Chin. Ann. Math. Ser. B
39(3), 2018, 461–486
DOI: 10.1007/s11401-018-0078-4

Chinese Annals of Mathematics, Series B

Bellman Systems with Mean Field Dependent Dynamics*

Alain BENSOUSSAN[1] Miroslav BULÍČEK[2] Jens FREHSE[3]

(*Paper dedicated to Professor Philippe G. Ciarlet*)

Abstract The authors deal with nonlinear elliptic and parabolic systems that are the Bellman like systems associated to stochastic differential games with mean field dependent dynamics. The key novelty of the paper is that they allow heavily mean field dependent dynamics. This in particular leads to a system of PDE's with critical growth, for which it is rare to have an existence and/or regularity result. In the paper, they introduce a structural assumptions that cover many cases in stochastic differential games with mean field dependent dynamics for which they are able to establish the existence of a weak solution. In addition, the authors present here a completely new method for obtaining the maximum/minimum principles for systems with critical growths, which is a starting point for further existence and also qualitative analysis.

Keywords Stochastic games, Bellman equation, Mean field equation, Nonlinear elliptic equations, Weak solution, Maximum principle
2000 MR Subject Classification 35J60, 35K55, 35J55, 35B65

1 Introduction

In recent literature, mainly scalar Bellman equations which are coupled with a Fokker-Planck equation are studied and we refer to starting paper [9] or to a survey [6], see also [7]. These equations model a Nash game with a large number of players behaving similarly, so that the decision can be approximate by a single decision make (a representative agent). So we have a scalar Bellman equation. The present paper considers a model suggested by Bensoussan, accompanied by co-authors (see [3]), where the decision of finite number of N players of large population of Nash-game-players are approximated by N representative agents. So the Bellman system is a system of backward parabolic equations coupled with a forward mean field equation.

In principle, the dependence of the coefficients on the data in the nonlinearities of the equation may be a functional one. But in order to have the first insight in the difficulties and

Manuscript received October 28, 2017. Revised November 15, 2017.
[1]Jindal School of Management, International Center for Decision and Risk Analysis University of Texas at Dallas, 800 W. Campbell Rd, SM30 Richardson, TX 75080-3021, USA; Department SEEM City University of Hong Kong, Hung Hom, Kowloon, Hong Kong, China.
E-mail: axb046100@utdallas.edu
[2]Mathematical Institute, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic. E-mail: mbul8060@karlin.mff.cuni.cz
[3]Institute for Applied Mathematics, Department of Applied Analysis, Endenicher Allee 60, 53115 Bonn, Germany. E-mail: erdbeere@iam.uni-bonn.de
*This work was supported by the National Science Foundation (Nos. DMS 1303775, DMS 1612880),

in order to simplify the presentation, we confine to a point-wise dependence of Hamiltonians with respect to the mean field variable here. Generally, obtaining the existence of solution of the Bellman mean field dependent system with growth with respect to the mean field variable in nonlinearities is the critical subject. Without additional assumptions only a poor growth behaviour is permitted. For the case of scalar Bellman equations, for obtaining the global solvability, [10–11] gave a quite exhaustive analysis of the growth conditions for the Hamiltonians concerning the dependence on ∇u (u is a generalized value function) and the mean field m, which appears in the pay off functional.

In comparison, in the present paper, we restrict ourselves to Hamiltonians which grow quadratically with respect to ∇u, which is from the point of view of PDE analysis the most interesting case. In a related paper [1], a mean field dependence of the pay off functional is assumed, but no such a dependence of m in the dynamics of the system is considered. Also in [1], the growth properties of the data with respect to m, were crucial to obtain global solvability of the problem.

In this paper we go much beyond the scalar theory and the theory developed in [1] and obtain the existence result for much broader class of problems. As a key tool for the existence of a solution we use the method of sub and super solution used in the context of Bellman systems in [2].

2 Derivation of the System

In this paper we study a system of partial differential equations which arises as a necessary condition of a Nash-Point-problem for Vlasov-McKean-functionals

$$\mathcal{J}^i(\boldsymbol{v}) = \int_0^T \int_\Omega m(\boldsymbol{v}) f^i(\cdot, \boldsymbol{v}, m(\boldsymbol{v}))\mathrm{d}x\mathrm{d}t + \int_\Omega u_T^i m(T,\cdot)\,\mathrm{d}x, \tag{2.1}$$

where $(0,T)$ is a given time interval, $\Omega \subset \mathbb{R}^d$ is a cube $(0,1)^d$ and $Q := (0,T)\times\Omega$ is a space-time cylinder. The function $\boldsymbol{v} := (v^1, \cdots, v^N)$ with $N \in \mathbb{N}$ is the vector of the control functions, i.e.,

$$v^i : Q \to \mathbb{R}^M$$

is the control of the i-th player, where $i = 1, \cdots, N$ and $M \in \mathbb{N}$ is given. For every $i = 1, \cdots, N$ the function

$$f^i : Q \times \mathbb{R}^{MN} \times \mathbb{R} \to \mathbb{R}$$

is the so-called pay off function of the i-th player. Finally, the function

$$m : Q \to \mathbb{R}$$

is the so-called mean field variable. This means that for a given $\boldsymbol{v}$, it is a weak nonnegative solution of the following parabolic equation ("mean field equation"):

$$\partial_t m - \Delta m + \operatorname{div}(m\boldsymbol{g}(\cdot, \boldsymbol{v}, m)) = 0 \tag{2.2}$$

that is supposed to be satisfied in the space-time cylinder Q, fulfills spatially periodic boundary conditions (with respect to the unit cube Ω) and is completed by the initial data

$$m(0) = m_0 \geq 0 \quad \text{in } \Omega. \tag{2.3}$$

Here

$$\boldsymbol{g} : Q \times \mathbb{R}^{NM} \times \mathbb{R} \to \mathbb{R}^d$$

is a given mapping and we postpone the discussion about its structure to the end of the section.

For a bounded $\boldsymbol{v}$, under natural assumptions on the mapping $\boldsymbol{g}$, the problem (2.2)–(2.3) has a unique solution

$$m \in L^\infty(0,T; L^\infty(\Omega)) \cap L^2(0,T; W^{1,2}_{\text{per}}(\Omega)) \cap W^{1,2}(0,T; (W^{1,2}_{\text{per}}(\Omega))^*), \tag{2.4}$$

which allows us to define (2.1). Since the mean field variable depends on the choice of $\boldsymbol{v}$, we will frequently also write $m = m(\boldsymbol{v})$, which should not be understood as an algebraic relation, whenever the couple $(m, \boldsymbol{v})$ solves (2.2).

In (2.4) we use the standard notation for Bochner, Sobolev and Lebesgue spaces, and the subscript "per" indicates the periodicity with respect to Ω and this notation will be kept through the whole paper. In what follows, we also omit writing the dependence of function on (t, x) explicitly to shorten all formulae, i.e., we use the following abbreviation $\boldsymbol{f}(m, \boldsymbol{v})$ for $\boldsymbol{f}(t, x, m(t, x), \boldsymbol{v}(t, x))$ or for $\boldsymbol{f}(\cdot, m, \boldsymbol{v})$ in what follows, where $\boldsymbol{f} := (f^1, \cdots, f^N)$.

Having the mean field variable m and the control function $\boldsymbol{v}$, we can define a "pre-version" of the so-called Bellman system for a further function $\boldsymbol{u} = (u^1, \cdots, u^N) : Q \to \mathbb{R}^N$ via the backward parabolic system

$$\begin{aligned} &- \partial_t \boldsymbol{u} - \Delta \boldsymbol{u} = \boldsymbol{f}(\boldsymbol{v}, m) + m \boldsymbol{f}_m(\boldsymbol{v}, m), \\ &\nabla \boldsymbol{u}[\boldsymbol{g}(\boldsymbol{v}, m) + m \boldsymbol{g}_m(\boldsymbol{v}, m)] =: L(m, \boldsymbol{v}, \nabla \boldsymbol{u}), \end{aligned} \tag{2.5}$$

which is supposed to be satisfied in Q, equipped with the Ω-periodic boundary conditions and completed by the initial condition

$$\boldsymbol{u}(T) = \boldsymbol{u}_T \quad \text{in } \Omega.$$

We call this system the "pre-Bellman equation" since $\boldsymbol{v}$ has not been replaced by a feed back formula

$$\boldsymbol{v}(t, x) := \boldsymbol{\omega}(t, x, \nabla \boldsymbol{u}(t, x), m(t, x)) \tag{2.6}$$

yet and we introduce the meaning of (2.6) in the next subsection. Moreover, $L = (L^1, \cdots, L^N)$ are the so-called modified Lagrangians[1].

It is evident, that (2.2) and (2.5) do not form a closed problem and one needs to connect $\boldsymbol{v}$ with m and $\nabla \boldsymbol{u}$ via some relationship. It will be shown in the next subsection that the necessary condition for classical Nash–Point problem may serve as such constraint. In addition

[1]We use here the word modified since they differ from standard Lagrangians in Bellman systems, which is however here caused by the fact that $\boldsymbol{f}$ and $\boldsymbol{g}$ depend on m.

assuming further certain qualitative properties of $\boldsymbol{f}$, we will be able to give a good meaning to the feed back formula (2.6) and thus to avoid the presence of the control variable $\boldsymbol{v}$ in the analysis. The main goal of the paper is to introduce certain structural assumptions on $\boldsymbol{f}$ and $\boldsymbol{g}$ such that they describe very general mean field dependent Bellman system on one hand, and for which we can establish the existence of a weak solution on the other hand.

2.1 Derivation of the full system

Finally, we need to close the problem (2.2) and (2.5) by an algebraic condition which is necessary condition for the Nash-point of functionals J's. The classical Nash-point problem reads: For given $\boldsymbol{u}_T$, find $\boldsymbol{v} \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^{MN}))$ such that

$$J^i(\boldsymbol{v}) \le J^i(v_1,\cdots,v_{i-1},z,v_{i+1},\cdots,v_N) \tag{2.7}$$

for all $z \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^M))$ and corresponding m's, the solutions[2] to (2.2) with $\boldsymbol{v}$ replaced by

$$\boldsymbol{z} := (v_1,\cdots,v_{i-1},z,v_{i+1},\cdots,v_N).$$

The classical version treats the case where f^i and $\boldsymbol{g}$ do not depend on m. In that case, the problem (2.7) is purely analytical (meaning stochastic free) formulation of a stochastic differential game driven by the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x} = \boldsymbol{g}(t,\boldsymbol{x},\boldsymbol{v}).$$

In recent years, interest came up to study cases with m-dependence of the pay-off $\boldsymbol{f}$ and/or the dynamics $\boldsymbol{g}$. From PDE's point of view, this leads to new interesting version of the Bellman system.

