal{O}}. \tag{2.16}$$

We recover that $\widehat{v}_\alpha(\,\cdot\,)$ is an optimal feedback.

2.3 Calculus of variations approach

To avoid confusion of notation, we shall consider the process defined by (2.13), with an initial condition x_0. The corresponding process $y(t)$ is a Markov process, whose probability distribution has a density denoted by $p_{v(\,\cdot\,)}(x,t)$ to emphasize the dependence on the feedback $v(\,\cdot\,)$, which is the solution to the Chapman-Kolmogorov equation

$$\begin{aligned}
&\frac{\partial p}{\partial t} - \triangle p + \mathrm{div}(g(x,v(x))p) = 0, \quad x\in\mathcal{O},\\
&\frac{\partial p}{\partial \nu} - g(x,v(x))\cdot\nu(x)p = 0, \quad x\in\Gamma,\\
&p(x,0) = \delta_{x_0}(x).
\end{aligned} \tag{2.17}$$

By the smoothing effect of diffusions, $p_{v(\,\cdot\,)}(x,t)$ is a function and not a distribution for any positive t. Moreover, $p_{v(\,\cdot\,)}(x,t)$ is, for any t, a density probability on $\mathcal{O}$. Now we can express

$$\begin{aligned}
u_{v(\,\cdot\,),\alpha}(x_0) &= E\int_0^{+\infty} \exp(-\alpha t\, f(y(t),v(y(t))))\mathrm{d}t\\
&= \int_0^{+\infty} \exp\Big(-\alpha t\Big(\int_{\mathcal{O}} p_{v(\,\cdot\,)}(x,t) f(x,v(x))\mathrm{d}x\Big)\Big)\mathrm{d}t.
\end{aligned} \tag{2.18}$$

Let us define

$$p_{v(\,\cdot\,),\alpha}(x) = \alpha \int_0^{+\infty} \exp(-\alpha t\, p_{v(\,\cdot\,)}(x,t))\mathrm{d}t, \tag{2.19}$$

which is the solution to

$$\begin{aligned} & -\triangle p_\alpha + \mathrm{div}(g(x,v(x))p_\alpha) + \alpha p_\alpha = \alpha \delta_{x_0}, \quad x \in \mathcal{O}, \\ & \frac{\partial p_\alpha}{\partial \nu} - g(x,v(x)) \cdot \nu(x) p_\alpha = 0, \quad x \in \Gamma. \end{aligned} \tag{2.20}$$

We then get the formula

$$\alpha u_{v(\,\cdot\,),\alpha}(x_0) = \int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) f(x,v(x))\mathrm{d}x. \tag{2.21}$$

We can then state the lemma as follows.

Lemma 2.1 *The functional $u_{v(\,\cdot\,),\alpha}(x_0)$ is Frechet differentiable in $v(\,\cdot\,)$ with the formula*

$$\alpha \frac{\mathrm{d}}{\mathrm{d}\theta} u_{v(\,\cdot\,)+\theta\widetilde{v}(\,\cdot\,),\alpha}(x_0)\Big|_{\theta=0} = \int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) \frac{\partial L}{\partial v}(x, v(x), Du_{v(\,\cdot\,),\alpha}(x))\widetilde{v}(x)\mathrm{d}x. \tag{2.22}$$

Proof We first show that $p_{v(\,\cdot\,),\alpha}(x)$ is Frechet-differentiable in $v(\,\cdot\,)$ for fixed x. Indeed, by direct differentiation, we check that

$$\widetilde{p}_\alpha(x) = \frac{\mathrm{d}}{\mathrm{d}\theta} p_{v(\,\cdot\,)+\theta\widetilde{v}(\,\cdot\,),\alpha}(x)\Big|_{\theta=0}$$

is the solution to

$$\begin{aligned} & -\triangle \widetilde{p}_\alpha + \mathrm{div}(g(x,v(x))\widetilde{p}_\alpha) + \alpha\widetilde{p}_\alpha + \mathrm{div}(g_v(x,v(x))\widetilde{v}(x) p_{v(\,\cdot\,),\alpha}(x)) = 0, \quad x \in \mathcal{O}, \\ & \frac{\partial \widetilde{p}_\alpha}{\partial \nu} - g(x,v(x)) \cdot \nu(x)\widetilde{p}_\alpha - g_v(x,v(x))\widetilde{v}(x) \cdot \nu(x) p_{v(\,\cdot\,),\alpha} = 0, \quad x \in \Gamma, \end{aligned} \tag{2.23}$$

in which

$$g_v(x,v) = \frac{\partial g}{\partial v}(x,v).$$

Therefore,

$$\alpha \frac{\mathrm{d}}{\mathrm{d}\theta} u_{v(\,\cdot\,)+\theta\widetilde{v}(\,\cdot\,),\alpha}(x_0)\Big|_{\theta=0} = \int_{\mathcal{O}} \widetilde{p}_\alpha(x) f(x,v(x))\mathrm{d}x + \int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) \frac{\partial f}{\partial v}(x,v(x))\widetilde{v}(x)\mathrm{d}x.$$

But

$$\int_{\mathcal{O}} \widetilde{p}_\alpha(x) f(x,v(x))\mathrm{d}x = \int_{\mathcal{O}} Du_{v(\,\cdot\,),\alpha}(x) \cdot g_v(x,v(x))\widetilde{v}(x)\mathrm{d}x,$$

and the result follows immediately.

Corollary 2.1 *A feedback $\widehat{v}_\alpha(\,\cdot\,)$, which minimizes $u_{v(\,\cdot\,),\alpha}(x_0)$, satisfies*

$$\widehat{v}_\alpha(x) = \widehat{v}(x, Du_\alpha(x)),$$

where $u_\alpha(x)$ is the solution to the Bellman equation (2.4).

Proof The Frechet derivative of $u_{v(\,\cdot\,),\alpha}(x_0)$ at $\widehat{v}_\alpha(\,\cdot\,)$ must vanish. From formula (2.22), we deduce

$$\frac{\partial L}{\partial v}(x, \widehat{v}_\alpha(x), Du_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x)) = 0.$$

But then $u_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x) = u_\alpha(x)$ and the result follows.

Remark 2.1 We note that the feedback $\widehat{v}_\alpha(x)$ is optimal for any value of x_0. In this approach, Bellman equation appears in expressing a necessary condition of optimality for a calculus of variations problem. This is not at all the traditional way, in which Bellman equation is introduced as a sufficient condition of optimality for the original stochastic control problem (2.3). This calculus of variations approach is rather superfluous for the standard stochastic control problem, since it leads to weaker results. In particular, we need to restrict the class of controls to the feedback controls, whereas we know that the optimality of the feedback controls holds against any non-anticipative controls. However, the calculus of variations approach can be extended to more general classes of control problems, as considered in this work, whereas the traditional approach can not.

3 More General Control Problems

3.1 Motivation

We consider the same objective function as before, but we would also like to control a functional of the path. As an example, we want to minimize the modified functional

$$\begin{aligned} J_\alpha(x_0, v(\,\cdot\,)) = {} & E\int_0^{+\infty} \exp(-\alpha t\, f(y(t), v(y(t))))\mathrm{d}t \\ & + \frac{\gamma}{2}\Big(E\int_0^{+\infty} \exp(-\alpha t\, h(y(t)))\mathrm{d}t - M\Big)^2, \end{aligned} \tag{3.1}$$

where $h(x)$ is continuous. We can regard the second term as transforming a constraint into a penalty term in the cost functional.

We restrict ourselves to the feedbacks $v(\,\cdot\,)$ and $y(0) = x_0$. Clearly, the dynamic programming approach fails for this problem, since $J_\alpha(x, v(\,\cdot\,))$ is not a solution to a PDE. However, we can extend the calculus of variations approach. Indeed, considering the probability $p_{v(\,\cdot\,),\alpha}(x)$ as a solution to (2.20), we can write $J_\alpha(x_0, v(\,\cdot\,))$ as

$$\begin{aligned} J_\alpha(x_0, v(\,\cdot\,)) &= \frac{1}{\alpha}\int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) f(x, v(x))\mathrm{d}x + \frac{\gamma}{2}\Big(\frac{1}{\alpha}\int p_{v(\,\cdot\,),\alpha}(x) h(x)\mathrm{d}x - M\Big)^2 \\ &= \frac{1}{\alpha}\int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) f(x, v(x))\mathrm{d}x + \frac{1}{\alpha}\Phi_\alpha(p_{v(\,\cdot\,),\alpha}), \end{aligned} \tag{3.2}$$

where

$$\Phi_\alpha(m) = \frac{\gamma}{2\alpha}\Big(\int_{\mathcal{O}} m(x)h(x)\mathrm{d}x - \alpha M\Big)^2 \tag{3.3}$$

is a functional on the set $L^1(\mathcal{O})$.

3.2 Calculus of variations problem

To avoid Dirac measures on the right-hand side, we shall consider the state equation $p_{v(\,\cdot\,),\alpha}(\,\cdot\,)$ as a solution to

$$\begin{aligned} & -\triangle p_\alpha + \mathrm{div}(g(x, v(x))p_\alpha) + \alpha p_\alpha = \alpha m_0, \quad x \in \mathcal{O}, \\ & \frac{\partial p_\alpha}{\partial \nu} - g(x, v(x)) \cdot \nu(x) p_\alpha = 0, \quad x \in \Gamma, \end{aligned} \tag{3.4}$$

in which m_0 is a probability density on $\mathcal{O}$. It corresponds clearly to equations (2.17)–(2.19) with initial condition m_0 instead of δ_{x_0}. It means that, going back to the reflected diffusion (2.13), we can not observe the initial state. However, since we apply a feedback on the state, we still consider that we can observe the state at any time strictly positive. We choose the feedback $v(\,\cdot\,)$ in order to minimize the payoff

$$\alpha J_\alpha(v(\,\cdot\,)) = \int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x) f(x, v(x))\mathrm{d}x + \Phi_\alpha(p_{v(\,\cdot\,),\alpha}). \tag{3.5}$$

The functional $\Phi_\alpha(m)$ is defined on $L^1(\mathcal{O})$, and we assume that it is Frechet-differentiable, with derivative in $L^\infty(\mathcal{O})$. Namely

$$\frac{\mathrm{d}\Phi_\alpha(m+\theta\widetilde{m})}{\mathrm{d}\theta}\Big|_{\theta=0} = \int_{\mathcal{O}} V_{m,\alpha}(x)\widetilde{m}(x)\mathrm{d}x, \tag{3.6}$$

where $V_{m,\alpha}(\,\cdot\,)$ is in $L^\infty(\mathcal{O})$. In the example (3.3), we simply have

$$V_{m,\alpha}(x) = \frac{\gamma}{\alpha} h(x)\Big(\int_{\mathcal{O}} m(\xi)h(\xi)\mathrm{d}\xi - \alpha M\Big). \tag{3.7}$$

Our problem is to minimize the functional $\alpha J_\alpha(v(\,\cdot\,))$. In fact, since there are no constraints on the feedback control, we will write a necessary condition of optimality for an optimal feedback.

3.3 Euler condition of optimality

We just check that the functional $\alpha J_\alpha(v(\,\cdot\,))$ has a Frechet derivative. We associate to a feedback $v(\,\cdot\,)$, i.e., the PDE with mean field term

$$\begin{aligned} &-\triangle u_{v(\,\cdot\,),\alpha}(x) - g(x, v(x))\cdot Du_{v(\,\cdot\,),\alpha}(x) + \alpha u_{v(\,\cdot\,),\alpha}(x) = f(x, v(x)) + V_{p_{v(\,\cdot\,),\alpha},\alpha}(x),\\ &\frac{\partial u_{v(\,\cdot\,),\alpha}}{\partial\nu}\Big|_\Gamma = 0. \end{aligned} \tag{3.8}$$

We see that, conversely to the case (2.12), the PDE depends explicitly on $p_{v(\,\cdot\,),\alpha}$.

Lemma 3.1 *The functional $\alpha J_\alpha(v(\,\cdot\,))$ has a Frechet differential given by*

$$\alpha\frac{\mathrm{d}}{\mathrm{d}\theta}J_\alpha(v(\,\cdot\,)+\theta\widetilde{v}(\,\cdot\,))\Big|_{\theta=0} = \int_{\mathcal{O}} p_{v(\,\cdot\,),\alpha}(x)\frac{\partial L}{\partial v}(x, v(x), Du_{v(\,\cdot\,),\alpha}(x))\widetilde{v}(x)\mathrm{d}x. \tag{3.9}$$

Proof The proof is similar to that of Lemma 2.1. The Lagrangian $L(x, v, q)$ is defined in (2.5).