Although, it is not known whether the problem (2.1) admits a Nash-point, we derive in what follows certain necessary conditions that must be fulfilled by the hypothetical Nash-point, which finally allow us to connect the pre-Bellman system (2.5) with the mean field equation (2.2) via the feed back formula (2.6) or its "equivalent".

Under natural assumptions on the data (see Subsection 2.2) and, say, $\boldsymbol{v} \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^{NM}))$, it is easy to see that the Gateaux derivatives of the J^i and of $m(\boldsymbol{v})$ exist. For

$$\boldsymbol{z}^i := (\underbrace{0,\cdots,0}_{i-1},z,\underbrace{0,\cdots,0}_{N-i})$$

with arbitrary smooth Ω-periodic function $z : Q \to \mathbb{R}^M$, we obtain that

$$M^i := \frac{\mathrm{d}}{\mathrm{d}s} m(\boldsymbol{v}+s\boldsymbol{z})\Big|_{s=0}$$

with $i = 1,\cdots,N$ satisfies

$$\partial_t M^i - \Delta M^i = -\operatorname{div}(M^i\boldsymbol{g}(\boldsymbol{v},m) + m\boldsymbol{g}_{v^i}(\boldsymbol{v},m)\cdot z + mM^i\boldsymbol{g}_m(\boldsymbol{v},m)) \quad \text{in } Q \tag{2.8}$$

and is completed by the initial condition

$$M^i(0,x) = 0 \quad \text{a.e. in } \Omega.$$

[2]We may define this also for other $L^p(L^q)$ spaces once the uniqueness of $m(\boldsymbol{v})$ is guaranteed.

Here, we use the subscript to abbreviate the notion of partial derivative, i.e., $\boldsymbol{g}_{\boldsymbol{v}}(\boldsymbol{v},m) := \partial_{\boldsymbol{v}}\boldsymbol{g}(\boldsymbol{v},m)$ and $\boldsymbol{g}_m(\boldsymbol{v},m) := \partial_m \boldsymbol{g}(\boldsymbol{v},m)$. Furthermore, assuming that $\boldsymbol{v}$ is the Nash-equilibrium, we have for all $i = 1, \cdots, N$ that

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}s} J^i(\boldsymbol{v} + s\boldsymbol{z}^i)\Big|_{s=0} \\ &= \int_Q (m f^i_{v^i}(\boldsymbol{v},m)\cdot z + M^i f^i(\boldsymbol{v},m) + mM^i f^i_m(\boldsymbol{v},m))\mathrm{d}x\mathrm{d}t \\ &\quad + \int_\Omega M^i(T)u^i(T)\,\mathrm{d}x \end{aligned} \tag{2.9}$$

for arbitrary smooth Ω-periodic function z. Notice that here $m := m(\boldsymbol{v})$, i.e., m solves (2.2) with $\boldsymbol{v}$. To evaluate the terms not involving explicitly z, we use the equations (2.5) and (2.8). First, multiplying (2.8) by u^i, integrating over Q and using integration by parts (note that $M^i(0) = 0$), we deduce

$$\begin{aligned} &\int_\Omega M^i(T)u^i(T)\,\mathrm{d}x + \int_Q \nabla M^i\cdot\nabla u^i \mathrm{d}x\mathrm{d}t - \int_Q M^i\partial_t u^i\mathrm{d}x\mathrm{d}t \\ &= \int_Q (M^i\boldsymbol{g}(\boldsymbol{v},m) + m\boldsymbol{g}_{v^i}(\boldsymbol{v},m)\cdot z + mM\boldsymbol{g}_m(\boldsymbol{v},m))\cdot\nabla u^i\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{2.10}$$

Next, multiplying the i-th equation in (2.5) by M^i, we observe

$$\begin{aligned} &-\int_Q \partial_t u^i M^i \mathrm{d}x\mathrm{d}t + \int_Q \nabla u^i\cdot\nabla M^i\mathrm{d}x\mathrm{d}t \\ &= \int_Q M^i(f^i(\boldsymbol{v},m) + mf^i_m(\boldsymbol{v},m))\mathrm{d}x\mathrm{d}t \\ &\quad + \int_Q M^i(\boldsymbol{g}(\boldsymbol{v},m) + m\boldsymbol{g}_m(\boldsymbol{v},m))\cdot\nabla u^i\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{2.11}$$

Finally, subtracting (2.11) from (2.10), we obtain the following identity:

$$\begin{aligned} \int_\Omega M^i(T)u^i(T)\,\mathrm{d}x &= -\int_Q M^i(f^i(\boldsymbol{v},m) + mf^i_m(\boldsymbol{v},m))\mathrm{d}x\mathrm{d}t \\ &\quad + \int_Q m\boldsymbol{g}_{v^i}(\boldsymbol{v},m)\cdot z\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{2.12}$$

Thus, using this relation in the necessary condition (2.9), we see that

$$\int_Q f^i_{v^i}(\boldsymbol{v},m)\cdot z + (\nabla u^i \otimes z)\cdot\boldsymbol{g}_{v^i}(\boldsymbol{v},m)\mathrm{d}x\mathrm{d}t = 0$$

for all $i = 1, \cdots, N$ and all smooth Ω-periodic z's. This consequently leads to the necessary compatibility condition

$$f^i_{v^i}(\boldsymbol{v},m) + \nabla u^i\cdot\boldsymbol{g}_{v^i}(\boldsymbol{v},m) = 0 \quad \text{in } Q. \tag{2.13}$$

Thus, now we have a closed system of equations. Namely, (2.2), (2.5) and (2.13) forms a well-defined problem for which we want to establish our analytical result. Indeed, the first one will deal just with (2.2), (2.5) and (2.13) and leads to the uniform a priori estimate for $(m, \boldsymbol{v}, \boldsymbol{u})$.

However, to obtain also the existence result, we shall require that for a given $(m, \nabla \boldsymbol{u})$, we can find a unique $\boldsymbol{v}$ solving (2.13). For such a solution we define the feed back formula (2.6) as

$$\boldsymbol{\omega}(m, \nabla \boldsymbol{u}) := \boldsymbol{v}$$

and replacing $\boldsymbol{v}$ in (2.5) by the feed back formula, we obtain

$$-\partial_t \boldsymbol{u} - \Delta \boldsymbol{u} = H(\nabla \boldsymbol{u}, m) := L(m, \boldsymbol{\omega}(m, \nabla \boldsymbol{u}), \nabla \boldsymbol{u}). \tag{2.14}$$

Similarly, we replace $\boldsymbol{v}$

We also prescribe the behaviour of $\boldsymbol{f}$ at 0, i.e., for all $i = 1, \cdots, N$, we assume that

$$f^i(m, \boldsymbol{v})|_{v^i=0} \le K(1 + m^{2s_0} + |\boldsymbol{v}|^{\alpha+1}) \tag{2.20}$$

and finally the coerciveness of $f^i_{v^i}$, i.e., we assume that

$$C_0(m^r + 1)|v^i|^2 \le f^i_{v^i}(m, \boldsymbol{v}) \cdot v^i + K(1 + m^{2s_0} + |\boldsymbol{v}|^{\alpha+1}). \tag{2.21}$$

Next, we focus on the assumptions on $\boldsymbol{g}$. The standard ones are related to the growth estimates, which will be supposed to be given by

$$\begin{aligned} &m|\boldsymbol{g}_m(m, \boldsymbol{v})| + |\boldsymbol{g}(m, \boldsymbol{v})| \le K((m^s + 1)|\boldsymbol{v}| + m^{s_0} + 1), \\ &|\boldsymbol{g}_{\boldsymbol{v}}(m, \boldsymbol{v})| \le K(m^s + 1). \end{aligned} \tag{2.22}$$

The forthcoming structural assumptions on $\boldsymbol{g}$ and also on $\boldsymbol{f}_{\boldsymbol{v}}$ are in fact the key restrictions of the paper. First, to simplify the further analysis, we shall assume that[3]

$$\boldsymbol{g} = \sum_{j=1}^{N} b_1(m) A^j(\cdot) v^j + \boldsymbol{b}_0(m), \tag{2.23}$$

where $\boldsymbol{b}_0$ does not depend on (t, x). The matrices $A^j : Q \to \mathbb{R}^{d\times M}$ are given functions of (t, x), i.e., $(A^j)_{ik} := A^j_{ik}$ with $i = 1, \cdots, d$ and $k = 1, \cdots, M$ and the meaning of $A^j v^j$ is

$$(A^j v^j)_i := \sum_{k=1}^{M} A^j_{ik} v^j_k.$$

Hence, as it is assumed we have a d-dimensional vector function $\boldsymbol{g} = (g_1, \cdots, g_d)$. The inhomogeneities $\boldsymbol{b}_0 : \mathbb{R} \to \mathbb{R}^d$ and $b_1 : \mathbb{R} \to \mathbb{R}$ are given. Concerning the assumptions on functions b_1 and $\boldsymbol{b}_0$, we require that

$$|b_1(m)| \le K(m^s + 1), \quad |\boldsymbol{b}_0(m)| \le K(m^{s_0} + 1) \tag{2.24}$$

and for matrices A^j we assume that

$$|A^j(\cdot)| \le K. \tag{2.25}$$

Furthermore, one of the essentially required properties of A^j is that they have the same range, i.e., we assume that for all $i, j = 1, \cdots, N$, almost all (t, x) and all $\boldsymbol{z} \in \mathbb{R}^{Nd}$ there holds

$$|\boldsymbol{z}A^j(t, x)| \le C_1|\boldsymbol{z}A^i(t, x)|. \tag{2.26}$$

For derivatives of $\boldsymbol{b}_0$, we need that

$$\begin{aligned} &|m\partial_m \boldsymbol{b}_0(m)| \le Km(m + 1)^{s_0-1}, \\ &|m^2\partial_{mm}\boldsymbol{b}_0(m)| \le Km^2(m + 1)^{s_0-2}. \end{aligned} \tag{2.27}$$

Finally, for the derivative of b_1 with respect to m, we introduce a certain "smallness assumption": There exists $\delta \in [0, 1)$ such that for all $b \in \mathbb{R}_+$ and all $\boldsymbol{v} \in \mathbb{R}^{NM}$ there holds

$$C_1\sqrt{N}\frac{|m\partial_m b_1(m)|}{|b_1(m)|}|\boldsymbol{v}|\sum_{i=1}^{M}|f^i_{\boldsymbol{v}^i}| \le \sum_{i=1}^{N}(f^i(m, \boldsymbol{v}) + m\partial_m f^i(m, \boldsymbol{v})) + K(1 + m^{2s_0}). \tag{2.28}$$

[3]The linearity of $\boldsymbol{g}$ with respect to $\boldsymbol{v}$ is in fact not a necessary assumption. A more delicate here is just behaviour of $\boldsymbol{g}$ with respect to m.