Then we can give a necessary condition of optimality for a feedback $\widehat{v}_\alpha(\,\cdot\,)$. We recall the notations (2.6)–(2.8). We consider the system

$$\begin{aligned} &-\triangle u_\alpha + \alpha u_\alpha = H(x, Du_\alpha) + V_{m_\alpha,\alpha}(x), \quad x\in\mathcal{O},\\ &\frac{\partial u_\alpha}{\partial\nu}\Big|_\Gamma = 0,\\ &-\triangle m_\alpha + \mathrm{div}(G(x, Du_\alpha)m_\alpha) + \alpha m_\alpha = \alpha m_0, \quad x\in\mathcal{O},\\ &\frac{\partial m_\alpha}{\partial\nu} - G(x, Du_\alpha)\cdot\nu(x)m_\alpha = 0, \quad x\in\Gamma. \end{aligned} \tag{3.10}$$

We then write

$$\widehat{v}_\alpha(x) = \widehat{v}(x, Du_\alpha(x)), \tag{3.11}$$

where we recall the definition of $\widehat{v}(x,q)$ as the solution to (2.6). $H(x,q)$ and $G(x,q)$ have been defined in (2.7) and (2.8), respectively.

We can state as follows.

Proposition 3.1 *For a feedback $\widehat{v}_\alpha(\,\cdot\,)$ to be optimal for the functional* (3.5)*, it is necessary that equations* (3.10) *and* (4.14) *hold.*

Proof From the expression (3.9) of the Frechet derivative, one must have

$$\frac{\partial L}{\partial v}(x, \widehat{v}_\alpha(x), Du_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x)) = 0.$$

Hence

$$\widehat{v}_\alpha(x) = \widehat{v}(x, Du_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x)).$$

If we set

$$u_\alpha(x) = u_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x), \quad m_\alpha(x) = p_{\widehat{v}_\alpha(\,\cdot\,),\alpha}(x),$$

and from the equations (3.4) and (3.8), it is clear that $(u_\alpha(\,\cdot\,),\ m_\alpha(\,\cdot\,))$ is a solution to the system (3.10). This completes the proof.

Remark 3.1 As mentioned in the case of standard dynamic programming, showing that the system (3.10) has a solution becomes a problem itself. The claim that it has a solution, as a consequence of necessary conditions of optimality, lies on the assumption that an optimal feedback for the control problem (3.2) exists, and is not fully rigorous. We will address this problem in the analytic part.

4 Nash Equilibrium

4.1 Definition of the problem

To avoid redundant notation, we will not write explicitly the index α. We will generalize the calculus of variations problem described in Subsection 3.2, and then provide applications and exemples. We consider N players, which decide on feedbacks $v^i(x)$ $(i = 1, \cdots, N,\ x \in \mathbb{R}^n)$. We shall use the notation

$$v = (v^1, \cdots, v^N) = (v^i, \overline{v}^i).$$

The second notation means that we emphasize the case of player i, so we indicate his decision v^i, and denote by $\overline{v}^i$ the vector of decisions of all other players. The decision v^i belongs to an Euclidean space $\mathbb{R}^{d_i}$. We next consider continuous functions $f^i(x,v) \in \mathbb{R}$ and $g^i(x,v) \in \mathbb{R}^n$.

An important difference from the case of a single player is that the decision of player i is not just the feedback $v^i(x)$, a measurable function from $\mathbb{R}^n$ to $\mathbb{R}^{d_i}$, but also the state $p^i(x)$, a probability density on $\mathcal{O}$ which is a continuous function. So player i chooses the pair $v^i(\,\cdot\,),\ p^i(\,\cdot\,)$. We will require some constraints between these two decisions, but it is important to proceed in this way, for the reasons which will be explained below. In a way similar to v, we shall use the notation

$$p = (p^1, \cdots, p^N) = (p^i, \overline{p}^i)$$

to refer to the vector of states.

Each player wants to minimize his payoff

$$J^i(v(\cdot);\, p(\cdot)) = \int_{\mathcal{O}} p^i(x) f^i(x, v(x))\mathrm{d}x + \Phi^i(p). \tag{4.1}$$

The functionals $\Phi^i(p)$ are defined on $(L^1(\mathcal{O}))^N$. The functionals have partial Frechet derivatives. More precisely, by our convention $\Phi^i(m) = \Phi^i(m^i, \overline{m}^i)$, we assume that

$$\frac{\mathrm{d}\Phi^i(m^i + \theta\widetilde{m}^i, \overline{m}^i)}{\mathrm{d}\theta}\Big|_{\theta=0} = \int_{\mathcal{O}} V^i_{[m]}(x)\widetilde{m}^i(x)\mathrm{d}x, \tag{4.2}$$

and the functions $V^i_{[m]}(x)$ are in $L^\infty(\mathcal{O})$.

Our concept of Nash equilibrium is as follows. A pair $(\widehat{v}(\cdot),\ \widehat{p}(\cdot))$ is a Nash equilibrium, if the following conditions are satisfied. Let $v^i(\cdot)$ be any feedback for player i. Define $p^i_{v^i(\cdot),\widehat{\overline{v}}(\cdot)^i}(x)$ as the solution to

$$\begin{aligned} &-\triangle p^i + \mathrm{div}(g^i(x, v^i(x), \widehat{\overline{v}}(x)^i)p^i) + \alpha p^i = \alpha m^i_0, \quad x \in \mathcal{O}, \\ &\frac{\partial p^i}{\partial \nu} - g^i(x, v^i(x), \widehat{\overline{v}}(x)^i)\cdot \nu(x)p^i = 0, \quad x \in \Gamma. \end{aligned} \tag{4.3}$$

We note that the feedbacks of all players except i are frozen at the values $\widehat{v}^j(\cdot)$ $(j \neq i)$. The player i can choose his own feedback $v^i(\cdot)$. His decision $p^i(\cdot)$ is not decided independent of $v^i(\cdot)$ and of the vector of other players' decisions $\widehat{\overline{v}}(\cdot)^i$. It is $p^i_{v^i(\cdot),\widehat{\overline{v}}(\cdot)^i}(\cdot)$. However, he considers the decisions of the other players as $\widehat{\overline{v}}(\cdot)^i$, $\widehat{\overline{p}}(\cdot)^i$. The important thing to notice is that, he can not influence either $\widehat{\overline{v}}(\cdot)^i$ as expected, or $\widehat{\overline{p}}(\cdot)^i$. Therefore, the first condition is

$$\widehat{p}^i(x) = p^i_{\widehat{v}^i(\cdot),\widehat{\overline{v}}(\cdot)^i}(x) = p^i_{\widehat{v}(\cdot)}(x). \tag{4.4}$$

The second condition is that

$$J^i(\widehat{v}(\cdot);\, \widehat{p}(\cdot)) \le J^i(v^i(\cdot), \widehat{\overline{v}}(\cdot)^i;\, p^i_{v^i(\cdot),\widehat{\overline{v}}(\cdot)^i}(\cdot), \widehat{\overline{p}}(\cdot)^i). \tag{4.5}$$

This condition explains why the $p^i(\cdot)$ is also considered as a decision variable. If only the feedbacks $v^i(\cdot)$ (and not the pair $(v^i(\cdot), p^i(\cdot))$) were decision variables, we would have in (4.5), the vector of functions $p^j_{v^i(\cdot),\widehat{\overline{v}}(\cdot)^i}(x)$, $j \neq i$, instead of $\widehat{\overline{p}}(\cdot)^i$. We do not know how to solve this problem. The difficulty arises from the fact that the functional $\Phi^i(m)$ depends on all the functions. If it were dependent on m^i only, it would not be necessary to make the difference. This occurs, in particular, in the case of the control problem, when there is only one player.

4.2 Necessary conditions for a Nash equilibrium

Let $v(\cdot) = (v^i(\cdot), \overline{v}^i(\cdot))$ and $p(\cdot) = (p^i(\cdot), \overline{p}^i(\cdot))$ be a pair of vector feedbacks and probabilities. We associate probabilities $p^i_{v^i(\cdot),\overline{v}(\cdot)^i}(\cdot)$ as a solution to

$$\begin{aligned} &-\triangle p^i + \mathrm{div}(g^i(x, v^i(x), \overline{v}(x)^i)p^i) + \alpha p^i = \alpha m^i_0, \quad x \in \mathcal{O}, \\ &\frac{\partial p^i}{\partial \nu} - g^i(x, v^i(x), \overline{v}(x)^i)\cdot \nu(x)p^i = 0, \quad x \in \Gamma. \end{aligned} \tag{4.6}$$

We furthermore define functions $u^i_{v^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i;\overline{p}^i(\,\cdot\,)}(x)$ by

$$\begin{aligned}
&-\triangle u^i - g^i(x, v^i(x), \overline{v}(x)^i)\cdot Du^i + \alpha u^i \\
&= f^i(x, v^i(x), \overline{v}(x)^i) + V^i_{[p^i_{v^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i}(\,\cdot\,),\overline{p}^i(\,\cdot\,)]}(x), \quad x \in \mathcal{O}, \\
&\frac{\partial u^i}{\partial \nu}\Big|_\Gamma = 0.
\end{aligned} \tag{4.7}$$

We then claim the following result.

Lemma 4.1 *The functional $J^i(v(\,\cdot\,); p(\,\cdot\,))$ satisfies*

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\theta} J^i(v^i(\,\cdot\,) + \theta\widetilde{v}^i(\,\cdot\,), \overline{v}(\,\cdot\,)^i; p^i_{v^i(\,\cdot\,)+\theta\widetilde{v}^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i}(\,\cdot\,), \overline{p}^i(\,\cdot\,))\Big|_{\theta=0} \\
&= \int_{\mathcal{O}} p^i_{v^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i}(x) \frac{\partial L^i}{\partial v^i}(x, v^i(x), \overline{v}(x)^i, Du^i_{v^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i;\overline{p}^i(\,\cdot\,)}(x))\mathrm{d}x
\end{aligned} \tag{4.8}$$

with

$$L^i(x, v, q^i) = f^i(x, v) + q^i \cdot g^i(x, v). \tag{4.9}$$

Proof The proof is similar to the case of a single player, since the vectors $\overline{v}(\,\cdot\,)^i$, $\overline{p}^i(\,\cdot\,)$ are fixed in the functional $J^i(v^i(\,\cdot\,), \overline{v}(\,\cdot\,)^i; p^i_{v^i(\,\cdot\,),\overline{v}(\,\cdot\,)^i}(\,\cdot\,), \overline{p}^i(\,\cdot\,))$.

We will now state necessary conditions for a pair $\widehat{v}(\,\cdot\,), \widehat{p}(\,\cdot\,)$ to be a Nash equilibrium. We first define a Nash equilibrium of the Lagrangian functions. Namely, we solve the system

$$\frac{\partial L^i}{\partial v^i}(x, v^i, \overline{v}^i, q^i) = 0, \quad i = 1, \cdots, N. \tag{4.10}$$

This defines functions $\widehat{v}^i(x, q)$ where $q = (q^1, \cdots, q^N)$. We define next the Hamiltonians

$$H^i(x, q) = L^i(x, \widehat{v}(x, q), q^i) \tag{4.11}$$

and

$$G^i(x, q) = g^i(x, \widehat{v}(x, q)). \tag{4.12}$$

We next introduce the system

$$\begin{aligned}
&-\triangle u^i + \alpha u^i = H^i(x, Du) + V^i_{[m]}(x), \quad x \in \mathcal{O}, \\
&\frac{\partial u^i}{\partial \nu}\Big|_\Gamma = 0, \\
&-\triangle m^i + \mathrm{div}(G^i(x, Du)m^i) + \alpha m^i = \alpha m^i_0, \quad x \in \mathcal{O}, \\
&\frac{\partial m^i}{\partial \nu} - G^i(x, Du)\cdot\nu(x) m^i = 0, \quad x \in \Gamma,
\end{aligned} \tag{4.13}$$

and define

$$\widehat{v}^i(x) = \widehat{v}^i(x, Du(x)), \quad \widehat{p}^i(x) = m^i(x). \tag{4.14}$$

By construction, we have

$$m^i(x) = p^i_{\widehat{v}^i(\,\cdot\,),\overline{\widehat{v}}(\,\cdot\,)^i}(x), \tag{4.15}$$

$$u^i(x) = u^i_{\widehat{v}^i(\,\cdot\,),\overline{\widehat{v}}(\,\cdot\,)^i;\overline{m}^i(\,\cdot\,)}(x). \tag{4.16}$$

We can then state as follows.