Although the above assumption seems to be complicated, it naturally appears in the first a priori estimate of the problem. In the next subsection, we shall show that in many cases, the assumption (2.28) still allows very general behaviour of all quantities with respect to m, namely, the case when one can dominate the function $\boldsymbol{g}$ in certain sense. In addition, since (2.28) may seem to be complicated, we can replace it by

$$|m\partial_m b_1(m)| \leq \gamma |b_1(m)| \quad \text{with } \gamma \leq \frac{C_0}{2(C_1^2+N^2)}, \tag{2.29}$$

since then (2.28) follows easily from the previous assumptions.

Please, observe here that if $\boldsymbol{g}$ is given by (2.23), then (2.13) reduces to

$$f^i_{v^i_j}(\boldsymbol{v},m) + b_1(m)\sum_{k=1}^d \partial_{x_k} u^i A^i_{kj} = 0 \quad \text{in } Q, \tag{2.30}$$

which must be valid for all $i=1,\cdots,N$ and all $j=1,\cdots,M$.

The above assumptions are sufficient for establishing formal a priori estimates. However, for getting also the existence of a weak solution, we need to give a well meaning to the feed back formula, or in other words, we need to guarantee the unique solvability of (2.30) for given m and ∇u. There can be introduced many assumptions that would lead to such a goal but we follow here the most standard one, which is the monotone operator approach. For a given m, we define the mapping $T:\mathbb{R}^{NM}\to\mathbb{R}^{NM}$ by

$$T(\boldsymbol{v}) = \Big(\frac{\partial}{\partial v^1} f^1(\boldsymbol{v},m),\cdots,\frac{\partial}{\partial v^N} f^N(\boldsymbol{v},m)\Big)$$

and assume that it is continuous with respect to $(m,\boldsymbol{v})$ and measurable with respect to (t,x) and the strictly monotone, i.e., for all $\boldsymbol{v}\neq\widetilde{\boldsymbol{v}}$,

$$(T(\boldsymbol{v})-T(\widetilde{\boldsymbol{v}}))\cdot(\boldsymbol{v}-\widetilde{\boldsymbol{v}}) > 0. \tag{2.31}$$

Then using also the assumption (2.21), we see that it satisfies

$$\frac{T(\nu)\cdot\nu}{|\nu|} \to \infty \quad \text{as } |\nu|\to\infty, \tag{2.32}$$

and from the standard monotone operator theory, we can conclude the existence of a unique $\boldsymbol{v}$ solving (2.30) and also consequently, we obtain that the feed back law (2.6) is well defined.

2.3 Prototypical example

Our prototype example is the following. For f, we assume a structure

$$f^i(m,\boldsymbol{v}) := (m+1)^r|v^i|^2 + B^i\cdot\boldsymbol{v} + K(1+m^{2s_0}), \tag{2.33}$$

where B^i's are arbitrary bounded measurable matrices. Next, for b_1 we consider

$$b_1(m) := (m+1)^s, \tag{2.34}$$

and for $\boldsymbol{b}_0$ and matrices A^i, we just require (2.25)–(2.27). With such a choice, all assumptions (2.16)–(2.27) and also (2.31) are satisfied and we shall just to show what is the meaning of (2.28). A very direct computation leads to the necessary condition

$$sm \leq \frac{\gamma}{2N}(1+(r+1)m),$$

which is surely satisfied whenever

$$s < \frac{r}{2N}. \tag{2.35}$$

2.4 Statement of main results

The main result of the paper is twofold. First, we give the uniform a priori estimate result which holds for all sufficiently smooth solutions provided that parameters satisfy the assumptions stated above.

Theorem 2.1 *Let $\boldsymbol{g}$ and $\boldsymbol{f}$ satisfy* (2.16)–(2.28)*. Then any sufficiently regular solution* $(m, \boldsymbol{v}, \boldsymbol{u})$ *to* (2.2)*,* (2.5) *and* (2.13) *satisfies the following estimate:*

$$\begin{aligned}\sup_{t\in(0,T)} (\|m(t)\|_\sigma + \|\boldsymbol{u}(t)\|_\infty) + \int_Q |\nabla \boldsymbol{u}|^2 + (m+1)^{\sigma-2}|\nabla m|^2 \mathrm{d}x\mathrm{d}t \\ + \int_Q m^{2s_0+1} + (m+1)(m^r+1)|\boldsymbol{v}|^2 \mathrm{d}x\mathrm{d}t \le C(\|\boldsymbol{u}_T\|_\infty, \|m_0\|_\sigma),\end{aligned} \tag{2.36}$$

where

$$\sigma := r - 2s + 1 \tag{2.37}$$

provided that

$$0 \le \min\left\{\frac{4(2s_0(d+2)-1)_+}{\sigma(d+2)-d-(2s_0-\sigma+1)_+(d+2)}, \frac{2s_0(d+2)}{\sigma(d+2)-d}\right\} < 1 \tag{2.38}$$

and

$$r \ge 2s. \tag{2.39}$$

Next, we state the second main theorem of the paper, which is the existence result.

Theorem 2.2 *Let $\boldsymbol{g}$ and $\boldsymbol{f}$ satisfy* (2.16)–(2.28) *and* (2.31)*. Assume that* (2.38)–(2.39) *is fulfilled. Then for arbitrary $\boldsymbol{u}_T \in L^\infty$ and nonnegative $m_0 \in L^\sigma(\Omega)$ with σ fulfilling* (2.37) *there exists a weak solution satisfying the estimate* (2.36) *and fulfilling or almost all $t \in (0, T)$,*

$$\langle \partial_t m, \varphi\rangle + \int_\Omega \nabla m \cdot \nabla\varphi - m\boldsymbol{g}(\boldsymbol{v}, m)\cdot\nabla\varphi = 0, \tag{2.40}$$

$$\begin{aligned}&- \langle \partial_t \boldsymbol{u}, \boldsymbol{z}\rangle + \int_\Omega \nabla\boldsymbol{u}\cdot\nabla\boldsymbol{z}\,\mathrm{d}x \\ &= \int_\Omega (\boldsymbol{f}(\boldsymbol{v}, m) + m\boldsymbol{f}_m(\boldsymbol{v}, m))\cdot\boldsymbol{z}\,\mathrm{d}x + \int_\Omega \nabla\boldsymbol{u}[\boldsymbol{g}(\boldsymbol{v}, m) + m\boldsymbol{g}_m(\boldsymbol{v}, m)]\cdot\boldsymbol{z}\,\mathrm{d}x,\end{aligned} \tag{2.41}$$

and completed by the relationship between $\boldsymbol{v}$ and $(m, \nabla\boldsymbol{u})$,

$$f^i_{v^i_j}(\boldsymbol{v}, m) + b_1(m)\sum_{k=1}^d \partial_{x_k} u^i A^i_{kj} = 0 \quad \text{in } Q. \tag{2.42}$$

To illustrate the power of the result, we just consider our prototypical example. First in case that $s_0 = 0$, we see that the only restriction is

$$2Ns < r.$$

In the opposite extreme case, i.e., if

$$r = s = 0,$$

then

$$s_0 < \frac{1}{2(d+2)}.$$

3 Algebraic Estimates for Lagrangians and Hamiltonians

In this section, we derive basic algebraic inequalities that are satisfied for Lagrangians and consequently also for Hamiltonians provided that $\boldsymbol{f}$ and $\boldsymbol{g}$ satisfy the assumption introduced in Subsection 2.2, namely the assumptions (2.16)–(2.28). The key observation is that under these assumptions, the Lagrangians satisfy the lower sum coerciveness and the proper upper estimates. It will be also evident from the estimates below why we require $2s \le r$ in main results of the paper.

Lemma 3.1 *Let $\boldsymbol{f}$ and $\boldsymbol{g}$ satisfy* (2.16)–(2.28). *Then there exists a constant $C > 0$ and an $\varepsilon_0 \in \big(0, \frac{1}{2N}\big)$ such that for almost all (t, x), all $(m, \boldsymbol{v}, \nabla \boldsymbol{u})$ fulfilling* (2.13) *and all $i = 1, \cdots, N$ there hold:*

(i) *Sum coerciveness*

$$\sum_{i=1}^N L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) \ge \frac{C_0}{2}(m^r+1)|\boldsymbol{v}|^2 - C(m^{2s_0}+1) - C\Big|\nabla \sum_{i=1}^N u^i\Big|^2 \Big(1 + \frac{m^{2s}+1}{m^r+1}\Big). \tag{3.1}$$

(ii) *Upper bound*

$$\begin{aligned} &L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) - \varepsilon_0 \sum_{j=1}^N L^j(m, \boldsymbol{v}, \nabla \boldsymbol{u}) \\ &\le C(1+m^{2s_0}) + C\Big|\nabla\Big(u^i - \varepsilon_0 \sum_{i=j}^N u^j\Big)\Big|^2 \Big(1 + \frac{m^{2s}+1}{m^r+1}\Big). \end{aligned} \tag{3.2}$$

(iii) *Global bound*

$$|L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u})| \le C\Big(1 + m^{2s_0} + (1+m^r)|\boldsymbol{v}|^2 + |\nabla \boldsymbol{u}|^2 \Big(1 + \frac{m^{2s}+1}{m^r+1}\Big)\Big). \tag{3.3}$$

Proof We start with the proof of (3.1). Using the definition of L in (2.5) we get the identity

$$\sum_{i=1}^N L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) = \sum_{i=1}^N (f^i(\boldsymbol{v}, m) + m f^i_m(\boldsymbol{v}, m)) + \sum_{i=1}^N \nabla u^i \cdot (\boldsymbol{g}(\boldsymbol{v}, m) + m \boldsymbol{g}_m(\boldsymbol{v}, m)).$$

Hence, using (2.18) for the first part and (2.22) for the second part, we observe that

$$\begin{aligned} \sum_{i=1}^N L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) &\ge C_0(m^r+1)|\boldsymbol{v}|^2 - K(m^{2s_0}+1) - K\Big|\nabla \sum_{i=1}^N u^i\Big|((m^s+1)|\boldsymbol{v}| + m^{s_0} + 1) \\ &\ge \frac{C_0}{2}(m^r+1)|\boldsymbol{v}|^2 - C(m^{2s_0}+1) - C\Big|\nabla \sum_{i=1}^N u^i\Big|^2 \Big(1 + \frac{m^{2s}+1}{m^r+1}\Big), \end{aligned}$$

where for the second estimate we used the Young inequality. This finishes the proof of (3.1).

Next, we look for (3.2). Using the definition of Lagrangians, we directly obtain for arbitrary $\varepsilon > 0$ the following identity:

$$\begin{aligned}\mathrm{I} &:= L^i(m,\boldsymbol{v},\nabla\boldsymbol{u}) - \varepsilon\Big(\sum_{i=1}^N L^i(m,\boldsymbol{v},\nabla\boldsymbol{u})\Big)\\ &= f^i(m,\boldsymbol{v}) + m f^i_m(m,\boldsymbol{v}) - \varepsilon\Big(\sum_{j=1}^N f^j(m,\boldsymbol{v}) + m f^j_m(m,\boldsymbol{v})\Big)\\ &\quad + \nabla\Big(u^i - \varepsilon\sum_{j=1}^N u^j\Big)\cdot(\boldsymbol{g}(m,\boldsymbol{v}) + m\boldsymbol{g}_m(m,\boldsymbol{v})).\end{aligned}$$

Next, using (2.17)–(2.18), (2.22) and the Young inequality, we have

$$\begin{aligned}\mathrm{I} &\le C(1+m^{2s_0}+|\boldsymbol{v}|^{\alpha+1}) + Kf^i(m,\boldsymbol{v}) - \varepsilon C_0(m^r+1)|\boldsymbol{v}|^2\\ &\quad + K\Big|\nabla\Big(u^i - \varepsilon\sum_{i=1}^N u^i\Big)\Big|(1+m^{s_0}+(m^s+1)|\boldsymbol{v}|)\\ &\le C(\varepsilon)(1+m^{2s_0}) + Kf^i(m,\boldsymbol{v}) - \frac{\varepsilon C_0}{2}(m^r+1)|\boldsymbol{v}|^2\\ &\quad + C(\varepsilon)\Big(1+\frac{m^{2s}+1}{m^r+1}\Big)\Big|\nabla\Big(u^i - \varepsilon\sum_{i=1}^N u^i\Big)\Big|^2.\end{aligned} \tag{3.4}$$

It remains to estimate the term with f^i. We use the convexity of f^i with respect to v^i and the assumption (2.20). Then to evaluate $f^i_{v^i}$ we also use the constraint (2.13), which however in our case reduces to (2.30) due to the structural assumptions (2.23). Doing so, we get

$$\begin{aligned}&Kf^i(m,\boldsymbol{v}) \le Kf^i(m,\boldsymbol{v})|_{v^i=0} + Kf^i_{v^i}(m,\boldsymbol{v})\cdot v^i\\ &\le C(1+m^{2s_0}+|\boldsymbol{v}|^{\alpha+1}) - Cb_1(m)\sum_{k=1}^d\sum_{j=1}^N \partial_{x_k}u^iA^i_{kj}v^i_j\\ &= C(1+m^{2s_0}+|\boldsymbol{v}|^{\alpha+1}) - Cb_1(m)\sum_{k=1}^d\sum_{j=1}^N \partial_{x_k}\Big(u^i - \varepsilon\sum_{\ell=1}^N u^\ell\Big)A^i_{kj}v^i_j\\ &\quad - \varepsilon Cb_1(m)\sum_{k=1}^d\sum_{j=1}^N\sum_{\ell=1}^N \partial_{x_k}u^\ell A^i_{kj}v^i_j.\end{aligned}$$

Then we estimate terms involving $A^\ell\nabla u^i$ with the help of (2.24), (2.26) and (2.30) as follows

$$\begin{aligned}Kf^i(m,\boldsymbol{v}) &\le C(1+m^{2s_0}+|\boldsymbol{v}|^{\alpha+1}) + C(m^s+1)\Big|\nabla\Big(u^i - \varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big||\boldsymbol{v}|\\ &\quad + \varepsilon C|v^i|\sum_{\ell=1}^N|\nabla u^\ell A^\ell||b_1(m)|\\ &\le C(1+m^{2s_0}+|\boldsymbol{v}|^{\alpha+1}) + C(m^s+1)\Big|\nabla\Big(u^i - \varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big||\boldsymbol{v}|\end{aligned}$$

$$+ \varepsilon C|v^i||\boldsymbol{f}_{\boldsymbol{v}}(m,\boldsymbol{v})|. \tag{3.5}$$

Finally, we focus on estimate of v^i. It follows from (2.21), (2.24), (2.26) and (2.30) that

$$\begin{aligned}
C_0(m^r+1)|v^i|^2 &\le f^i_{v^i}\cdot v^i + K(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha})\\
&= -b_1(m)\nabla u^i A^i\cdot v^i + K(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha})\\
&= -b_1(m)\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)A^i\cdot v^i\\
&\quad -\varepsilon b_1(m)\sum_{\ell=1}^N \nabla u^\ell A^i\cdot v^i + K(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha})\\
&\le C(m^s+1)\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big||v^i|\\
&\quad +\varepsilon C|v^i||b_1(m)|\sum_{\ell=1}^N|\nabla u^\ell A^\ell| + K(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha})\\
&\le C\frac{m^{2s}+1}{m^r+1}\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big|^2 + C_0\frac{m^r+1}{2}|v^i|^2\\
&\quad +\varepsilon C|\boldsymbol{v}||\boldsymbol{f}_{\boldsymbol{v}}(m,\boldsymbol{v})| + K(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha}),
\end{aligned}$$

where for the last inequality, we used the Young inequality and the structural constraint (2.30). Hence absorbing the corresponding term to the left-hand side, using (2.16) and the Young inequality, we get

$$\begin{aligned}
(m^r+1)|v^i|^2 &\le C\frac{(m^{2s}+1)}{m^r+1}\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big|^2\\
&\quad +\varepsilon C(m^r+1)|\boldsymbol{v}|^2 + C(1+m^{2s_0}+|\boldsymbol{v}|^{1+\alpha}).
\end{aligned} \tag{3.6}$$

Next, using the Young inequality in (3.5), we find

$$\begin{aligned}
Kf^i(m,\boldsymbol{v}) &\le C(\varepsilon)(1+m^{2s_0}) + C\frac{m^{2s}+1}{\varepsilon^2(m^r+1)}\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big|^2\\
&\quad + C\varepsilon^2(m^r+1)|\boldsymbol{v}|^2 + \varepsilon C|v^i||\boldsymbol{f}_{\boldsymbol{v}}(m,\boldsymbol{v})|\\
&\le C(\varepsilon)(1+m^{2s_0}) + C\frac{m^{2s}+1}{\varepsilon^2(m^r+1)}\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big|^2\\
&\quad + C\varepsilon^2(m^r+1)|\boldsymbol{v}|^2 + \varepsilon^{\frac12}(m^r+1)|v^i|^2 + C\varepsilon^{\frac32}\frac{|\boldsymbol{f}_{\boldsymbol{v}}(m,\boldsymbol{v})|^2}{m^r+1}.
\end{aligned} \tag{3.7}$$

Finally, we substitute (3.6) into (3.7) to estimate term with v^i and with the help of the assumption (2.16), we obtain

$$Kf^i(m,\boldsymbol{v}) \le C(\varepsilon)\Big(1+m^{2s_0}+\frac{m^{2s}+1}{m^r+1}\Big|\nabla\Big(u^i-\varepsilon\sum_{\ell=1}^N u^\ell\Big)\Big|^2\Big) + C\varepsilon^{\frac32}(m^r+1)|\boldsymbol{v}|^2. \tag{3.8}$$

Consequently, (3.8) combined with (3.4) directly implies

$$\mathrm{I} \le C(\varepsilon)(1+m^{2s_0}) - \frac{C_0}{2}(\varepsilon - C\varepsilon^{\frac{3}{2}})(m^r+1)|\boldsymbol{v}|^2 + C(\varepsilon)\frac{m^{2s}+1}{m^r+1}\Big|\nabla\Big(u^i - \varepsilon\sum_{i=1}^N u^i\Big)\Big|^2. \tag{3.9}$$

Thus, if we choose $\varepsilon =: \varepsilon_0 \in \big(0, \frac{2}{N}\big)$ such that

$$\varepsilon_0 - C\varepsilon_0^{\frac{3}{2}} \ge 0,$$

which is always possible and the maximal value of such ε_0 depends only on the generic constant C, we obtain the desired estimate (3.2).

The estimate (3.3) is then a simple combination of (3.1)–(3.2). The proof is finished.

4 Uniform a Priori Estimates—Proof of Theorem 2.1

This section is devoted to the estimates for solution $(m, \boldsymbol{u}, \boldsymbol{v})$ of (2.2), (2.5) and (2.13) that depend only on data of the problem provided that the assumptions (2.16)–(2.28) are satisfied. Here, we proceed rather formally, considering that the solution is sufficiently regular. The justification of such a procedure then will be provided in the proof of the existence result, i.e., in the proof of Theorem 2.2.

4.1 Estimates for m

We start with estimates for m. First, if $m_0 \ge 0$ almost everywhere in Ω, then the standard minimum principle for parabolic equations implies that (sufficiently smooth) solution satisfies $m \ge 0$ almost everywhere in Q as well. Next, setting $\varphi := 1$ in (2.40), we get with the help of nonnegativity of m that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|m(t)\|_1 = 0.$$

Consequently, we have

$$\sup_{t\in(0,T)} \|m(t)\|_1 \le \|m_0\|_1 \le C. \tag{4.1}$$

Next, in order to obtain estimates on $(\boldsymbol{u}, \boldsymbol{v})$ we need to improve the information about m since the Lagrangians (or Hamiltonians) depend heavily on m. Our goal is to prove the starting point inequality

$$\begin{aligned} E &:= \sup_{t\in(0,T)} \|m(t)+1\|_\sigma^\sigma + \int_Q (m+1)^{\sigma-2}|\nabla m|^2 \mathrm{d}x\mathrm{d}t \\ &\le C\int_Q (m+1)^{2s_0+1}\mathrm{d}x\mathrm{d}t + C\|\boldsymbol{u}\|^2_{L^\infty(Q)}\Big(1 + \int_Q (m+1)^{2s_0-\sigma+2}\mathrm{d}x\mathrm{d}t\Big), \end{aligned} \tag{4.2}$$

where the constant C depends only on data given by assumptions on $\boldsymbol{f}$ and $\boldsymbol{g}$.