Proposition 4.1 *A Nash equilibrium* $(\widehat{v}(\,\cdot\,), \widehat{p}(\,\cdot\,))$ *of functionals* (4.1) *in the sense of conditions* (4.4)–(4.5) *must satisfy the relations* (4.14).

Proof In view of (4.5) and the formula giving the Frechet differential (4.8), we must have

$$\frac{\partial L^i}{\partial v^i}(x, \widehat{v}^i(x), \overline{\widehat{v}}(x)^i, Du^i_{\widehat{v}^i(\,\cdot\,),\overline{\widehat{v}}(\,\cdot\,)^i;\overline{\widehat{p}}^i(\,\cdot\,)}(x)) = 0.$$

In view of (4.4), the functions $m^i(x)$ and $u^i(x)$ defined by (4.15) and (4.16) are solutions to (4.13), and conditions (4.14) are satisfied. This completes the proof.

4.3 Examples

We give here an example of the functional $\Phi^i(m)$. We set

$$\Phi^i(m) = \frac{\gamma}{2}\Big(\int_{\mathcal{O}} m^i(x)h^i(x)\mathrm{d}x - \frac{1}{N}\sum_{j=1}^N \int_{\mathcal{O}} m^j(x)h^j(x)\mathrm{d}x\Big)^2. \tag{4.17}$$

When this functional is incorporated into the payoff (4.1), player i aims at equalizing a quantity of interest with all corresponding ones of other players. This functional has a Frechet differential in m^i given by

$$V^i_{[m]}(x) = \gamma\Big(1 - \frac{1}{N}\Big)\Big(\int_{\mathcal{O}} m^i(\xi)h^i(\xi)\mathrm{d}\xi - \frac{1}{N}\sum_{j=1}^N \int_{\mathcal{O}} m^j(\xi)h^j(\xi)\mathrm{d}\xi\Big)h^i(x). \tag{4.18}$$

4.4 Probabilistic interpretation

We can give a probabilistic interpretation to the Nash game (4.1) in the sense of (4.4)–(4.5). We consider feedbacks $v^i(\,\cdot\,)$, and construct on a probability space $\Omega, \mathcal{A}, P$ trajectories $y^i(t) \in \overline{\mathcal{O}}$, which are independent and have probability densities $p^i(t)$ defined on $\mathcal{O}$. These densities as well as the feedbacks are decisions. Then we set

$$p^i = \alpha \int_0^{+\infty} \exp(-\alpha t\, p^i(t))\mathrm{d}t.$$

If we consider the functional (4.17), we have the interpretation

$$\Phi^i(p) = \frac{\gamma}{2}\Big(\alpha E\int_0^{+\infty} \exp(-\alpha t\, h^i(y^i(t)))\mathrm{d}t - \frac{\alpha}{N}\sum_{j=1}^N E\int_0^{+\infty} \exp(-\alpha t\, h^j(y^j(t)))\mathrm{d}t\Big)^2,$$

so the functional $J^i(v(\,\cdot\,);p(\,\cdot\,))$ defined by (4.1) has the following interpretation:

$$\begin{aligned}J^i(v(\,\cdot\,);p(\,\cdot\,)) = \alpha E\int_0^{+\infty} &\exp(-\alpha t\, f^i(y^i(t), v(y^i(t))))\mathrm{d}t\\ &+ \frac{\gamma}{2}\Big(\alpha E\int_0^{+\infty}\exp(-\alpha t\, h^i(y^i(t)))\mathrm{d}t\\ &- \frac{\alpha}{N}\sum_{j=1}^{N} E\int_0^{+\infty}\exp(-\alpha t\, h^j(y^j(t)))\mathrm{d}t\Big)^2,\end{aligned} \tag{4.19}$$

in which

$$v(y^i(t)) = (v^1(y^i(t)), \cdots, v^N(y^i(t))).$$

It is important to notice that, although the feedbacks relate to the different players, each player i considers that they operate on his trajectory $y^i(t)$. Moreover, condition (4.4) means that player i sees his trajectory $y^i(t)$ as the solution to

$$\begin{aligned}&\mathrm{d}y^i(t) = g^i(y^i(t), v(y^i(t)))\mathrm{d}t + \sqrt{2}\mathrm{d}w^i(t) - \nu(y^i(t))1\!\!1_{y(t)\in\Gamma}\mathrm{d}\xi^i(t),\\ &y^i(0) = y_0^i,\end{aligned} \tag{4.20}$$

where the Wiener processes $w^i(\,\cdot\,)$ are independent standard, and y_0^i are independent random variables, also independent of the Wiener processes, with probability density m_0^i. Since $y^i(t)$ is a reflected process, the pair $y^i(t), \xi^i(t)$ has to be defined jointly, in a unique way.

Remark 4.1 In problem (4.19)–(4.20), it is important to emphasize that player i considers the trajectories of other players $y^j(t)$ as given. His own trajectory $y^i(t)$ is defined by (4.20), in which he takes into account all feedbacks. However, he does not take into account his own influence on the trajectories of other players. Taking into account this influence would be a much more complex problem.

5 Analytic Framework

We shall develop here a theory to solve systems of the type (4.13), and define the set of assumptions. This will extend the results given in [2]. However, many techniques are similar to those developed in this reference. For the convenience of the reader, we shall indicate the main steps without all the details. Since we shall treat boundary conditions with local charts, it will be helpful to replace the Laplacian operator by a general second order operator in the divergence form. So we consider functions $a_{kl}(x)$ $(k, l = 1, \cdots, n)$ defined on $\mathbb{R}^n$, which satisfy

$$a_{kl}(\,\cdot\,)\text{ bounded},\quad \sum_{k,l=1}^{n} a_{kl}(x)\xi_k\xi_l \ge \underline{a}|\xi|^2,\quad \forall \xi\in\mathbb{R}^n. \tag{5.1}$$

We shall consider the matrix $a(x)$, whose elements are the quantities $a_{kl}(x)$, and write

$$\underline{a}I \le a(x) \le \overline{a}I, \tag{5.2}$$

where I is the identity matrix. Note that $a(x)$ is not necessarily symmetric.

5.1 Assumptions

We denote by $\mathcal{O}$ a smooth bounded open domain of $\mathbb{R}^n$. We write $\Gamma = \partial\mathcal{O}$. We define the second order linear operator

$$A\varphi(x) = -\mathrm{div}(a(x)\mathrm{grad}\varphi(x)), \quad x \in \mathcal{O},$$

and the boundary operator

$$\frac{\partial\varphi}{\partial\nu_A}(x) = \nu(x)\cdot a(x)\mathrm{grad}\varphi(x), \quad x \in \Gamma,$$

where $\nu(x)$ is the unit pointed outward normal vector on a point $x \in \Gamma$. The adjoint operator is defined by

$$A^*\varphi(x) = -\mathrm{div}(a^*(x)\mathrm{grad}\varphi(x)), \quad x \in \mathcal{O},$$

where $a^*(x)$ is the transpose of the matrix $a(x)$. The corresponding boundary operator is

$$\frac{\partial\varphi}{\partial\nu_{A^*}}(x) = \nu(x)\cdot a^*(x)\mathrm{grad}\varphi(x), \quad x \in \Gamma.$$

For $i = 1, \cdots, n$, we define the functions $H^i(x,q)$, $G^i(x,q)$, $q \in \mathbb{R}^{nN}$ with the following assumptions:

$$H^i(x,q) : \mathbb{R}^n \times \mathbb{R}^{nN} \to \mathbb{R}, \quad \text{measurable}, \tag{5.3}$$

$$|H^i(x,q)| \le K^i|q||q^i| + \sum_{j=1}^{i} K^i_j|q^j|^2 + k^i(x), \quad i = 1,\cdots,N-1, \tag{5.4}$$

where q^i $(i = 1, \cdots, N)$ are vectors of $\mathbb{R}^n$, representing the components of q. The functions $k^i(\,\cdot\,) \in L^p(\mathcal{O})$, $p > \frac{n}{2}$. We next assume

$$|H^N(x,q)| \le K^N|q|^2 + k^N(x), \quad k^N(\,\cdot\,) \in L^p(\mathcal{O}), \quad p > \frac{n}{2}. \tag{5.5}$$

We also assume that

$$|H^i(x,q)|_{q^i=0} \le C_0, \quad i = 1, \cdots, N. \tag{5.6}$$

Concerning $G^i(x,q)$, we assume

$$G^i(x,q) : \mathbb{R}^n \times \mathbb{R}^{nN} \to \mathbb{R}^n, \quad \text{measurable}, \tag{5.7}$$

$$|G^i(x,q)| \le K|q| + K. \tag{5.8}$$

We next consider the functionals $V^i_{[m]}(\,\cdot\,) :\ L^1(\mathcal{O};\mathbb{R}^N) \to L^1(\mathcal{O})$, such that

$$\|V^i_{[m]}\|_{L^\infty(\mathcal{O})} \le l(\|m\|), \tag{5.9}$$

where

$$\|m\| = \|m\|_{L^1(\mathcal{O};\mathbb{R}^N)} = \sup_{i=1}^{N} \int_{\mathcal{O}} |m^i(x)|\mathrm{d}x.$$

We also assume the convergence property

$$\text{if } m_j \to m \text{ pointwise, } \|m_j\|_{L^\infty(\mathcal{O};\mathbb{R}^N)} \le C, \quad \text{then } V^i_{[m_j]} \to V^i_{[m]} \text{ in } L^1(\mathcal{O}). \tag{5.10}$$

We finally consider

$$m^i_0 \in L^p(\mathcal{O}), \quad p > \frac{n}{2}, \quad m^i_0 \ge 0. \tag{5.11}$$

5.2 Preliminaries

We first state some technical results, the proof of which can be found in [2]. Without loss of generality, the assumptions (5.4)–(5.5) can be changed into

$$H^i(x,q) = Q^i(x,q)\cdot q^i + H_0^i(x,q), \quad i=1,\cdots,N \tag{5.12}$$

with

$$|Q^i(x,q)| \le K^i|q|, \tag{5.13}$$
$$Q^N(x,q) = Q^{N-1}(x,q), \tag{5.14}$$
$$|H_0^i(x,q)| \le \sum_{j=1}^{i} K_j^i |q^j|^2 + k^i(x), \quad i=1,\cdots,N, \tag{5.15}$$

in which all quantities have been defined in (5.4)–(5.5), except K_i^N $(i=1,\cdots,N-1)$ and K_N^N defined as follows:

$$K_i^N = K^N + \frac{K^{N-1}}{2}, \quad K_N^N = K^N + K^{N-1}. \tag{5.16}$$

So, from now on, we assume that (5.12)–(5.15) hold.

We shall also use the following technical property. Define the function

$$\beta(x) = \exp x - x - 1.$$

Let $s\in\mathbb{R}^N$. The components are defined as s^i $(i=1,\cdots,N)$. Let

$$X^N(s) = \exp[\beta(\gamma^N s^N) + \beta(-\gamma^N s^N)],$$

where γ^N is a positive constant. We then define recursively

$$X^i(s) = \exp[X^{i+1}(s) + \beta(\gamma^i s^i) + \beta(-\gamma^i s^i)], \quad i=1,\cdots,N-1,$$

where γ^i are positive constants. We have the lemma below.

Lemma 5.1 *One has*

$$\frac{\partial X^i}{\partial s^j} = \begin{cases} 0, & \text{if } j<i, \\ X^i\cdots X^j\gamma^j(\exp(\gamma^j s^j) - \exp(-\gamma^j s^j)), & \text{if } j\ge i. \end{cases}$$

Hence

$$|X^i(s) - X^i(0)| \le c(|s|)|s|^2, \tag{5.17}$$
$$X^i(s) \ge X^i(0) \ge 1, \tag{5.18}$$

where the constant c depends on the norm of the vector s and all constants $\gamma^1,\cdots,\gamma^N$. To avoid ambiguity later, we denote $X^i(0) = X_0^i$.

The proof is left to the reader.

5.3 Regularity result

We are interested in the system

$$\begin{aligned}
&Au^i + \alpha u^i = H^i(x, Du) + V^i_{[m]}(x), \quad x \in \mathcal{O},\\
&\frac{\partial u^i}{\partial \nu_A}\Big|_\Gamma = 0,\\
&Am^i + \operatorname{div}(G^i(x, Du)m^i) + \alpha m^i = \alpha m^i_0, \quad x \in \mathcal{O},\\
&\frac{\partial m^i}{\partial \nu_{A^*}} - G^i(x, Du)\cdot \nu(x) m^i = 0, \quad x \in \Gamma.
\end{aligned} \tag{5.19}$$

We interpret (5.19) in the weak sense

$$\begin{aligned}
&\int_\mathcal{O} a(x)Du^i(x)\cdot D\varphi^i(x)\mathrm{d}x + \alpha\int_\mathcal{O} u^i(x)\varphi^i(x)\mathrm{d}x\\
=&\int_\mathcal{O}(H^i(x, Du) + V^i_{[m]}(x))\varphi^i(x)\mathrm{d}x,
\end{aligned} \tag{5.20}$$

$$\begin{aligned}
&\int_\mathcal{O} a^*(x)Dm^i(x)\cdot D\psi^i(x)\mathrm{d}x - \int_\mathcal{O} m^i(x)G^i(x, Du)\cdot D\psi^i(x)\mathrm{d}x + \alpha\int_\mathcal{O} m^i(x)\psi^i(x)\mathrm{d}x\\
=&\,\alpha\int_\mathcal{O} m^i_0(x)\psi^i(x)\mathrm{d}x
\end{aligned} \tag{5.21}$$

for any pair $\varphi^i(\cdot) \in H^1 \cap L^\infty(\mathcal{O})$, $\psi^i \in W^{1,\infty}$, $i = 1, \cdots, N$.

We state the important regularity result concerning the u^i.

Theorem 5.1 *We assume that* (5.1) *and* (5.3)–(5.11) *hold. Suppose that there exists a solution u, m to the system* (5.20)–(5.21)*, such that $u, m \in H^1(\mathcal{O}; \mathbb{R}^N)$, $m \geq 0$. Then one has*

$$u \in W^{1,r} \cap L^\infty(\mathcal{O}; \mathbb{R}^N),\ 2 \leq r < r_0, \quad u \in C^{0,\delta}(\overline{\mathcal{O}}; \mathbb{R}^N),\ 0 < \delta \leq \delta_0 < 1, \tag{5.22}$$

where the constants r_0, δ_0 depend only on the constants in the assumptions and the data. They do not depend on the H^1 norm of u, m. The norm of u in the functional spaces $W^{1,r} \cap L^\infty$ and $C^{0,\delta}$ does not depend on the H^1 norm of m.

Remark 5.1 This result extends the traditional additional results of regularity of H^1 solutions to (5.20). The functions m^i appear as an external factor. In view of the weak coupling, only the positivity of m^i is important.

6 A Priori Estimates

The proof of Theorem 5.1 will rely on a priori estimates. Although very close to the treatment in [2] which is done for Dirichlet problems, we develop the main steps of the proof. This will also be helpful at the existence phase. We will indeed consider an approximation procedure, and we shall have to check that the same estimates hold. That will be instrumental in passing to the limit.

6.1 Preliminary steps

Taking $\psi^i = 1$ in (5.21), we obtain

$$\int_\mathcal{O} m^i(x)\mathrm{d}x = \int_\mathcal{O} m^i_0(x)\mathrm{d}x.$$

Since $m \geq 0$, we get immediately

$$m \in L^1(\mathcal{O};\mathbb{R}^N), \quad \|m\|_{L^1(\mathcal{O};\mathbb{R}^N)} = \|m_0\|_{L^1(\mathcal{O};\mathbb{R}^N)}. \tag{6.1}$$

From the assumption (5.9), it follows that

$$\|V^i_{[m]}\|_{L^\infty(\mathcal{O})} \leq l(\|m_0\|). \tag{6.2}$$

Using the assumption (5.6) in the first equation of (5.19) and the standard maximum principle arguments for Neumann elliptic problems, we deduce easily

$$\|u^i\|_{L^\infty(\mathcal{O})} \leq \frac{C_0 + l(\|m_0\|)}{\alpha}. \tag{6.3}$$

Considering the vector $u(x)$ of components $u^i(x)$, we call

$$\|u\|_{L^\infty(\mathcal{O})} = \| \, |u| \, \|_{L^\infty(\mathcal{O})},$$

where $|u|$ is the vector norm. Hence

$$\|u\|_{L^\infty(\mathcal{O})} \leq \sqrt{N}\frac{C_0 + l(\|m_0\|)}{\alpha} = \rho. \tag{6.4}$$

6.2 Basic inequality

Thanks to (6.2), we can simply set $f^i(x) = V^i_{[m]}(x) - \alpha u^i(x)$, and consider that f^i is a given bounded function. We take advantage of the weak coupling of m and u in the first set of the equations (5.20). We now consider a constant vector $c \in \mathbb{R}^N$. This constant vector will be chosen in specific applications of the basic inequality. The only thing that we require is $|c| \leq \rho$. Define next $\widetilde{u} = u - c$, and consider the functions $X^i(s)$ introduced in Subsection 5.2. We associate to these functions $X^i(x) = X^i(\widetilde{u}(x))$. This is a slight abuse of notation, to shorten the notation. The basic inequlity is summarized in the following lemma.

Lemma 6.1 *Let $\Psi \in H^1(\mathcal{O}) \cup L^\infty(\mathcal{O})$, with $\Psi \geq 0$. We have the inequality*

$$\int_{\mathcal{O}} a(x) DX^1 \cdot D\Psi \mathrm{d}x + \underline{a} \int_{\mathcal{O}} \Psi |Du|^2 \mathrm{d}x \leq C(\rho) \int_{\mathcal{O}} \Psi \sum_{i=1}^N (|k^i| + |f^i|) \, \mathrm{d}x, \tag{6.5}$$

where $C(\rho)$ is a constant depending only on ρ and the various constants in (5.13)–(5.16). *This inequlity is obtained for a specific choice of the constants γ^i in the definition of $X^i(s)$. This inequality is valid for any constant vector c with $|c| \leq \rho$.*

We note

$$|Du|^2 = \sum_{i=1}^N |Du^i|^2.$$

Proof The proof is rather technical. Details can be found in [2]. We only sketch here the main steps to facilitate the reading. Consider the functions $X^i(x)$. We have

$$DX^i = \sum_{j=i}^N \gamma^j X^i \cdots X^j (\exp(\gamma^j \widetilde{u}^j) - \exp(-\gamma^j \widetilde{u}^j)) Du^j.$$

We take, in (5.20),

$$\varphi^i = \Psi\gamma^i(\exp(\gamma^i\widetilde{u}^i) - \exp(-\gamma^i\widetilde{u}^i))\prod_{j=1}^{i} X^j.$$

After tedious calculations, we obtain the expression

$$\begin{aligned}
&\sum_{i=1}^{N}\int_{\mathcal{O}} a(x)Du^i \cdot D\varphi^i \mathrm{d}x \\
&= \int_{\mathcal{O}} a(x)DX^1 \cdot D\Psi \mathrm{d}x + \sum_{i=1}^{N}\int_{\mathcal{O}} \Psi a(x)DF^i \cdot DF^i \prod_{j=1}^{i} X^j \mathrm{d}x \\
&\quad + \sum_{i=1}^{N}\int_{\mathcal{O}} \Psi a(x)Du^i \cdot Du^i \prod_{j=1}^{i} X^j (\gamma^i)^2(\exp(\gamma^i\widetilde{u}^i) + \exp(-\gamma^i\widetilde{u}^i))\mathrm{d}x
\end{aligned} \tag{6.6}$$

with $F^i = \log X^i$. On the other hand, from (5.20), we have

$$\sum_{i=1}^{N}\int_{\mathcal{O}} a(x)Du^i \cdot D\varphi^i \mathrm{d}x = \sum_i \int_{\mathcal{O}} (H^i(x, Du) + f^i)\varphi^i \mathrm{d}x.$$

Using (5.12) and performing calculations, we can set

$$\begin{aligned}
&\sum_i \int_{\mathcal{O}} (H^i(x, Du) + f^i)\varphi^i \mathrm{d}x \\
&= \int_{\mathcal{O}} \sum_{i=1}^{N-1} (Q^i - Q^{i-1})DF^i \prod_{j=1}^{i} X^j \mathrm{d}x \\
&\quad + \int_{\mathcal{O}} \Psi \sum_{i=1}^{N} (H_0^i(x, Du) + f^i)\gamma^i(\exp(\gamma^i\widetilde{u}^i) - \exp(-\gamma^i\widetilde{u}^i)) \prod_{j=1}^{i} X^j,
\end{aligned} \tag{6.7}$$

in which $Q^0 = 0$. Using the assumptions (5.13)–(5.16), we can check the inequality

$$\begin{aligned}
&\int_{\mathcal{O}} a(x)DX^1 \cdot D\Psi \mathrm{d}x + \int_{\mathcal{O}} \Psi \sum_{j=1}^{N} |Du^j|^2 B^j(x) \mathrm{d}x \\
&\leq \int_{\mathcal{O}} \Psi \sum_{i=1}^{N} (k^i + f^i)\gamma^i(\exp(\gamma^i\widetilde{u}^i) - \exp(-\gamma^i\widetilde{u}^i)) \prod_{j=1}^{i} X^j \mathrm{d}x,
\end{aligned}$$

where $B^j(x)$ is the long expression

$$\begin{aligned}
B^j(x) &= \underline{a}(\gamma^j)^2(\exp(\gamma^j\widetilde{u}^j) + \exp(-\gamma^j\widetilde{u}^j)) \prod_{h=1}^{j} X^h - \sum_{i=1}^{N-1} \frac{(K^i + K^{i-1})^2}{4\underline{a}^2} \prod_{h=1}^{i} X^h \\
&\quad - \sum_{i=j}^{N} K_j^i \gamma^i(\exp(\gamma^i\widetilde{u}^i) - \exp(-\gamma^i\widetilde{u}^i)) \prod_{h=1}^{i} X^h.
\end{aligned}$$

Rearranging the above expressions, we have

$$\begin{aligned} B^j(x) \geq \Big[&\underline{a}(\gamma^j)^2 + \frac{\underline{a}(\gamma^j)^2 - 2\gamma^j K_j^j}{2}(\exp(\gamma^j \widetilde{u}^j) + \exp(-\gamma^j \widetilde{u}^j)) \\ &- \sum_{i=1}^{j} \frac{(K^i + K^{i-1})^2}{4\underline{a}^2} \sum_{i=j+1}^{N-1} \frac{(K^i + K^{i-1})^2}{4\underline{a}^2} \prod_{h=j+1}^{i} X^h \\ &- \sum_{i=j+1}^{N} K_j^i \gamma^i (\exp(\gamma^i \widetilde{u}^i) - \exp(-\gamma^i \widetilde{u}^i)) \prod_{h=j+1}^{i} X^h \Big] \prod_{h=1}^{j} X^h. \end{aligned}$$

We then chose recursively the constants γ^j, such that

$$\begin{aligned} &\underline{a}\gamma^j - 2K_j^j > 0, \quad \gamma^j > 1, \\ &\underline{a}(\gamma^j)^2 - 2\gamma^j K_j^j - \sum_{i=1}^{j} \frac{(K^i + K^{i-1})^2}{4\underline{a}^2} \\ &> \sum_{i=j+1}^{N-1} \frac{(K^i + K^{i-1})^2}{4\underline{a}^2} \prod_{h=j+1}^{i} X^h - \sum_{i=j+1}^{N} K_j^i \gamma^i (\exp(\gamma^i \widetilde{u}^i) - \exp(-\gamma^i \widetilde{u}^i)) \prod_{h=j+1}^{i} X^h. \end{aligned}$$

This is possibly backward recursively, and the choice of these constants depends only on ρ and the various constants in the assumptions. With this choice of the constants, we get $B^j(x) \geq \underline{a}$ and the result follows easily.