Proof of (4.2) We first set $\varphi := (m+1)^{\sigma-1}$ in (2.40), where σ is given by (2.37), i.e.,

$$\sigma := r + 1 - 2s \ge 1.$$

With such a choice of φ, we get the identity

$$\frac{1}{\sigma}\frac{\mathrm{d}}{\mathrm{d}t}\|m(t)+1\|_\sigma^\sigma + (\sigma-1)\int_\Omega (m+1)^{\sigma-2}|\nabla m|^2 \,\mathrm{d}x$$

$$
\begin{aligned}
&= (\sigma-1)\int_\Omega m(m+1)^{\sigma-2}\boldsymbol{g}(m,\boldsymbol{v})\cdot\nabla m \,\mathrm{d}x\\
&= (\sigma-1)\int_\Omega m(m+1)^{\sigma-2}b_1(m)\sum_{j=1}^M A^j v^j\cdot\nabla m \,\mathrm{d}x\\
&\quad + (\sigma-1)\int_\Omega m(m+1)^{\sigma-2}\boldsymbol{b}_0(m)\cdot\nabla m \,\mathrm{d}x,
\end{aligned}
$$

where the second equality follows from the structural assumption on $\boldsymbol{g}$ (see (2.23)). Clearly, the second term on the right-hand side vanishes due to the integration by parts and spatially periodic boundary conditions. For the first integral we use the Young inequality to conclude

$$
\begin{aligned}
&\frac{1}{\sigma}\frac{\mathrm{d}}{\mathrm{d}t}\|m(t)+1\|_\sigma^\sigma + \frac{(\sigma-1)}{2}\int_\Omega (m+1)^{\sigma-2}|\nabla m|^2 \,\mathrm{d}x\\
&\le \frac{(\sigma-1)}{2}\int_\Omega m^2(m+1)^{\sigma-2}b_1^2(m)\Big|\sum_{j=1}^N A^j v^j\Big|^2 \,\mathrm{d}x.
\end{aligned}
$$

Thus, using the bounds for b_1 and A^j, namely (2.24)–(2.25), we obtain after integration over $(0,T)$

$$
\begin{aligned}
&\sup_{t\in(0,T)}\|m(t)+1\|_\sigma^\sigma + \int_Q (m+1)^{\sigma-2}|\nabla m|^2 \mathrm{d}x\mathrm{d}t\\
&\le C(\sigma,\|m_0\|_q,\Omega)\Big(1+\int_Q m(m+1)^{2s+\sigma-1}|\boldsymbol{v}|^2\mathrm{d}x\mathrm{d}t\Big).
\end{aligned}
\tag{4.3}
$$

Next, we set $\varphi := u^i$ in (2.40) and $z := m$ in the i-th equation in (2.41), integrate the result with respect to time and use integration by parts to obtain the following two identities:

$$
\begin{aligned}
&\int_\Omega m(T)u^i(T) - m(0)u^i(0)\,\mathrm{d}x - \int_Q m\partial_t u^i \mathrm{d}x\mathrm{d}t\\
&+\int_Q \nabla m\cdot\nabla u^i - m\boldsymbol{g}(\boldsymbol{v},m)\cdot\nabla u^i \mathrm{d}x\mathrm{d}t = 0
\end{aligned}
\tag{4.4}
$$

and

$$
\begin{aligned}
&-\int_Q \partial_t u^i m \mathrm{d}x\mathrm{d}t + \int_Q \nabla u^i\cdot\nabla m \mathrm{d}x\mathrm{d}t\\
&=\int_Q m(f^i(\boldsymbol{v},m) + m f^i_m(\boldsymbol{v},m))\mathrm{d}x\mathrm{d}t\\
&\quad+\int_Q m\nabla u^i\cdot\boldsymbol{g}(\boldsymbol{v},m) + m^2\boldsymbol{g}_m(\boldsymbol{v},m)\cdot\nabla u^i\mathrm{d}x\mathrm{d}t.
\end{aligned}
\tag{4.5}
$$

Subtracting (4.4) from (4.5) we arrive at identity

$$
\begin{aligned}
&\int_Q m(f^i(\boldsymbol{v},m) + m f^i_m(\boldsymbol{v},m))\mathrm{d}x\mathrm{d}t\\
&= -\int_\Omega m(T)u^i(T) - m(0)u^i(0)\,\mathrm{d}x - \int_Q m^2\boldsymbol{g}_m(\boldsymbol{v},m)\cdot\nabla u^i\\
&\le C\|\boldsymbol{u}\|_{L^\infty(Q)} - \int_Q m^2\boldsymbol{g}_m(\boldsymbol{v},m)\cdot\nabla u^i\mathrm{d}x\mathrm{d}t,
\end{aligned}
\tag{4.6}
$$

where for the last inequality we used the Hölder inequality and the uniform bound (4.1). Next, we evaluate the last integral. Using the structural assumption (2.23) and the growth assumptions (2.24)–(2.27), we see that

$$\begin{aligned}
&-\int_Q m^2 \boldsymbol{g}_m(\boldsymbol{v}, m)\cdot \nabla u^i \mathrm{d}x\mathrm{d}t \\
&= -\int_Q m^2 \partial_m \boldsymbol{b}_0(m)\cdot \nabla u^i \mathrm{d}x\mathrm{d}t - \sum_{j=1}^N \int_Q m^2 \partial_m b_1(m) A^j v^j \cdot \nabla u^i \mathrm{d}x\mathrm{d}t \\
&= \int_Q u^i \partial_m (m^2 \partial_m \boldsymbol{b}_0(m))\cdot \nabla m \mathrm{d}x\mathrm{d}t \\
&\quad - \sum_{j=1}^N \int_Q m^2 \partial_m b_1(m) A^j v^j \cdot \nabla u^i \mathrm{d}x\mathrm{d}t \\
&\le \int_Q K|\boldsymbol{u}| m (m+1)^{s_0-1} |\nabla m| \mathrm{d}x\mathrm{d}t \\
&\quad + \int_Q m^2 |\partial_m b_1(m)| |\boldsymbol{v}| \Big(\sum_{j=1}^N |\nabla u^i A^j|^2 \Big)^{\frac{1}{2}} \mathrm{d}x\mathrm{d}t \\
&=: \mathrm{I}_1 + \mathrm{I}_2.
\end{aligned}$$

To estimate I_2 we use (2.26) and the constraint (2.30) to get

$$\begin{aligned}
\mathrm{I}_2 &\le C_1 \int_Q \frac{m^2 |\partial_m b_1(m)|}{b_1(m)} |\boldsymbol{v}| \Big(\sum_{j=1}^N |b_1(m) \nabla u^i A^i|^2 \Big)^{\frac{1}{2}} \mathrm{d}x\mathrm{d}t \\
&= C_1 \sqrt{N} \int_Q \frac{m^2 |\partial_m b_1(m)|}{b_1(m)} |f^i_{v^i}| |\boldsymbol{v}| \mathrm{d}x\mathrm{d}t.
\end{aligned}$$

For the term I_1 we use the Young and the Hölder inequalities to obtain for arbitrary $\varepsilon > 0$,

$$\mathrm{I}_1 \le \varepsilon \int_Q (m+1)^{\sigma-2} |\nabla m|^2 \mathrm{d}x\mathrm{d}t + C(\varepsilon) \|\boldsymbol{u}\|^2_{L^\infty(Q)} \int_Q (m+1)^{2s_0-\sigma+2} \mathrm{d}x\mathrm{d}t.$$

Finally, substituting these estimates into (4.6), summing the result over $i = 1, \cdots, N$ and using (2.28), we obtain

$$\begin{aligned}
&\int_Q \sum_{i=1}^N m (f^i(\boldsymbol{v}, m) + m f^i_m(\boldsymbol{v}, m)) \mathrm{d}x\mathrm{d}t \\
&\le N\varepsilon \int_Q (m+1)^{\sigma-2} |\nabla m|^2 \mathrm{d}x\mathrm{d}t \\
&\quad + C(\varepsilon) \|\boldsymbol{u}\|^2_{L^\infty(Q)} \Big(1 + \int_Q (m+1)^{2s_0-\sigma+2} \mathrm{d}x\mathrm{d}t \Big) \\
&\quad + C_1 \sqrt{N} \int_Q \frac{m^2 |\partial_m b_1(m)|}{b_1(m)} |\boldsymbol{v}| \sum_{i=1}^N |f^i_{v^i}| \mathrm{d}x\mathrm{d}t \\
&\le N\varepsilon \int_Q (m+1)^{\sigma-2} |\nabla m|^2 \mathrm{d}x\mathrm{d}t \\
&\quad + C(\varepsilon) \|\boldsymbol{u}\|^2_{L^\infty(Q)} \Big(1 + \int_Q (m+1)^{2s_0-\sigma+2} \mathrm{d}x\mathrm{d}t \Big)
\end{aligned}$$

$$+\int_Q \delta \sum_{i=1}^N m(f^i(\boldsymbol{v},m) + m f^i_m(\boldsymbol{v},m)) + K(1+m^{2s_0})\mathrm{d}x\mathrm{d}t. \tag{4.7}$$

Consequently, since $\delta < 1$ we can absorb the last term by the left-hand side. Therefore, we can use the sum coerciveness, i.e., the assumption (2.18), to deduce that

$$\int_Q m(m^r+1)|\boldsymbol{v}|^2\mathrm{d}x\mathrm{d}t \le C\|\boldsymbol{u}\|_{L^\infty(Q)} + C(\varepsilon)\|\boldsymbol{u}\|^2_{L^\infty(Q)}\int_Q (m+1)^{2s_0-\sigma+2}\mathrm{d}x\mathrm{d}t$$
$$+ C\int_Q (m+1)^{2s_0+1}\mathrm{d}x\mathrm{d}t + C\varepsilon\int_Q (m+1)^{\sigma-2}|\nabla m|^2\mathrm{d}x\mathrm{d}t. \tag{4.8}$$

Finally, using the estimate (4.8) in (4.3), we see that (recalling (2.37))

$$\begin{aligned}&\sup_{t\in(0,T)}\|m(t)+1\|^\sigma_\sigma + \int_Q (m+1)^{\sigma-2}|\nabla m|^2\mathrm{d}x\mathrm{d}t\\ &\le C\Big(1+\int_Q m(m^r+1)|\boldsymbol{v}|^2\mathrm{d}x\mathrm{d}t\Big)\\ &\le C\int_Q (m+1)^{2s_0+1} + \varepsilon(m+1)^{\sigma-2}|\nabla m|^2\mathrm{d}x\mathrm{d}t\\ &\quad + C(\varepsilon)\|\boldsymbol{u}\|^2_{L^\infty(Q)}\Big(1+\int_Q (m+1)^{2s_0-\sigma+2}\mathrm{d}x\mathrm{d}t\Big).\end{aligned}$$

Thus, setting $\varepsilon > 0$ sufficiently small, we obtain the desired inequality (4.2).

Our next goal is to improve the information coming from (4.2) such that the estimate on E depends only on $\boldsymbol{u}$ and not on m. It means that we want to show that

$$E \le C\Big(1+\|\boldsymbol{u}\|_{L^\infty(Q)}^{\frac{2(\sigma(d+2)-d)}{\sigma(d+2)-d-((2s_0-\sigma+1)_+(d+2))}}\Big) \tag{4.9}$$

provided that

$$2s_0 + 2s < r + \frac{2}{d+2}, \tag{4.10}$$
$$s_0 + 2s < r + \frac{1}{d+2}. \tag{4.11}$$

The estimate (4.9) will be shown by a certain interpolation of (4.1)–(4.2).