6.3 $W^{1,r}$ estimates

We begin with the $W^{1,r}$ estimate, $2 \leq r < r_0$. Let $\tau(x)$ be a smooth function with $0 \leq \tau(x) \leq 1$ and

$$\begin{aligned} \tau(x) &= 1, \quad \text{if } |x| \leq 1, \\ \tau(x) &= 0, \quad \text{if } |x| \geq 2. \end{aligned}$$

To any point x_0, we associate the ball of center x_0 and radius R, denoted by $B_R(x_0)$. We assume that $R \leq R_0$ but R can be arbitrarily small. We define the cut-off function

$$\tau_R(x) = \tau\Big(\frac{x - x_0}{R}\Big).$$

We then apply the basic inequality (6.5) with $\Psi = \tau_R^2$. We deduce easily

$$\underline{a} \int_{\mathcal{O} \cap B_R} |Du|^2 \mathrm{d}x \leq \frac{\widehat{a}}{R} \int_{\mathcal{O} \cap B_{2R}} |DX^1| \mathrm{d}x + C(\rho) \int_{\mathcal{O} \cap B_{2R}} \sum_{i=1}^{N} (|k^i| + |f^i|) \mathrm{d}x. \tag{6.8}$$

We take $c = c_R$, to be defined below. We have

$$DX^1(x) = \sum_{j=1}^{N} \frac{\partial X^1}{\partial s^j}(u(x) - c_R) Du^j(x).$$

Using Lemma 5.1, we get

$$|DX^1(x)| \leq C^1(\rho) |Du(x)| |u(x) - c_R|.$$

We have then

$$\int_{\mathcal{O}\cap B_{2R}} |DX^1|\mathrm{d}x \le C^1(\rho)\Big(\int_{\mathcal{O}\cap B_{2R}} |Du|^\mu \mathrm{d}x\Big)^{\frac{1}{\mu}}\Big(\int_{\mathcal{O}\cap B_{2R}} |u-c_R|^\lambda \mathrm{d}x\Big)^{\frac{1}{\lambda}}$$

for any $\lambda > 1$ and $\frac{1}{\lambda}+\frac{1}{\mu} = 1$. We next define c_R. We consider points x_0, such that $|\mathcal{O}\cap B_R(x_0)| > 0$. Therefore, $|\mathcal{O}\cap B_{2R}(x_0)| > 0$. We consider two cases, that is, the case of $B_{2R}(x_0)\subset\mathcal{O}$, and the case of $B_{2R}(x_0)\cap(R^n-\mathcal{O})\neq\varnothing$. In the second case, by the smoothness of the domain, $\Gamma\cap B_{2R}(x_0)\neq\varnothing$. We pick a point $x_0'\in\Gamma\cap B_{2R}(x_0)$, and note that $B_{2R}(x_0)\subset B_{4R}(x_0')\subset B_{6R}(x_0)$. Again, from the smoothness of the domain, we have (the sphere condition)

$$|B_{4R}(x_0')\cap\mathcal{O}|\ge c_0R^n,\quad |B_{4R}(x_0')\cap(R^n-\mathcal{O})|\ge c_0R^n,$$

where c_0 is a constant. We then define c_R by

$$c_R = \begin{cases} \dfrac{1}{|B_{2R}|}\displaystyle\int_{B_{2R}} u(x)\mathrm{d}x, & \text{if } B_{2R}(x_0)\subset\mathcal{O},\\ \dfrac{1}{|B_{4R}(x_0')\cap\mathcal{O}|}\displaystyle\int_{B_{4R}(x_0')\cap\mathcal{O}} u(x)\mathrm{d}x, & \text{if } B_{2R}(x_0)\cap(R^n-\mathcal{O})\neq\varnothing. \end{cases}$$

We can state the Poincaré's inequality

$$\Big(\int_{\mathcal{O}\cap B_{2R}} |u-c_R|^\lambda \mathrm{d}x\Big)^{\frac{1}{\lambda}}$$
$$\le c_1R^{n(\frac{1}{\lambda}-\frac{1}{\nu})+1}\Big(\int_{\mathcal{O}\cap B_{6R}} |Du|^\nu \mathrm{d}x\Big)^{\frac{1}{\nu}},\quad \forall\nu,\ \text{such that } 1\le\nu\le 2,\ n\Big(\frac{1}{\lambda}-\frac{1}{\nu}\Big)+1\ge 0.$$

We will apply this inequlity with $n\big(\frac{1}{\lambda}-\frac{1}{\nu}\big)+1 = 0$, i.e., $\nu = \frac{\lambda n}{n+\lambda}$. From the conditions $1\le\nu\le 2$, this is possible only when $n\ge 2$ and

$$\frac{n}{n-1}\le\lambda\le\frac{2n}{n-2}.$$

In that case, we have

$$\Big(\int_{\mathcal{O}\cap B_{2R}} |u-c_R|^\lambda \mathrm{d}x\Big)^{\frac{1}{\lambda}} \le c_1\Big(\int_{\mathcal{O}\cap B_{6R}} |Du|^{\frac{\lambda n}{n+\lambda}}\mathrm{d}x\Big)^{\frac{n+\lambda}{\lambda n}}.$$

We now chose λ, such that $\frac{\lambda n}{n+\lambda} = \mu = \frac{\lambda}{\lambda-1}$. This implies $\lambda = \frac{2n}{n-1}$, which is compatible with the restrictions on λ. Collecting the above results, we can assert that

$$\int_{\mathcal{O}\cap B_{2R}} |DX^1|\mathrm{d}x \le C^2(\rho)\Big(\int_{\mathcal{O}\cap B_{6R}} |Du|^{\frac{2n}{n+1}}\mathrm{d}x\Big)^{\frac{n+1}{n}},$$

where $C^2(\rho)$ is another constant, depending only on ρ. We next note that

$$\int_{\mathcal{O}\cap B_{2R}}\sum_{i=1}^N(|k^i|+|f^i|)\mathrm{d}x\le CR^{n(1-\frac{1}{p})}.$$

Therefore, collecting the above results, we can assert from (6.8) that

$$\int_{\mathcal{O}\cap B_R} |Du|^2\mathrm{d}x \le C^3(\rho)\Big(R^{n(1-\frac{1}{p})}+\frac{1}{R}\Big(\int_{\mathcal{O}\cap B_{6R}} |Du|^{\frac{2n}{n+1}}\mathrm{d}x\Big)^{\frac{n+1}{n}}\Big),\quad \forall R\le R_0. \tag{6.9}$$

Note that this inequlity is trivial if $|\mathcal{O} \cap B_R(x_0)| = 0$. According to Gehring's result (see [2]), we assert that

$$\int_{\mathcal{O}} |Du|^r \mathrm{d}x \leq C(r, \rho), \quad \forall 2 \leq r < r_0, \tag{6.10}$$

where r_0 depends only on ρ and the data.

6.4 $C^{0,\delta}$ estimates

We now turn to the Hölder regularity. To treat the Hölder regularity up to the boundary, we have to use local maps. The regularity is then reduced to interior regularity and regularity on balls centered on the boundary, which can be transformed into half-planes by a straightening operation. We shall again limit ourselves to the main ideas, leaving details to the reference [2]. We begin with the interior regularity.

Let $\widetilde{\mathcal{O}}$ be a smooth domain such that $\overline{\widetilde{\mathcal{O}}} \subset \mathcal{O}$. We shall prove the Hölder regularity on $\overline{\widetilde{\mathcal{O}}}$. Since $\widetilde{\mathcal{O}}$ is arbitrary, that will prove the Hölder regularity on $\mathcal{O}$. Let $x_0 \in \overline{\widetilde{\mathcal{O}}}$. We shall apply the Green function to the Dirichlet problem in $\mathcal{O}$. It is denoted by $G = G^{x_0}$ and defined by

$$\int_{\mathcal{O}} a(x) D\varphi \cdot DG \,\mathrm{d}x = \varphi(x_0), \quad \forall \varphi \in C_0^\infty(\mathcal{O}). \tag{6.11}$$

We shall use the following properties of Green functions (see [2] for details):

$$\begin{aligned} &G \in W_0^{1,\mu}(\mathcal{O}), \quad \forall \mu,\ 1 \leq \mu < \frac{n}{n-1}, \\ &G \in L^\nu(\mathcal{O}), \qquad \forall \nu,\ 1 \leq \nu < \frac{n}{n-2}. \end{aligned} \tag{6.12}$$

Assume $n \geq 3$. Then

$$c_0|x - x_0|^{2-n} \leq G(x) \leq c_1|x - x_0|^{2-n}, \quad \forall x \in Q,\ \forall Q \text{ neighbourhood of } x_0 \text{ with } \overline{Q} \subset \mathcal{O}, \tag{6.13}$$

where the constants c_0, c_1 depend only on $\underline{a}$ and $\overline{a}$.

We next consider the balls $B_R(x_0)$. We assume that $R \leq R_0$, with $2R_0 < \operatorname{dist}(\overline{\widetilde{\mathcal{O}}}, \mathbb{R}^n - \mathcal{O})$. This implies $\overline{B_{2R}(x_0)} \subset \mathcal{O}$. We consider the cut-off function $\tau_R(x)$ as that defined in subsection 6.3.

In the basic inequality (6.5), we choose

$$c = c_R = \frac{1}{|B_{2R} - B_{\frac{R}{2}}|} \int_{B_{2R} - B_{\frac{R}{2}}} u \mathrm{d}x, \quad \Psi = G\tau_R^2,$$

so we get

$$\int_{\mathcal{O}} a(x) DX^1 \cdot D(G\tau_R^2) \mathrm{d}x + \underline{a} \int_{\mathcal{O}} G\tau_R^2 |Du|^2 \mathrm{d}x \leq C(\rho) \int_{\mathcal{O}} G\tau_R^2 \sum_{i=1}^N (|k^i| + |f^i|) \,\mathrm{d}x. \tag{6.14}$$

Clearly $\int_{\mathcal{O}} G\tau_R^2 |Du|^2 \mathrm{d}x \geq \int_{B_{\frac{R}{2}}} G|Du|^2 \mathrm{d}x$, and since $\frac{B_R(x_0)}{2} \subset \overline{B_{2R}(x_0)} \subset \mathcal{O}$, we can use the estimate (6.13) to assert

$$\underline{a} \int_{\mathcal{O}} G\tau_R^2 |Du|^2 \mathrm{d}x \geq C \int_{B_{\frac{R}{2}}} |Du|^2 |x - x_0|^{2-n} \mathrm{d}x, \tag{6.15}$$

where C is a constant. Next

$$\int_{\mathcal{O}} G\tau_R^2 \sum_{i=1}^{N}(|k^i|+|f^i|)\,\mathrm{d}x \le \sum_{i=1}^{N}\int_{B_{2R}} G|k^i|\mathrm{d}x + \sum_{i=1}^{N}\|f^i\|\int_{B_{2R}} G\mathrm{d}x,$$

and from the estimates (6.12) and Hölder's inequality,

$$\int_{B_{2R}} G\mathrm{d}x \le CR^{\frac{n}{\mu'}}, \qquad \forall \mu' \text{ such that } \frac{n}{\mu'} < 2,$$
$$\int_{B_{2R}} G|k^i|\mathrm{d}x \le CR^{n(\frac{1}{p'}-\frac{1}{\nu})}, \quad \forall \nu < \frac{n}{n-2}.$$

Collecting the above results, we can assert that

$$\int_{\mathcal{O}} G\tau_R^2 \sum_{i=1}^{N}(|k^i|+|f^i|)\mathrm{d}x \le CR^{\beta}, \quad \beta < 2. \tag{6.16}$$

Next we have

$$\begin{aligned}\int_{\mathcal{O}} a(x)DX^1\cdot D(G\tau_R^2)\mathrm{d}x &= \int_{\mathcal{O}} a(x)DX^1\cdot DG\,\tau_R^2\mathrm{d}x + 2\int_{\mathcal{O}} a(x)DX^1\cdot D\tau_R\tau_R G\mathrm{d}x\\ &= \mathrm{Z}+\mathrm{I}.\end{aligned}$$

We have, as seen in the previous section,

$$|DX^1(x)| \le C|Du(x)||u(x)-c_R|.$$

Therefore, using again the estimates on the Green function (6.13), we have

$$|\mathrm{I}| \le C\int_{B_{2R}-B_{\frac{R}{2}}} |Du(x)|\frac{|u(x)-c_R|}{R}|x-x_0|^{2-n}\mathrm{d}x.$$

Note that

$$\begin{aligned}&\int_{B_{2R}-B_{\frac{R}{2}}} \frac{|u(x)-c_R|^2}{R^2}|x-x_0|^{2-n}\mathrm{d}x\\ &\le CR^{-n}\int_{B_{2R}-B_{\frac{R}{2}}} |u(x)-c_R|^2\mathrm{d}x\\ &\le CR^{2-n}\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2\mathrm{d}x\\ &\le C\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x,\end{aligned}$$

by Poincaré's inequality. Therefore,

$$|\mathrm{I}| \le C\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x. \tag{6.17}$$

We now turn to the term Z. Recalling the term $X_0^1 \geq 1$ (see (5.18)), we write

$$\begin{aligned}
\mathrm{Z} &= \int_{\mathcal{O}} a(x) D(X^1 - X_0^1) \cdot DG\, \tau_R^2 \mathrm{d}x \\
&= \int_{\mathcal{O}} a(x) D(\tau_R^2(X^1 - X_0^1)) \cdot DG \mathrm{d}x - 2 \int_{\mathcal{O}} a(x) D\tau_R \cdot DG(X^1 - X_0^1)\tau_R \mathrm{d}x \\
&\geq -2 \int_{\mathcal{O}} a(x) D\tau_R \cdot DG(X^1 - X_0^1)\tau_R \mathrm{d}x,
\end{aligned}$$

where we have made use of the Green function's definition (6.11). Recalling (5.17), we have $|X^1(x) - X_0^1| \leq |u(x) - c_R|^2$. Therefore,

$$\begin{aligned}
&\int_{\mathcal{O}} a(x) D\tau_R \cdot DG(X^1 - X_0^1)\tau_R \mathrm{d}x \\
&\leq \frac{C}{R} \int_{B_{2R} - B_R} |u - c_R|^2 |DG| \tau_R \mathrm{d}x \\
&\leq C \int_{B_{2R} - B_R} \frac{|u - c_R|^2}{R^2} G \mathrm{d}x + C \int_{B_{2R} - B_R} |u - c_R|^2 |DG|^2 G^{-1}\, \tau_R^2 \mathrm{d}x \\
&\leq C \int_{B_{2R} - B_{\frac{R}{2}}} |Du|^2 |x - x_0|^{2-n} \mathrm{d}x + C\mathrm{Y}.
\end{aligned}$$

Therefore, we have

$$\mathrm{Z} \geq -C \int_{B_{2R} - B_{\frac{R}{2}}} |Du|^2 |x - x_0|^{2-n} \mathrm{d}x - C\mathrm{Y}. \tag{6.18}$$

We now estimate

$$\mathrm{Y} = \int_{B_{2R} - B_R} |u - c_R|^2 |DG|^2 G^{-1}\, \tau_R^2 \mathrm{d}x.$$

We introduce a new cut-off function

$$\xi(x) = \begin{cases} 0 & \text{for } |x| \leq \dfrac{1}{2}, \\ \tau(x) & \text{for } |x| \geq 1, \end{cases}$$

and denote $\xi_R(x) = \xi\big(\frac{x - x_0}{R}\big)$. Hence

$$\begin{aligned}
&\xi_R(x) = \tau_R(x) \quad \text{on } B_{2R} - B_R, \\
&\xi_R(x) = 0 \qquad\quad \text{on } B_{\frac{R}{2}}.
\end{aligned}$$

In the Green function equation (6.11), we take

$$\varphi = G^{-\frac{1}{2}} |u - c_R|^2 \xi_R^2.$$

Noting that $\varphi(x_0) = 0$, it follows that

$$\int_{\mathcal{O}} G^{-\frac{1}{2}} a(x) D(|u - c_R|^2 \xi_R^2) \cdot DG \mathrm{d}x = \frac{1}{2} \int_{\mathcal{O}} G^{-\frac{3}{2}} a(x) DG \cdot DG\, |u - c_R|^2 \xi_R^2 \mathrm{d}x. \tag{6.19}$$

Next, in (5.20), we take

$$\varphi^i = (u^i - c_R^i) G^{\frac{1}{2}} \xi_R^2.$$

We obtain, after rearrangements,

$$\begin{aligned}&\sum_{i=1}^{N}\int_{\mathcal{O}}a(x)Du^i\cdot Du^iG^{\frac{1}{2}}\xi_R^2\mathrm{d}x+\frac{1}{4}\int_{\mathcal{O}}a(x)D(|u-c_R|^2\xi_R^2)\cdot DGG^{-\frac{1}{2}}\mathrm{d}x\\&-\frac{1}{2}\int_{\mathcal{O}}a(x)D\xi_R\cdot DG\xi_R|u-c_R|^2G^{-\frac{1}{2}}\mathrm{d}x+\int_{\mathcal{O}}a(x)D(|u-c_R|^2)\cdot D\xi_R\xi_RG^{\frac{1}{2}}\mathrm{d}x\\=&\int_{\mathcal{O}}\sum_{i=1}^{N}(H^i+f^i)(u^i-c_R^i)G^{\frac{1}{2}}\xi_R^2\mathrm{d}x.\end{aligned}$$

Hence,

$$\begin{aligned}&\int_{\mathcal{O}}a(x)D(|u-c_R|^2\xi_R^2)\cdot DGG^{-\frac{1}{2}}\mathrm{d}x\\\le&\ 2\int_{\mathcal{O}}a(x)D\xi_R\cdot DG\xi_R|u-c_R|^2G^{-\frac{1}{2}}\mathrm{d}x+CR^{\frac{2-n}{2}}\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2\mathrm{d}x+CR^{1+\frac{n}{2}-\frac{n}{p}}.\end{aligned}$$

Therefore, from (6.19), we can write

$$\begin{aligned}&\frac{1}{2}\int_{\mathcal{O}}G^{-\frac{3}{2}}a(x)DG\cdot DG\,|u-c_R|^2\xi_R^2\mathrm{d}x\\\le&\ 2\int_{\mathcal{O}}a(x)D\xi_R\cdot DG\xi_R|u-c_R|^2G^{-\frac{1}{2}}\mathrm{d}x+CR^{\frac{2-n}{2}}\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2\mathrm{d}x+CR^{1+\frac{n}{2}-\frac{n}{p}},\end{aligned}$$

from which one easily deduces

$$\begin{aligned}&\int_{\mathcal{O}}G^{-\frac{3}{2}}|DG|^2\,|u-c_R|^2\xi_R^2\mathrm{d}x\\\le&\ C\int_{B_{2R}-B_{\frac{R}{2}}}\frac{|u-c_R|^2}{R^2}G^{\frac{1}{2}}\mathrm{d}x+CR^{\frac{2-n}{2}}\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2\mathrm{d}x+CR^{1+\frac{n}{2}-\frac{n}{p}}.\end{aligned}$$

Hence

$$\int_{\mathcal{O}}G^{-\frac{3}{2}}|DG|^2\,|u-c_R|^2\xi_R^2\mathrm{d}x\le CR^{\frac{2-n}{2}}\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2\mathrm{d}x+CR^{1+\frac{n}{2}-\frac{n}{p}}.$$

Since $\tau_R=\xi_R$ on $B_{2R}-B_R$, we have

$$\begin{aligned}\mathrm{Y}&=\int_{B_{2R}-B_R}|u-c_R|^2|DG|^2G^{-1}\,\xi_R^2\mathrm{d}x\\&\le\int_{B_{2R}-B_{\frac{R}{2}}}|u-c_R|^2|DG|^2G^{-1}\,\xi_R^2\mathrm{d}x\\&\le CR^{\frac{2-n}{2}}\int_{B_{2R}-B_{\frac{R}{2}}}|u-c_R|^2|DG|^2G^{-\frac{3}{2}}\,\xi_R^2\mathrm{d}x\\&\le CR^{2-n}\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2\mathrm{d}x+CR^{2-\frac{n}{p}}\\&\le C\int_{B_{2R}-B_{\frac{R}{2}}}|Du|^2|x-x_0|^{2-n}\mathrm{d}x+CR^{2-\frac{n}{p}}.\end{aligned}$$

Therefore, from (6.18), we obtain

$$\mathrm{Z} \geq -C\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x - CR^{2-\frac{n}{p}}.$$

From (6.14)–(6.17), we obtain

$$\int_{B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x \leq C\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x - \mathrm{Z} + CR^{\beta}.$$

Noting that $0 < 2 - \frac{n}{p} < 2$, and changing the constant β to another possible constant strictly less than 2, we get

$$\int_{B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x \leq C\int_{B_{2R}-B_{\frac{R}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x + CR^{\beta},$$

or

$$\int_{B_R} |Du|^2|x-x_0|^{2-n}\mathrm{d}x \leq C\int_{B_{4R}-B_R} |Du|^2|x-x_0|^{2-n}\mathrm{d}x + CR^{\beta}, \quad \forall R \leq \frac{R_0}{2}. \tag{6.20}$$

Now, going back to the basic inequality (6.5) and taking $\Psi = G$, we deduce

$$\int_{\mathcal{O}} a(x)DX^1 \cdot DG\mathrm{d}x + \underline{a}\int_{\mathcal{O}} G|Du|^2\mathrm{d}x \leq C(\rho)\int_{\mathcal{O}} G\sum_{i=1}^{N}(|k^i| + |f^i|)\,\mathrm{d}x.$$

From the definition of the Green function , the first integral is positive, and the third integral is bounded. Hence

$$\int_{\mathcal{O}} G|Du|^2\mathrm{d}x \leq C,$$

and also $\int_{2R_0} G|Du|^2\mathrm{d}x \leq C$. Hence

$$\int_{B_{\frac{R_0}{2}}} |Du|^2|x-x_0|^{2-n}\mathrm{d}x \leq C. \tag{6.21}$$

From (6.20)–(6.21), using the hole filling technique (see [2]), we can find $\delta_0 \leq \frac{\beta}{2}$, depending only on the data and ρ, such that for $\delta < \delta_0$, one has

$$\mathbb{R}^{2-n-2\delta}\int_{B_R(x_0)} |Du|^2\mathrm{d}x \leq C, \quad \forall R \leq \frac{R_0}{2},\ x_0 \in \overline{\tilde{\mathcal{O}}}.$$

From Hölder's inequality, it follows that

$$\int_{B_R(x_0)} |Du|\mathrm{d}x \leq CR^{n-1+\delta}, \quad \forall R \leq \frac{R_0}{2},\ x_0 \in \overline{\tilde{\mathcal{O}}}.$$

From Morrey's theorem, we obtain that $u^i \in C^{0,\delta}(\overline{\tilde{\mathcal{O}}})$.

So the interior Hölder regularity has been proven. To proceed on the closure, we consider a system of local maps, and prove the regularity on each of them. So we consider a ball B, centered on a point of the boundary, and we assume that there exists a diffeomorphism Ψ from B into $\mathbb{R}^n$, such that

$$\mathcal{O}^+ = \Psi(B \cap \mathcal{O}) \subset \{y \in \mathbb{R}^n \mid y_n > 0\},$$
$$\Gamma' = \Psi(B \cap -) \subset \{y \in \mathbb{R}^n \mid y_n = 0\}.$$

We also define the set obtained from $\mathcal{O}^+$ by reflection, namely,

$$\mathcal{O}^- = \{y \mid y_n < 0,\ (y_1, \cdots, y_{n-1}, -y_n) \in \mathcal{O}^+\},$$

and set

$$\mathcal{O}' = \mathcal{O}^+ \cup \mathcal{O}^- \cup \Gamma'.$$

Then $\mathcal{O}'$ is a bounded domain of $\mathbb{R}^n$. We consider, in (2.22), functions φ^i, which are in $H^1(\mathcal{O} \cap B)$, such that $\varphi^i|_{\mathcal{O}\cap\partial B} = 0$. These functions are extended by 0 on $\mathcal{O} - \mathcal{O} \cap B$. Therefore, (2.22) becomes

$$\int_{B\cap\mathcal{O}} a(x)Du^i \cdot D\varphi^i \mathrm{d}x = \int_{B\cap\mathcal{O}} (H^i(x, Du) + f^i)\varphi^i \mathrm{d}x. \tag{6.22}$$

We then make the change of coordinates $x = \Psi^{-1}(y)$. We call $v^i(y) = u^i(\Psi^{-1}(y))$. Consider the matrix

$$J_\Psi(x) = \mathrm{matrix}\Big(\frac{\partial \Psi_k}{\partial x_l}\Big),$$

and set

$$\widetilde{a}(y) = \frac{J_\Psi(\Psi^{-1}(y))a(\Psi^{-1}(y))J_\Psi^*(\Psi^{-1}(y))}{|\det J_\Psi(\Psi^{-1}(y))|},$$
$$\widetilde{H^i}(y, Dv) = \frac{H^i(\Psi^{-1}(y), J_\Psi(\Psi^{-1}(y))Dv)}{|\det J_\Psi(\Psi^{-1}(y))|},$$
$$\widetilde{f^i}(y) = \frac{f^i(\Psi^{-1}(y))}{|\det J_\Psi(\Psi^{-1}(y))|}.$$

Naturally, the notation Dv refers to the gradient with respect to the variables v. Moreover,

$$J_\Psi(\Psi^{-1}(y))Dv = (J_\Psi(\Psi^{-1}(y))Dv^1, \cdots, J_\Psi(\Psi^{-1}(y))Dv^n),$$

so, in fact,

$$\widetilde{H^i}(y, q) = \frac{H^i(\Psi^{-1}(y), J_\Psi(\Psi^{-1}(y))q^1, \cdots, J_\Psi(\Psi^{-1}(y))q^n)}{|\det J_\Psi(\Psi^{-1}(y))|}.$$

The system (2.19) can be written as

$$\int_{\mathcal{O}^+} \widetilde{a}(y)Dv^i \cdot D\widetilde{\varphi}^i \mathrm{d}y = \int_{\mathcal{O}^+} (\widetilde{H^i}(y, Dv) + \widetilde{f^i})\widetilde{\varphi}^i \mathrm{d}y \tag{6.23}$$

for any $\widetilde{\varphi}^i(y) \in H^1 \cap L^\infty(\mathcal{O}^+)$, such that $\widetilde{\varphi}^i(y) = 0$ on $\Psi(\mathcal{O} \cap \partial B) = \partial\mathcal{O}^+ - \Gamma'$.