Proof of (4.9) First, we recall the following parabolic interpolation inequality:

$$\begin{aligned}\int_0^T \|u\|^{\frac{2(d+2)}{d}}_{\frac{2(d+2)}{d}}\mathrm{d}t &\le \int_0^T \|u\|_2^{\frac{4}{d}}\|u\|^2_{1,2}\,\mathrm{d}t\\ &\le C\Big(\sup_{t\in(0,T)}\|u(t)\|_2^2 + \int_0^T \|\nabla u\|_2^2\,\mathrm{d}t\Big)^{\frac{d+2}{d}}.\end{aligned} \tag{4.12}$$

Next, we apply (4.12) onto the function $u := (m+1)^{\frac{\sigma}{2}}$ with σ given by (2.37) to get

$$\begin{aligned}\int_Q (m+1)^{\frac{\sigma(d+2)}{d}}\mathrm{d}x\mathrm{d}t &= \int_0^T \|(m+1)^{\frac{\sigma}{2}}\|^{\frac{2(d+2)}{d}}_{\frac{2(d+2)}{d}}\,\mathrm{d}t\\ &\le C\Big(\sup_{t\in(0,T)}\|(m(t)+1)^{\frac{\sigma}{2}}\|_2^2 + \int_0^T \|\nabla(m+1)^{\frac{\sigma}{2}}\|_2^2\,\mathrm{d}t\Big)^{\frac{d+2}{d}}\end{aligned}$$

$$\le C\Big(\sup_{t\in(0,T)}\|(m(t)+1)\|_\sigma^\sigma+\int_Q (m+1)^{\sigma-2}|\nabla m|^2 \mathrm{d}x\mathrm{d}t\Big)^{\frac{d+2}{d}}$$
$$= CE^{\frac{d+2}{d}}. \tag{4.13}$$

Finally, we also use the a priori estimate (4.1) and the interpolation inequality

$$\|\cdot\|_q \le \|\cdot\|_1^{1-\frac{q-1}{q}\frac{\sigma(d+2)}{\sigma(d+2)-d}} \|\cdot\|_{\frac{\sigma(d+2)}{d}}^{\frac{q-1}{q}\frac{\sigma(d+2)}{\sigma(d+2)-d}},$$

which is valid for all $1 \le q \le \frac{\sigma(d+2)}{d}$, to obtain

$$\int_Q (m+1)^q \mathrm{d}x\mathrm{d}t \le \|m+1\|_{L^1(Q)}^{q-\frac{\sigma(d+2)(q-1)}{\sigma(d+2)-d}} \|m+1\|_{L^{\frac{\sigma(d+2)}{d}}(Q)}^{\frac{\sigma(d+2)(q-1)}{\sigma(d+2)-d}}$$
$$\le C\Big(\int_Q (m+1)^{\frac{\sigma(d+2)}{d}} \mathrm{d}x\mathrm{d}t\Big)^{\frac{d(q-1)}{\sigma(d+2)-d}}$$
$$\le E^{\frac{(q-1)(d+2)}{\sigma(d+2)-d}}, \tag{4.14}$$

where for the last inequality we used (4.13).

Next, we use (4.14) to handle the right-hand side of (4.2). Assuming that

$$2s_0+1 \le \frac{\sigma(d+2)}{d}, \tag{4.15}$$

we get from (4.14),

$$\int_Q (m+1)^{2s_0+1}\mathrm{d}x\mathrm{d}t \le CE^{\frac{2s_0(d+2)}{\sigma(d+2)-d}} \le C\Big(1+\frac{1}{4}E\Big) \tag{4.16}$$

provided that

$$\frac{2s_0(d+2)}{\sigma(d+2)-d} < 1, \tag{4.17}$$

which by using of (2.37) can be shown to be equivalent to (4.10). Notice that (4.17) directly implies the validity of (4.15). Hence, we can absorb the first term on the right-hand side of (4.2) to get

$$E \le C\Big(1+\|\boldsymbol{u}\|^2_{L^\infty(Q)}\Big(1+\int_Q (m+1)^{2s_0-\sigma+2}\mathrm{d}x\mathrm{d}t\Big)\Big). \tag{4.18}$$

In case that

$$2s_0-\sigma+2 \le 1 \Leftrightarrow 2(s_0+s) \le r, \tag{4.19}$$

we can use (4.1) to conclude (4.9) directly. If (4.19) is not true, we again use (4.14) to get from (4.18)

$$E \le C(1+\|\boldsymbol{u}\|^2_{L^\infty(Q)} E^{\frac{(2s_0-\sigma+1)(d+2)}{\sigma(d+2)-d}}),$$

which after using the Young inequality leads to (4.9), provided that

$$\frac{(2s_0-\sigma+1)(d+2)}{\sigma(d+2)-d} < 1. \tag{4.20}$$

Note that (4.20) is a stronger assumption that (4.19) and it can be shown by using (2.37) that (4.20) is equivalent to (4.11). Hence the proof of (4.9) is complete.

4.2 Estimates for $\boldsymbol{u}$

This subsection is devoted to the uniform estimates on $\boldsymbol{u}$, which will still depend on m. To be more precise, we want to show that for arbitrary $p > d+2$ we have the estimate

$$\|\boldsymbol{u}\|_{L^\infty(Q)} \le C + C(p)\|m^{2s_0}\|^2_{L^p(Q)}. \tag{4.21}$$

Proof of (4.21) We start with the estimates for below for the quantity

$$w := \sum_{i=1}^N u^i.$$

It is not difficult to observe from (2.5), (3.1) and the fact that $2s \le r$, that w satisfies almost everywhere in Q,

$$-\partial_t w - \Delta w = \sum_{i=1}^N L^i(\boldsymbol{v}, m, \nabla \boldsymbol{u}) \ge -C(|\nabla w|^2 + m^{2s_0} + 1). \tag{4.22}$$

Next, let us consider w_1 a solution to

$$-\partial_t w_1 - \Delta w_1 = -Cm^{2s_0} \tag{4.23}$$

completed by zero initial condition, i.e., $w_1(T) = 0$. Then by a standard parabolic estimate, we obtain that

$$\|w_1\|_{L^\infty(Q)} + \|\nabla w_1\|_{L^\infty(Q)} \le C(p)\|m^{2s_0}\|_{L^p(Q)} \tag{4.24}$$

whenever $p > d+2$. Then subtracting (4.23) from (4.22), we obtain

$$-\partial_t(w-w_1) - \Delta(w-w_1) \ge -C(|\nabla w|^2 + 1) \ge -C|\nabla(w-w_1)|^2 - C(1+\|\nabla w_1\|^2_\infty). \tag{4.25}$$

Hence from the theory for subsolutions to parabolic equation (see [8]), we obtain

$$w - w_1 \ge -C(T)\max\{\|w(T)\|_\infty, (1+\|\nabla w_1\|^2_{L^\infty(Q)})\},$$

which together with (4.24) and the assumption that $\boldsymbol{u}(T) \in L^\infty$ leads to the final estimate from below

$$w \ge -C - C(p)\|m^{2s_0}\|^2_{L^p(Q)}, \tag{4.26}$$

which is valid for arbitrary $p > d+2$.

Next, we focus on estimates from above. Keeping the notation for w, we can derive from (2.5) that

$$-\partial_t(u^i - \varepsilon_0 w) - \Delta(u^i - \varepsilon_0 w) = L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) - \varepsilon_0 \sum_{j=1}^N L^j(m, \boldsymbol{v}, \nabla \boldsymbol{u}).$$

Hence, using (3.2), we get

$$-\partial_t(u^i - \varepsilon_0 w) - \Delta(u^i - \varepsilon_0 w)$$

$$\leq C(1+m^{2s_0}) + C\Big(1+\frac{m^{2s}+1}{m^r+1}\Big)|\nabla(u^i-\varepsilon_0 w)|^2$$
$$\leq C(1+m^{2s_0}) + C|\nabla(u^i-\varepsilon_0 w)|^2, \tag{4.27}$$

where the second inequality follows from the assumption $2s \leq r$. Thus, we can repeat step by step the procedure for w and using the fact that $\boldsymbol{u}(T) \in L^\infty(\Omega)$, we obtain

$$u^i - \varepsilon_0 w \leq C + C(p)\|m^{2s_0}\|^2_{L^p(Q)}. \tag{4.28}$$

Finally, we derive the uniform bound (4.21). First, summing (4.28) over $i = 1, \cdots, N$, we have

$$(1-N\varepsilon_0)w \leq CN + C(p)N\|m^{2s_0}\|^2_{L^p(Q)}.$$

Since $\varepsilon_0 < \frac{1}{2N}$, we can combine this estimate with (4.26) to get

$$|w| \leq C + C(p)\|m^{2s_0}\|^2_{L^p(Q)}. \tag{4.29}$$

Consequently, it follows from (4.28) that

$$u^i \leq C + C(p)\|m^{2s_0}\|^2_{L^p(Q)}. \tag{4.30}$$

Finally to obtain also estimate from below for u^i, we use (4.29)–(4.30) and get

$$u^i = w - \sum_{j\neq i} u^j \overset{(4.29)-(4.30)}{\geq} C + C(p)\|m^{2s_0}\|^2_{L^p(Q)},$$

which together with (4.30) implies the desired estimate (4.21). Hence the proof is complete.

4.3 Uniform L^∞ bounds

This subsection is devoted to the uniform bound for $\boldsymbol{u}$, which directly implies the part of uniform estimates stated in (2.36). Here, we combine (4.14), (4.9) and (4.21) to obtain the final bound. We go back to (4.21) and estimate the right-hand side. Although we need to choose $p > d+2$, we formally provide all computation for $p = d+2$ and in the final restriction on the size of s_0 (or σ) we just use the strict inequality sign. Hence, we need to estimate the term on the right-hand side of (4.21), i.e., the integral

$$\|m^{2s_0}\|^2_{d+2} = \Big(\int_Q m^{2s_0(d+2)}\mathrm{d}x\mathrm{d}t\Big)^{\frac{2}{d+2}}. \tag{4.31}$$

If $2s_0(d+2) < 1$ then the integral on the right-hand side of (4.21) is bounded due to (4.1) and therefore we immediately get

$$\|\boldsymbol{u}\|_{L^\infty(Q)} \leq C(\|\boldsymbol{u}_0\|_\infty, \|m_0\|_\infty). \tag{4.32}$$

Hence, in what follows, we assume the opposite case. Assuming that

$$2s_0(d+2) < \frac{\sigma(d+2)}{d} \Leftrightarrow r+1 > 2ds_0 + 2s, \tag{4.33}$$

we can use (4.14) with $q := 2s_0(d+2)$ and we deduce

$$\Big(\int_Q (m+1)^{2s_0(d+2)}\mathrm{d}x\mathrm{d}t\Big)^{\frac{2}{d+2}} \le CE^{\frac{(2s_0(d+2)-1)_+ 2}{\sigma(d+2)-d}} \tag{4.34}$$

and it follows from (4.21) that

$$\|\boldsymbol{u}\|_{L^\infty(Q)} \le C(1+E^{\frac{(2s_0(d+2)-1)_+ 2}{\sigma(d+2)-d}}). \tag{4.35}$$

Inserting this estimate into the right-hand side of (4.9), we also deduce

$$\|\boldsymbol{u}\|_{L^\infty(Q)} \le C(1+\|\boldsymbol{u}\|_{L^\infty(Q)}^{\frac{4(2s_0(d+2)-1)_+}{\sigma(d+2)-d-(2s_0-\sigma+1)_+(d+2)}}). \tag{4.36}$$

Hence, in case

$$\frac{4(2s_0(d+2)-1)_+}{\sigma(d+2)-d-(2s_0-\sigma+1)_+(d+2)} < 1, \tag{4.37}$$

we can absorb the right-hand side by left-hand side to obtain

$$\|\boldsymbol{u}\|_{L^\infty(Q)} \le C(\|\boldsymbol{u}(T)\|_\infty, \|m_0\|_\infty), \tag{4.38}$$

which implies a part of (2.36). Notice that (4.37) is a stronger assumption than (4.20) and therefore all needed assumptions, i.e., the assumptions (4.17) and (4.37), are already encoded in (2.38). Furthermore, using (4.2) and (4.9), we obtain also the bound for m and ∇m stated in (2.36). Finally, from (4.8), we deduce the bound for term with $m(m^r+1)|\boldsymbol{v}|^2$ in (2.36).

4.4 Uniform estimates for $\nabla \boldsymbol{u}$

This subsection is devoted to the last remaining part of (2.36), i.e., the part of the estimate with $\nabla\boldsymbol{u}$. Keeping the notation from the previous sections, we start with estimates for ∇w. Using (3.1) and the fact that $2s \le r$ we have

$$-\partial_t w - \Delta w \ge (m^r+1)|\boldsymbol{v}|^2 - C(|\nabla w|^2 + m^{2s_0}+1). \tag{4.39}$$

Next, we multiply (4.39) by $\mathrm{e}^{-2Cw} \ge 0$, integrate over Q and use integration by parts to obtain

$$\begin{aligned}&\int_Q (m^r+1)|\boldsymbol{v}|^2\mathrm{e}^{-2Cw} + 2C\mathrm{e}^{-2Cw}|\nabla w|^2\mathrm{d}x\mathrm{d}t\\ &\le \frac{1}{2C}\int_Q \partial_t \mathrm{e}^{-2Cw} + C(|\nabla w|^2 + m^{2s_0}+1)\mathrm{e}^{-2Cw}\mathrm{d}x\mathrm{d}t.\end{aligned} \tag{4.40}$$

Thus, we see that we can absorb the term with ∇w by the left-hand side and due to the L^∞ bound for $\boldsymbol{u}$ (see (4.38)), and $L^{2s_0(d+2)}$ bound for m (see (4.34)), we deduce from (4.40) that

$$\int_Q (m^r+1)|\boldsymbol{v}|^2 + |\nabla w|^2\mathrm{d}x\mathrm{d}t \le C(\|\boldsymbol{u}_T\|_\infty, \|m_0\|_\sigma). \tag{4.41}$$

Notice that the first in (4.41) together with (4.8) leads to the estimate (2.36) for term involving $|\boldsymbol{v}|^2$.

Next, we use the inequality (4.27), i.e.,

$$-\partial_t(u^i - \varepsilon_0 w) - \Delta(u^i - \varepsilon_0 w) \le C(1+m^{2s_0}) + C|\nabla(u^i-\varepsilon_0 w)|^2, \tag{4.42}$$

which we multiply by $\mathrm{e}^{2C(u^i-\varepsilon_0 w)}$ and integrate over Q. Repeating step by step the procedure (4.40)–(4.41) and using uniform bounds on $\boldsymbol{u}$ and m, we get

$$\int_Q |\nabla(u^i-\varepsilon_0 w)|^2 \mathrm{d}x\mathrm{d}t \le C(\|\boldsymbol{u}_T\|_\infty, \|m_0\|_\sigma). \tag{4.43}$$

Finally, combining (4.41) and (4.43), we have for all $i=1,\cdots,N$,

$$\begin{aligned}\int_Q |\nabla u^i|^2 \mathrm{d}x\mathrm{d}t &\le 2\int_Q |\nabla(u^i-\varepsilon_0 w)|^2 + \varepsilon_0^2|\nabla w|^2 \mathrm{d}x\mathrm{d}t \\ &\le C(\|\boldsymbol{u}_T\|_\infty, \|m_0\|_\sigma),\end{aligned}$$

which finishes the proof of (2.36).

5 Existence of Solution—Proof of Theorem 2.2

This section is devoted to the proof of the existence of a solution to (2.2), (2.5) and (2.30). Notice that due to the assumption (2.31)–(2.32), we know that (2.32) is equivalent to

$$\boldsymbol{v} = \boldsymbol{\omega}(m, \nabla \boldsymbol{u})$$

with a Carathéodory mapping $\boldsymbol{\omega}$. Therefore we can omit (2.30) and replace $\boldsymbol{v}$ by $\boldsymbol{\omega}(m,\nabla\boldsymbol{u})$ in (2.2) and (2.5) and solve the problem only for unknowns $(m,\boldsymbol{u})$. In fact, this is also the way how one can get the existence of a solution to an approximative problem. Nevertheless, for the sake of simplicity and to simplify the notation, we keep writing $\boldsymbol{v}$ in what follows.

Second, we do not provide the complete and rigorous proof here. We rather emphasize those steps that are different from the known procedure for Bellman systems. To be more specific, we provide here the proof of weak sequential stability, which is the key property of the system of equations we have in mind. It means that we shall consider a sequence of $(m^n,\boldsymbol{u}^n,\boldsymbol{v}^n)$ of smooth solutions to (2.2), (2.5) and (2.30) (which is however equivalent to (2.15), once the mapping $\boldsymbol{\omega}$ is well defined) corresponding sequence of initial data

$$\begin{aligned} &\boldsymbol{u}_T^n \to \boldsymbol{u}_T && \text{strongly in } L^1(\Omega;\mathbb{R}^N),\\ &\boldsymbol{u}_T^n \rightharpoonup^* \boldsymbol{u}_T && \text{weakly}^* \text{ in } L^\infty(\Omega;\mathbb{R}^N),\\ &m_0^n \to m_0 && \text{stongly in } L^\sigma(\Omega)\end{aligned} \tag{5.1}$$

with nonnegative m_0^n. Our goal is to show that

$$(m^n,\boldsymbol{u}^n,\boldsymbol{v}^n) \to (m,\boldsymbol{u},\boldsymbol{v}), \tag{5.2}$$

strongly in $L^1(Q)\times L^1(0,T;W^{1,2}(\Omega;\mathbb{R}^N))\times L^1(Q;\mathbb{R}^{NM})$, where the triple $(m,\boldsymbol{u},\boldsymbol{v})$ solves again (2.2), (2.5) and (2.30) with initial data $(\boldsymbol{u}_T,m_0)$. Indeed such a result suggests that the rigorous existence proof is doable. Indeed, approximating Hamiltonians H by a sequence of bounded functions $\{H^n\}_{n=1}^\infty$ and similarly $\boldsymbol{g}$ by a sequence of bounded $\{\boldsymbol{g}^n\}_{n=1}^\infty$, one may consider that for a such approximative system it is classical to obtain the existence of solution $(m^n,\boldsymbol{u}^n,\boldsymbol{v}^n)$ and the only remaining part of the proof is then the weak sequential stability. For details, how one can approximate Hamiltonians properly, we refer to [2].

5.1 Uniform a priori estimates

In this part we just use the result of Theorem 2.1, which holds for sufficiently smooth solutions. Indeed, we may assume that

$$\begin{aligned}&\sup_{t\in(0,T)}(\|m^n(t)\|_\sigma+\|\boldsymbol{u}^n(t)\|_\infty)+\int_Q|\nabla\boldsymbol{u}^n|^2+(m^n+1)^{\sigma-2}|\nabla m^n|^2\mathrm{d}x\mathrm{d}t\\&+\int_Q((m^n)^{r+1}+1)|\boldsymbol{v}^n|^2+(m^n)^{2s_0(d+2)}\mathrm{d}x\mathrm{d}t\\&\le C(\|\boldsymbol{u}^n_T\|_\infty,\|m^n_0\|_\sigma)\le C\end{aligned}\tag{5.3}$$

such that $(m^n,\boldsymbol{u}^n,\boldsymbol{v}^n)$ satisfies for all $\varphi\in\mathcal{C}_0^\infty(-\infty;T;W^{1,\infty}_{\mathrm{per}}(\Omega))$,

$$\int_Q -m^n\partial_t\varphi+\nabla m^n\cdot\nabla\varphi-m^n\boldsymbol{g}(\boldsymbol{v}^n,m^n)\cdot\nabla\varphi\mathrm{d}x\mathrm{d}t=\int_\Omega m^n_0\varphi(0)\,\mathrm{d}x,\tag{5.4}$$

for all $\varphi\in\mathcal{C}_0^\infty(0;\infty;W^{1,\infty}(\Omega))$,

$$\int_Q \boldsymbol{u}^n\partial_t\varphi+\nabla\boldsymbol{u}^n\cdot\nabla\varphi\mathrm{d}x\mathrm{d}t-\int_\Omega\boldsymbol{u}^n_T\varphi(T)\,\mathrm{d}x=\int_Q L(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n)\varphi\mathrm{d}x\mathrm{d}t\tag{5.5}$$

and almost everywhere in Q,

$$f^i_{v^i_j}(\boldsymbol{v}^n,m^n)+b_1(m^n)\sum_{k=1}^d\partial_{x_k}(u^n)^iA^i_{kj}=0\quad\text{in } Q\tag{5.6}$$

with $\boldsymbol{g}$ given as

$$\boldsymbol{g}(\boldsymbol{v}^n,m^n):=\sum_{j=1}^N b_1(m^n)A^j(v^n)^j+\boldsymbol{b}_0(m^n)\tag{5.7}$$

and $L(m^n,\boldsymbol{v}^n,\boldsymbol{u}^n)$ given as

$$\begin{aligned}L(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n):=&\,\boldsymbol{f}(\boldsymbol{v}^n,m^n)+m^n\boldsymbol{f}_m(\boldsymbol{v}^n,m^n)\\&+\nabla\boldsymbol{u}^n[\boldsymbol{g}(\boldsymbol{v}^n,m^n)+m^n\boldsymbol{g}_{m^n}(\boldsymbol{v}^n,m^n)].\end{aligned}\tag{5.8}$$

Next, we focus on the estimate for the time derivatives. First, using (3.3) and uniform bounds (2.36), we see that

$$\int_Q|L(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n)|\,\mathrm{d}x\,\mathrm{d}t\le C.\tag{5.9}$$