We then proceed with a reflexion procedure. Writing $y = (y', y_n)$, we define, for $y_n < 0$,

$$\widetilde{a}_{kk}(y', y_n) = \widetilde{a}_{kk}(y', -y_n), \quad \forall i,$$
$$\widetilde{a}_{kl}(y', y_n) = \widetilde{a}_{kl}(y', -y_n), \quad \forall k, l \neq n,$$
$$\widetilde{a}_{kn}(y', y_n) = \widetilde{a}_{kn}(y', -y_n), \quad \forall k \neq n,$$
$$\widetilde{H^i}(y', y_n; q^1, \cdots, q^{n-1}, q^n) = \widetilde{H^i}(y', -y_n; q^1, \cdots, q^{n-1}, -q^n),$$
$$\widetilde{f^i}(y', y_n) = \widetilde{f^i}(y', -y_n).$$

If we extend the solutions $v^i(y)$ to (6.23) for $y_n < 0$, by setting

$$v^i(y', y_n) = v^i(y', -y_n),$$

then it is easy to convince oneself that the functions $v^i(y)$ are in $H^1 \cap L^\infty(\mathcal{O}')$, and satisfy

$$\int_{\mathcal{O}'} \widetilde{a}(y) Dv^i \cdot D\widetilde{\varphi}^i \mathrm{d}y = \int_{\mathcal{O}'} (\widetilde{H^i}(y, Dv) + \widetilde{f}^i)\widetilde{\varphi}^i \mathrm{d}y, \quad \forall \widetilde{\varphi}^i \in H_0^1 \cap L^\infty(\mathcal{O}'). \tag{6.24}$$

Moreover, the functions $\widetilde{a}(y), \widetilde{H^i}(y,q)$ and $\widetilde{f}^i(y)$ satisfy the same assumptions as $a(x)$, $H^i(x,q)$ and $f^i(x)$, respectively. Therefore, we can obtain the interior $C^{0,\delta}$ regularity of $v^i(y)$ on $\mathcal{O}'$. We thus obtain the $C^{0,\delta}$ regularity including points of the interior of Γ'. By taking a covering of the boundary Γ of $\mathcal{O}$ by a finite number of local maps, we complete the proof of the $C^{0,\delta}(\overline{\mathcal{O}})$ of the function u. This completes the proof of Theorem 5.1.

6.5 Alternative assumptions

Assumptions (5.12)–(5.16) are not the only possible ones. Those were made in order to apply the method used in Lemma 6.1, which we call "exponential domination". They have been introduced in [6]. A certain form of exponential domination can be found already in [12]. An alternative condition replaces the growth condition for the Hamiltonians from below (resp. above) by the "sum coerciveness" of the Hamiltonians. This was first used in [3–4], and thereafter in [6–7]. In the case, the dimension $n = 2$ and the conditions, up to now, are better than those in the n-dimensional case. The first conditions are as usual. The conditions can be written as

$$|H^i(x,q)| \le K(|q|^2 + 1), \tag{6.25}$$

$$H^i(x,q) = H^{i0}(x,q) + q^i \cdot G(x,q), \tag{6.26}$$

$$|G(x,q)| \le K(|q| + 1). \tag{6.27}$$

In applications to the control theory, the term $q^i \cdot G(x,q)$ is derived from the dynamics, and the term $H^{i0}(x,q)$ is derived from the cost of the controls and the influence of nonmarket interaction.

In addition, from below (alternatively from above), the following "sum coerciveness" of the $H^{i0}(x,q)$ is assumed

$$\sum_{i=1}^N H^{i0}(x,q) \ge c_0 |Bq|^2 - K, \quad c_0 > 0, \tag{6.28}$$

and $B : \mathbb{R}^{nN} \to \mathbb{R}^m$ satisfies

$$|Bq| \le K(|q| + 1). \tag{6.29}$$

Of course, this applies to $B =$ identity, but in applications, B can be degenerate, i.e., $B^{-1}(0)$ may be nontrivial.

For G, we need the slightly stronger growth condition

$$|G(x,q)| \le K(|Bq| + 1). \tag{6.30}$$

In applications, B is the map, which assigns to the variables q the corresponding Nash equilibrium for controls in the Lagrangians (see (4.10)). In this context, to a certain extent, (6.30) and (6.28) are natural.

Finally, in [6–7], for $n = 2$, we assume that the above inequality holds. Then we have

$$H^{i0}(x,q) \leq K(|q^i||Bq| + 1), \tag{6.31}$$

which also has a reasonable interpretation in control theory.

In this framework, in [6–7], one obtains the $C^{0,\delta}$ regularity for Bellman systems.

$$Au^i + \alpha u^i = H^i(x, Du)$$

is recalled under the restriction $n = 2$. The techniques can be used for the present mean field setting. Hence, the results of Theorem 5.1 will hold under the assumptions (6.25)–(6.31).

To get rid of the dimension condition, a partial progress was achieved in [8]. They use the same assumptions as (6.25)–(6.26), but they replace (6.31) with

$$H^i(x,q) \leq K(|q^i|^2 + q^i \cdot G_0(x,q) + 1), \tag{6.32}$$

$$|G_0(x,q)| \leq K(|q| + 1), \tag{6.33}$$

and $G_0(x,q)$ can be different from $G(x,q)$ above, which increases applicability. Then (6.25)–(6.30) and (6.32) can be used for our mean field setting for $n \geq 2$, in order to obtain Theorem 5.1.

Concerning weak solutions, [1] showed that the conditions (6.25) and (6.31) of the 2-dimensional case imply the existence of a weak solution $u \in L^\infty \cap H^1$ and the strong convergence of the approximations in H^1, also in dimension $n \geq 3$.

There are several slight generalizations. One may replace (6.26) and (6.28) by

$$\sum_{i=1}^N H^i(x,q) \geq c_0|Bq|^2 - K|Bq|\Big|\sum_{i=1}^N q^i\Big| - K\Big|\sum_{i=1}^N q^i\Big|^2 - K,$$

where the function $G(x,q)$ is not needed. Perturbations of type $\big|\sum_{i=1}^N q^i\big|^2$ are allowed in this setting.

6.6 Full regularity for u

We can complete Theorem 5.1, and state the full regularity of u, provided that an additional assumption is made.

Theorem 6.1 *We make all the assumptions of Theorem* 5.1 *and*

$$a(x) \in W^{1,\infty}(\mathcal{O}), \quad k^i(x) \in L^\infty(\mathcal{O}), \quad i = 1, \cdots, N. \tag{6.34}$$

Then $u \in W^{2,r}(\mathcal{O}; \mathbb{R}^N), \forall 1 \leq r < \infty$. *The norm of* u *in the functional space depends only on the data and the constants in the assumptions.*

Proof The proof is based on the linear theory of elliptic equations. It follows from a bootstrap argument based on the Miranda-Nirenberg interpolation result and the regularity theory of linear elliptic equations. The details can be found in [2].

7 Study of the Field Equations

By field equations, we consider the equations (5.21).

7.1 Generic equation

We shall make the assumptions of Theorem 6.1. We can then assume that the functions $G^i(x, Du)$ are bounded. From Theorem 6.1, the bound depends only on the data, not on the $H^1(\mathcal{O})$ norm of m.

We can see in the equations (5.21) that there exists no coupling in the functions m^i. So it is sufficient to consider a generic problem

$$\begin{aligned}&\int_{\mathcal{O}} a^*(x)Dm(x)\cdot D\psi(x)\mathrm{d}x - \int_{\mathcal{O}} m(x)G(x)\cdot D\psi(x)\mathrm{d}x + \alpha\int_{\mathcal{O}} m(x)\psi(x)\mathrm{d}x\\ &= \alpha\int_{\mathcal{O}} m_0(x)\psi(x)\mathrm{d}x,\end{aligned} \tag{7.1}$$

where $G(x)$ is bounded and $m^0 \geq 0$ is in $L^p(\mathcal{O})$, $p > \frac{n}{2}$. However, we assume that there exists a positive $H^1(\mathcal{O})$ solution to (7.1). The test function $\psi(x)$ in (7.1) can be taken in $H^1(\mathcal{O})$.

7.2 L^∞ bound

An important step is as follows.

Proposition 7.1 *We make the assumptions of Theorem* 6.1*. A positive $H^1(\mathcal{O})$ solution to* (7.1) *is in $L^\infty(\mathcal{O})$ with a norm, which depends only on the data and the constants, and not on the $H^1(\mathcal{O})$ norm of m.*

Proof The proof relies on the properties of the Green function for the Neumann problem. For any $x_0 \in \mathcal{O}$, consider the solution $\Sigma = \Sigma^{x_0}$ to the equation

$$\int_{\mathcal{O}} a(x)D\Sigma\cdot D\psi\mathrm{d}x + \alpha\int_{\mathcal{O}}\Sigma\psi\mathrm{d}x = \alpha\psi(x_0),\quad \forall\psi\in H^1(\mathcal{O})\cap C^0(\overline{\mathcal{O}}). \tag{7.2}$$

The function $\Sigma = \Sigma^{x_0}$ is the Green function associated to the point x_0. We have included the coefficient α for convenience. In (6.11), we had considered the Green function for the Dirichlet problem. We shall use properties, similar to (6.12),

$$\Sigma \in L^{\frac{n}{n-2}(1-s)},\quad D\Sigma\in L^{\frac{n(1-s)}{n-1-s}},\quad \forall\, 0<s<1. \tag{7.3}$$

We take $s < \frac{1}{n-1}$ having exponents strictly larger than 1. The second exponent is strictly less than 2, as soon as $n \geq 3$. We shall take $\psi = m$ in (7.2), and $\psi = \Sigma$ in (7.1). This is formal, since we do not have the smoothness required. The correct approach is to approximate Σ with smoother functions, in smoothing the Dirac measure which comes in (7.2). We skip this step, which is classical. Note that $\Sigma \geq 0$. Comparing the two relations, we obtain

$$\alpha m(x_0) = \alpha\int_{\mathcal{O}} m_0\Sigma\mathrm{d}x + \int_{\mathcal{O}} mG\cdot D\Sigma\mathrm{d}x. \tag{7.4}$$

We stress that this writing is formal, since m is not continuous, and the third integral is not well defined. For the a priori estimates, it is sufficient.

We note first that

$$\int_{\mathcal{O}} m_0 \Sigma \mathrm{d}x \leq \|m_0\|_{L^p} \|\Sigma\|_{L^{\frac{p}{p-1}}}.$$

Using the first property (7.3), thanks to the assumption $p > \frac{n}{2}$, $\frac{p}{p-1} < \frac{n}{n-2}$ and the integral on the right-hand side is well defined.

Now, for any L,

$$\begin{aligned}\int_{\mathcal{O}} mG \cdot D\Sigma \mathrm{d}x &= \int_{\mathcal{O}\cap\{G\cdot D\Sigma \geq L\}} mG \cdot D\Sigma \mathrm{d}x + \int_{\mathcal{O}\cap\{G\cdot D\Sigma \leq L\}} mG \cdot D\Sigma \mathrm{d}x \\ &\leq L \int_{\mathcal{O}} m_0 \mathrm{d}x + \|m\|_\infty \int_{\mathcal{O}\cap\{G\cdot D\Sigma \geq L\}} G \cdot D\Sigma \mathrm{d}x.\end{aligned}$$

Set $z = (G \cdot D\Sigma)^+$. From the second property (7.3), we have

$$\int_{\mathcal{O}} z^{\frac{n(1-s)}{n-1-s}} \mathrm{d}x \leq C_s.$$

Therefore, we check easily that

$$\int_{\mathcal{O}\cap\{z\geq L\}} z \mathrm{d}x \leq C_s \frac{1}{L^{\frac{1-s(n-1)}{n-1-s}}}.$$

Collecting the above results, and choosing L sufficiently large, we deduce from (7.4) that $\|m\|_\infty \leq C$, where the constant depends only on the data, not on the $H^1(\mathcal{O})$ norm of m. This completes the proof.

7.3 Regularity of m

We can write (7.1) as

$$\int_{\mathcal{O}} a^*(x) Dm(x) \cdot D\psi(x) \mathrm{d}x + \alpha \int_{\mathcal{O}} m(x)\psi(x) \mathrm{d}x = \alpha \int_{\mathcal{O}} m_0 \psi(x) \mathrm{d}x + \int_{\mathcal{O}} g(x) \cdot D\psi(x) \mathrm{d}x, \quad (7.5)$$

where $g(x)$ is a bounded function, with a bound depending only on the data. It follows immediately that the $H^1(\mathcal{O})$ norm of m depends only on the data and constants of the assumptions. We can then state it as follows.