Consequently, we can deduce from (5.5) and also from (5.3) that for some $q>d$,

$$\int_0^T\|\partial_t\boldsymbol{u}\|_{(W^{1,q}_{\mathrm{per}}(\Omega;\mathbb{R}^N))^*}\,\mathrm{d}t\le C.\tag{5.10}$$

Similarly, using (2.22) and (5.3), we see that for some $q\in(1,\infty)$,

$$\int_0^T\|\partial_t m\|^{q'}_{(W^{1,q}_{\mathrm{per}}(\Omega))^*}\,\mathrm{d}t\le C.\tag{5.11}$$

5.3 Identification of L and the proof of (5.26)

This last part is devoted to the proof of (5.26). We follow the procedure developed in [2] with the necessary changes due to the presence of the mean field variable m. We also proceed here slightly formally and refer the interested reader to [2], where the very similar procedure is made rigorously. Defining

$$w^n := \sum_{i=1}^N (u^n)^i,$$

it follows from (5.5) that in the sense of distributions we have

$$-\partial_t w^n - \Delta w^n = \sum_{i=1}^N L(m^n, \boldsymbol{v}^n, \nabla \boldsymbol{u}^n). \tag{5.27}$$

Next, we multiply (5.27) by

$$\mathrm{e}^{-2Cw^n}$$

and obtain

$$\frac{1}{2C}\partial_t \mathrm{e}^{-2Cw^n} + \frac{1}{2C}\Delta \mathrm{e}^{-2Cw^n} = \sum_{i=1}^N L(m^n, \boldsymbol{v}^n, \nabla \boldsymbol{u}^n)\mathrm{e}^{-2Cw^n} + 2C\mathrm{e}^{-2Cw^n}|\nabla w^n|^2. \tag{5.28}$$

Consequently, multiplying the resulting identity by arbitrary nonnegative $\varphi \in W_0^{1,1}(0,T : L^\infty(\Omega) \cap W^{1,2}_{\mathrm{per}}(\Omega))$ and integrating the result over Q, we get by using integration by parts that

$$\begin{aligned} &-\frac{1}{2C}\int_Q \mathrm{e}^{-2Cw^n}\partial_t \varphi + \nabla \mathrm{e}^{-2Cw^n}\cdot \nabla \varphi \mathrm{d}x\mathrm{d}t \\ &= \int_Q \varphi \sum_{i=1}^N L(m^n, \boldsymbol{v}^n, \nabla \boldsymbol{u}^n)\mathrm{e}^{-2Cw^n} + 2C\mathrm{e}^{-2Cw^n}|\nabla w^n|^2. \end{aligned} \tag{5.29}$$

Next, we let $n \to \infty$. Thanks to (5.12)–(5.14), it is easy to deduce that

$$\lim_{n\to\infty}\int_Q \mathrm{e}^{-2Cw^n}\partial_t \varphi + \nabla \mathrm{e}^{-2Cw^n}\cdot \nabla \varphi \mathrm{d}x\mathrm{d}t = \int_Q \mathrm{e}^{-2Cw}\partial_t \varphi + \nabla \mathrm{e}^{-2Cw}\cdot \nabla \varphi \mathrm{d}x\mathrm{d}t. \tag{5.30}$$

Next, using (3.1), we get that

$$\sum_{i=1}^N L(m^n, \boldsymbol{v}^n, \nabla \boldsymbol{u}^n)\mathrm{e}^{-2Cw^n} + 2C\mathrm{e}^{-2Cw^n}|\nabla w^n|^2 \ge -C((m^n)^{2s_0} + 1).$$

Consequently, we see that the right-hand side of (5.29) is bounded from below by a strongly convergent function, so using the point-wise convergence result (5.23) and the Fatou lemma, we get

$$\begin{aligned} &\liminf_{n\to\infty}\int_Q \varphi\Big(\sum_{i=1}^N L(m^n, \boldsymbol{v}^n, \nabla \boldsymbol{u}^n)\mathrm{e}^{-2Cw^n} + 2C\mathrm{e}^{-2Cw^n}|\nabla w^n|^2\Big) \\ &\ge \int_Q \varphi\Big(\sum_{i=1}^N L(m, \boldsymbol{v}, \nabla \boldsymbol{u})\mathrm{e}^{-2Cw} + 2C\mathrm{e}^{-2Cw}|\nabla w|^2\Big). \end{aligned} \tag{5.31}$$

Hence, combining (5.30)–(5.31), we see that for all nonnegative $\varphi \in W_0^{1,1}(0,T : L^\infty(\Omega) \cap W^{1,2}_{\mathrm{per}}(\Omega))$ there holds

$$\begin{aligned} &-\frac{1}{2C}\int_Q \mathrm{e}^{-2Cw}\partial_t\varphi + \nabla \mathrm{e}^{-2Cw}\cdot\nabla\varphi \mathrm{d}x\mathrm{d}t \\ &\geq \int_Q \varphi\Big(\sum_{i=1}^N L(m,\boldsymbol{v},\nabla\boldsymbol{u})\mathrm{e}^{-2Cw} + 2C\mathrm{e}^{-2Cw}|\nabla w|^2\Big). \end{aligned} \tag{5.32}$$

Finally, taking $\varphi := \mathrm{e}^{2Cw}\psi$ with arbitrary $\psi \in W_0^{1,1}(0,T; L^\infty(\Omega)\cap W^{1,2}_{\mathrm{per}}(\Omega))$, we deduce that[4]

$$\int_Q w\partial_t\psi + \nabla w\cdot\psi \mathrm{d}x\mathrm{d}t \geq \int_Q \psi\Big(\sum_{i=1}^N L(m,\boldsymbol{v},\nabla\boldsymbol{u})\Big)\psi\mathrm{d}x\mathrm{d}t, \tag{5.33}$$

which implies that in the sense of distributions

$$-\partial_t w - \Delta w \geq \sum_{i=1}^N L(m,\boldsymbol{v},\nabla\boldsymbol{u}). \tag{5.34}$$

Similarly, we deduce the opposite type inequalities. It follows from (5.5) that for all $i = 1,\cdots,N$,

$$\begin{aligned} &-\partial_t((u^i)^n - \varepsilon_0 w^n) - \Delta((u^i)^n - \varepsilon_0 w^n) \\ &= L^i(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n) - \varepsilon_0\sum_{i=1}^N L^i(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n). \end{aligned} \tag{5.35}$$

Denoting $z^n := (u^i)^n - \varepsilon_0 w^n$ and multiplying (5.35) by e^{2Cz^n} we get

$$\begin{aligned} &-\frac{1}{2C}\partial_t \mathrm{e}^{2Cz^n} - \frac{1}{2C}\Delta \mathrm{e}^{2Cz^n} \\ &= \mathrm{e}^{2Cz^n}\Big(L^i(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n) - \varepsilon_0\sum_{i=1}^N L^i(m^n,\boldsymbol{v}^n,\nabla\boldsymbol{u}^n)\Big) - \mathrm{e}^{2Cz^n}|\nabla z^n|^2. \end{aligned} \tag{5.36}$$

Hence, using (3.2), we see that the right-hand side is bounded by an convergent sequence and therefore we can proceed similarly as before by using the Fatou lemma to obtain

$$-\partial_t(u^i - \varepsilon_0 w) - \Delta(u^i - \varepsilon_0 w) \leq L^i(m,\boldsymbol{v},\nabla\boldsymbol{u}) - \varepsilon_0\sum_{i=1}^N L^i(m,\boldsymbol{v},\nabla\boldsymbol{u}). \tag{5.37}$$

Thus summing with respect to $i = 1,\cdots,N$ and dividing by $(1-\varepsilon_0)$ we see that

$$-\partial_t w - \Delta w \leq \sum_{i=1}^N L^i(m,\boldsymbol{v},\nabla\boldsymbol{u}),$$

which combined with (5.34) gives

$$-\partial_t w - \Delta w = \sum_{i=1}^N L^i(m,\boldsymbol{v},\nabla\boldsymbol{u}). \tag{5.38}$$

[4]In fact we must mollify the test function with respect to the time variable and then to pas to the limit. Since such a procedure was explained in details in [2], we do not provide the complete proof here.

Consequently also we obtain from (5.37) that

$$-\partial_t u^i - \Delta u^i \le L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}). \tag{5.39}$$

Finally using (5.38)–(5.39), we get

$$\begin{aligned}-\partial_t u^i - \Delta u^i &= -\partial_t w - \Delta w + \sum_{j\neq i}(\partial_t u^j + \Delta u^j)\\ &\ge \sum_{j=1}^{N} L^j(m, \boldsymbol{v}, \nabla \boldsymbol{u}) - \sum_{j\neq i} L^j(m, \boldsymbol{v}, \nabla \boldsymbol{u})\\ &= L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}).\end{aligned}$$

Hence, (5.39) implies that

$$-\partial_t u^i - \Delta u^i = L^i(m, \boldsymbol{v}, \nabla \boldsymbol{u}) \tag{5.40}$$

and (2.41) follows. This completes the proof of Theorem 2.2.

References

[1] Bensoussan, A., Breit, D. and Frehse, J., Parabolic Bellman-systems with mean field dependence, *Appl. Math. Optim.*, **73**(3), 2016, 419–432.

[2] Bensoussan, A., Bulíček, 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, 1729–1750.

[3] Bensoussan, A., Frehse, J. and Yam, P., Mean field games and mean field type control theory, SpringerBriefs in Mathematics, Springer-Verlag, New York, 2013.

[4] Boccardo, L., Dall'aglio, A., Gallouët, T. and Orsina, L., Nonlinear parabolic equations with measure data, *J. Funct. Anal.*, **147**(2), 1997, 237–258.

[5] Boccardo, L. and Gallouët, T., Nonlinear elliptic and parabolic equations involving measure data, *J. Funct. Anal.*, **87**(1), 1989, 149–169.

[6] Gomes, D. A., Pimentel, E. A. and Sánchez-Morgado, H., Time-dependent mean-field games in the subquadratic case, *Comm. Partial Differential Equations*, **40**(1), 2015, 40–76.

[7] Gomes, D. A., Pimentel, E. A. and Voskanyan, V., Regularity Theory for Mean-Field Game Systems, Springer International Publishing, New York, 2016.

[8] Ladyzhenskaya, O. A. and Ural′tseva, N. N., Linear and quasilinear elliptic equations, Translated from the Russian by Scripta Technica, Leon Ehrenpreis (eds), Academic Press, New York, 1968.

[9] Lasry, J.-M. and Lions, P.-L., Jeux à champ moyen, I, Le cas stationnaire, *C. R. Math. Acad. Sci. Paris*, **343**(9), 2006, 619–625.

[10] Porretta, A., On the planning problem for a class of mean field games, *C. R. Math. Acad. Sci. Paris*, **351**(11–12), 2013, 457–462.

[11] Porretta, A., Weak solutions to fokker–planck equations and mean field games, *ARMA*, **216**(1), 2015, 1–62.