Theorem 7.1 *We make the assumptions of Theorem* (6.1). *Then the solution* m *to* (7.5) *belongs to* $W^{2,p}(\mathcal{O}) \oplus W^{1,r}(\mathcal{O}),\ \forall r < \infty$.

The norm depends only on the data and the constants of the assumptions.

Proof This is an immediate consequence of the regularity of the solutions to the linear problems of type (7.5).

8 Existence of Solutions

We can now address the issue of existence of solutions to the system (5.20)–(5.21), with smooth solutions and positive m^i. We shall assume, in addition to the assumptions of Theorem 6.1, that

$$H^i(x, q),\ G^i(x, q) \text{ are continuous in } q \text{ (Caratheodory)}. \quad (8.1)$$

8.1 Approximation procedure

We begin by defining an approximation procedure. We introduce the following notations:

$$H^{i,\epsilon}(x,q) = \frac{H^i(x,q)}{1+\epsilon|H^i(x,q)|}, \quad V^{i,\epsilon}_{[m]}(x) = V^i_{[\frac{m}{1+\epsilon|m|}]}(x), \tag{8.2}$$

and the function on R

$$h^\epsilon(\mu) = \frac{\mu^+}{1+\epsilon|\mu|}, \tag{8.3}$$

where $|m|$ is the norm of the vector m. Clearly

$$\Big\| \frac{m}{1+\epsilon|m|} \Big\|_{L^1(\mathcal{O};\mathbb{R}^N)} \leq \frac{|\mathcal{O}|}{\epsilon}.$$

Hence,

$$\|V^{i,\epsilon}_{[m]}\|_{L^\infty} \leq l\Big(\frac{|\mathcal{O}|}{\epsilon}\Big), \quad |H^{i,\epsilon}(x,q)| \leq \frac{1}{\epsilon}, \quad \forall m, x, q. \tag{8.4}$$

We then define a function T^ϵ from $H^1(\mathcal{O};\mathbb{R}^N) \times L^2(\mathcal{O};\mathbb{R}^N)$ into itself as follows. We write

$$(u,m) = T^\epsilon(v,\mu),$$

and u, m are the solutions to

$$\begin{aligned} &\int_{\mathcal{O}} a(x)Du^i(x)\cdot D\varphi^i(x)\mathrm{d}x + \alpha\int_{\mathcal{O}} u^i(x)\varphi^i(x)\mathrm{d}x \\ &= \int_{\mathcal{O}} (H^{i,\epsilon}(x,Dv) + V^{i,\epsilon}_{[\mu]}(x))\varphi^i(x)\mathrm{d}x, \end{aligned} \tag{8.5}$$

$$\begin{aligned} &\int_{\mathcal{O}} a^*(x)Dm^i(x)\cdot D\psi^i(x)\mathrm{d}x + \alpha\int_{\mathcal{O}} m^i(x)\psi^i(x)\mathrm{d}x \\ &= \int_{\mathcal{O}} h^\epsilon(\mu^i(x))G^i(x,Du)\cdot D\psi^i(x)\mathrm{d}x + \alpha\int_{\mathcal{O}} m^i_0(x)\psi^i(x)\mathrm{d}x. \end{aligned} \tag{8.6}$$

Note that the problems (8.5)–(8.6) are defined in sequence. In the right-hand side of (8.6), there is Du, not Dv. At any rate, u^i and m^i are solutions to linear problems. Using the linearity and the regularity theory of linear elliptic equations, we can assert that

$$u \in W^{2,r}(\mathcal{O};\mathbb{R}^N), \quad m \in W^{1,r}(\mathcal{O};\mathbb{R}^N) \oplus W^{2,p}(\mathcal{O};\mathbb{R}^N), \quad \forall r < \infty. \tag{8.7}$$

Moreover, the norm in these functional spaces is bounded by a fixed number, depending on ϵ, but not on the arguments v, μ. The map T^ϵ is continuous (thanks to (8.1)), and the image $T^\epsilon(v,\mu)$ remains in a fixed compact convex subset of $H^1(\mathcal{O};\mathbb{R}^N) \times L^2(\mathcal{O};\mathbb{R}^N)$. From Leray-Schauder theorem, the map T^ϵ has a fixed point. Therefore, we have obtained the following lemma.

Lemma 8.1 *Under the assumptions of Theorem* 6.1 *and* (8.1)*, there exists a pair* u^ϵ, m^ϵ, *belonging to the functional spaces as in* (8.7) *and satisfying the system of equations*

$$\begin{aligned}
&\int_{\mathcal{O}} a(x)Du^{i,\epsilon}(x)\cdot D\varphi^i(x)\mathrm{d}x+\alpha\int_{\mathcal{O}} u^{i,\epsilon}(x)\varphi^i(x)\mathrm{d}x\\
&=\int_{\mathcal{O}}(H^{i,\epsilon}(x,Du^\epsilon)+V^{i,\epsilon}_{[m^\epsilon]}(x))\varphi^i(x)\mathrm{d}x,
\end{aligned}\tag{8.8}$$

$$\begin{aligned}
&\int_{\mathcal{O}} a^*(x)Dm^{i,\epsilon}(x)\cdot D\psi^i(x)\mathrm{d}x+\alpha\int_{\mathcal{O}} m^{i,\epsilon}(x)\psi^i(x)\mathrm{d}x\\
&=\int_{\mathcal{O}} h^\epsilon(m^{i,\epsilon}(x))G^i(x,Du^\epsilon)\cdot D\psi^i(x)\mathrm{d}x+\alpha\int_{\mathcal{O}} m_0^i(x)\psi^i(x)\mathrm{d}x.
\end{aligned}\tag{8.9}$$

8.2 Main result

We can now state the main existence result.

Theorem 8.1 *Under the assumptions of Theorem* 6.1 *and* (8.1), *there exists a solution* (u,m) *to the system of equations* (5.20)–(5.21), *such that*

$$u\in W^{2,r}(\mathcal{O};\mathbb{R}^N),\quad m\in W^{1,r}(\mathcal{O};\mathbb{R}^N)\oplus W^{2,p}(\mathcal{O};\mathbb{R}^N),\quad \forall r<\infty.\tag{8.10}$$

Proof The first thing to observe is that the fixed point (u^ϵ,m^ϵ), i.e., the solution to (8.8)–(8.9), satisfies $m^\epsilon\geq 0$. This is easily seen by taking $\psi^i=(m^{i,\epsilon})^-$ in the equation (8.9) and noting that

$$\int_{\mathcal{O}} h^\epsilon(m^{i,\epsilon}(x))G^i(x,Du^\epsilon)\cdot D(m^{i,\epsilon})^-(x)\mathrm{d}x=0.$$

Therefore, we can write (8.9) as follows:

$$\begin{aligned}
&\int_{\mathcal{O}} a^*(x)Dm^{i,\epsilon}(x)\cdot D\psi^i(x)\mathrm{d}x+\alpha\int_{\mathcal{O}} m^{i,\epsilon}(x)\psi^i(x)\mathrm{d}x\\
&=\int_{\mathcal{O}}\frac{m^{i,\epsilon}(x)}{1+\epsilon|m^{i,\epsilon}(x)|}G^i(x,Du^\epsilon)\cdot D\psi^i(x)\mathrm{d}x+\alpha\int_{\mathcal{O}} m_0^i(x)\psi^i(x)\mathrm{d}x,
\end{aligned}\tag{8.11}$$

and $m^{i,\epsilon}\geq 0$. But then, by taking $\psi^i=1$, we get

$$\int_{\mathcal{O}} m^{i,\epsilon}(x)\mathrm{d}x=\int_{\mathcal{O}} m_0^i(x)\mathrm{d}x,$$

and we deduce $|V^{i,\epsilon}_{[m^\epsilon]}(x)|\leq l(\|m_0\|)$. Noting that $H^{i,\epsilon}(x,q)$ satisfies all the estimates of $H^i(x,q)$ and (5.4)–(5.6), we can apply all the techniques to (8.8) to derive the a priori estimates in Theorems 5.1 and 6.1. We can then assert that

$$\|u^\epsilon\|_{W^{2,r}(\mathcal{O},\mathbb{R}^N)}\leq C,\quad \forall r<\infty,$$

and the constant does not depend on ϵ. Similarly, the reasoning made in Proposition 7.1 and Theorem 7.1 carries over for m^ϵ in (8.11). We then obtain that

$$\|m^\epsilon\|_{W^{1,r_0}}\leq C,\quad r_0>\frac{n}{2}.$$

We can then extract subsequences, still denoted u^ϵ, m^ϵ, such that, among other properties,

$$u^\epsilon \to u, \quad Du^\epsilon \to Du, \quad \text{pointwise } \|u^\epsilon\|_{L^\infty}, \|Du^\epsilon\|_{L^\infty} \leq C,$$
$$m^\epsilon \to m, \quad \text{pointwise } Dm^\epsilon \to Dm \quad \text{weakly in } L^2 \; \|m^\epsilon\| \leq C.$$

Using (5.10) and the continuity properties of $G^i(x,q), H^i(x,q)$, it is easy to go to the limit as $\epsilon \to 0$ in equations (8.8), (8.11), and obtain a solution to (5.20)–(5.21) with the regularity (8.10). This completes the proof of Theorem 8.1.

References

[1] Bensoussan, A., Bulicěk, M. and Frehse, J., Existence and Compactness for weak solutions to Bellman systems with critical growth, *Discrete Contin. Dyn. Syst., Ser. B*, **17**(6), 2012, 1–21.

[2] Bensoussan, A. and Frehse, J., Regularity results for nonlinear elliptic systems and applications, Appl. Math. Sci., **151**, Springer-Verlag, Berlin, 2002.

[3] Bensoussan, A. and Frehse, J., Ergodic Bellman systems for stochastic games, Markus Feestricht Volume Differential Equations, Dynamical Systems and Control Sciences, Lecture Notes in Pure and Appl. Math., K. D. Elworty, W., Norrie Everitt, E. Bruce Lee, M. Dekker (eds.), Vol. 152, 1993, 411–421.

[4] Bensoussan, A. and Frehse, J., Ergodic Bellman systems for stochastic games in arbitrary dimension, *Proc. Royal Soc., London, Math. Phy. Sci. A*, **449**, 1995, 65–67.

[5] Bensoussan, A. and Frehse, J., Smooth solutions of systems of quasilinear parabolic equations, *ESAIM: Control, Optimization and Calculus of Variations*, **8**, 2002, 169–193.

[6] Bensoussan, A., Frehse, J. and Vogelgesang, J., Systems of Bellman equations to stochastic differential games with noncompact coupling, *Discrete Contin. Dyn. Syst., Ser. A*, **274**, 2010, 1375–1390.

[7] Bensoussan, A., Frehse, J. and Vogelgesang, J., Nash and stackelberg differential games, *Chin. Ann. Math.*, **33B**(3), 2012, 317–332.

[8] Bulicěk, M. and Frehse, J., On nonlinear elliptic Bellman systems for a class of stochastic differential games in arbitrary dimension, *Math. Models Methods Appl. Sci.*, **21**(1), 2011, 215–240.

[9] Guéant, O., Lasry, J. M. and Lions, P. L., Mean field games and applications, Paris-Princeton Lectures on Mathematical Sciences 2010, A. R. Carmona, et al. (eds.), 2011, 205–266.

[10] Huang, M., Caines, P. E. and Malhamé, R. P., Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria, *IEEE Trans. Automat. Control*, **52**(9), 2007, 1560–1571.

[11] Huang, M., Caines, P. E. and Malhamé, R. P., An invariance principle in large population stochastic dynamic games, *J. Syst. Sci. Comp.*, **20**(2), 2007, 162–172.

[12] Ladyzhenskaya, O. A. and Uraltseva, N. N., Linear and Quasilinear Elliptic Equations, Academic Press, New York, 1968.

[13] Lasry, J. M. and Lions, P. L., Jeux champ moyen I- Le cas stationnaire, *Comptes Rendus de l'Académie des Sciences, Series* I, **343**, 2006, 619–625.

[14] Lasry, J. M. and Lions, P. L., Jeux à champ moyen II- Horizn fini et contrôle optimal, *Comptes Rendus de l'Académie des Sciences, Series* I, **343**, 2006, 679–684.

[15] Lasry, J. M. and Lions, P. L., Mean field games, *Japanese J. Math.*, **2**(1), 2007, 229–260.