TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 376, Number 4, April 2023, Pages 2481–2523
https://doi.org/10.1090/tran/8760
Article electronically published on January 24, 2023

WELL-POSEDNESS OF MEAN FIELD GAMES MASTER EQUATIONS INVOLVING NON-SEPARABLE LOCAL HAMILTONIANS

DAVID M. AMBROSE AND ALPÁR R. MÉSZÁROS

ABSTRACT. In this paper we construct short time classical solutions to a class of master equations in the presence of non-degenerate individual noise arising in the theory of mean field games. The considered Hamiltonians are non-separable and *local* functions of the measure variable, therefore the equation is restricted to absolutely continuous measures whose densities lie in suitable Sobolev spaces. Our results hold for smooth enough Hamiltonians, without any additional structural conditions as convexity or monotonicity.

1. INTRODUCTION

The theory of mean field games was initiated around the same time by J.-M. Lasry and P.-L. Lions on the one hand ([27–29]) and by P. Caines, M. Huang and R. Malhamé ([26]) on the other hand. The main motivation of both groups was to characterize limits of Nash equilibria of stochastic (or deterministic) differential games, when the number of agents tends to infinity.

A central object in this theory is the so-called *master equation* introduced by P.-L. Lions in his lectures at Collège de France (cf. [30]). This is a nonlocal Hamilton-Jacobi equation set on the space of Borel probability measures, which encodes all the information about the game. One of the main features of the master equation is that it serves as an important tool to show the convergence/mean field limit of Nash equilibria of games with finite number of agents as the number of agents increases to infinity (cf. [11,15,16]). In particular, solutions to the master equation can be used to obtain fine quantitative estimates on the rate of convergence and these solutions typically provide ε-Nash equilibria for games with large but finite number of agents. In the same time, this equation contains all the information about the finite dimensional mean field game system and it also describes a precise quantitative stability of this system with respect to the initial distribution of the agents.

The concept of master equations has a long history in kinetic theory and mean field limits of particle systems (see for instance [32] and the references therein). The past couple of years have witnessed a great increase of literature on master equations arising in the theory of mean field games. Depending on the techniques used in these works to show the well-posedness of the corresponding master equations, one may group these results into three possible categories. We refer to a non-exhaustive list of works as follows: probabilistic ideas for problems including individual or common noise were used in [12,13,20,33]; variational techniques

Received by the editors May 18, 2021, and, in revised form, April 28, 2022.
2020 *Mathematics Subject Classification.* Primary 49N80; Secondary 91A16, 35K40.

(based on optimal transport or optimal control theory in Hilbert spaces, for problems without noise or with individual noise) were exploited in [8,19,23,31]; and finally PDE techniques were used in [10,11] to attack problems with common or individual noise. In most of these references, a special hypothesis is assumed on the Hamiltonian $\mathcal{H}$ appearing in the master equation, namely it is such that the momentum variable it is separated from the measure variable, i.e. it has the typical form of $\mathcal{H}(t,x,p,m) := H(t,x,p) + f(t,x,m)$. Moreover, in all the cases previously considered in the literature, the dependence of the Hamiltonian on the measure variable is *always* assumed to be (nonlocal) regularizing. To the best of our knowledge, [10,12] are the only works, where the authors show the short time well-posedness of master equations involving non-separable regularizing Hamiltonians under mild assumptions (these are supposed to be smooth enough with some additional growth condition of polynomial type in the momentum variable). The recent work [20] constructs global in time classical solutions to master equations involving non-separable Hamiltonians, under an additional, so-called displacement monotonicity assumption on the Hamiltonian. Finally, [31] constructs local in time classical solutions to the master equation in the deterministic setting for smooth regularizing Hamiltonians that have the separable structure.

In this paper, we show the existence and uniqueness of classical solutions of a class of so-called *first order* master equations (i.e. driven by non-degenerate individual noise, and so derivatives with respect to the measure variable appear only at first order) with non-separable Hamiltonians that depend *locally* on the measure variable. Because of this local dependence, we clearly need to restrict the domain of the master function to the set of absolutely continuous probability measures whose densities lie in a suitable Sobolev space $H^s(\mathbb{T}^d)$ (for some $s > 1$). Ours seems to be the first result on the well-posedness of master equations where the Hamiltonian is a local function of the measure variable. The choice of the physical space $\mathbb{T}^d$ is for convenience, to avoid non-compactness issues. Nevertheless, we expect our results to hold true, without major complications, in the setting of $\mathbb{R}^d$ as well (under suitable moment bounds on the measures).

It is a very natural and interesting question whether can one extend the short-time results to global in time ones, in presence of (the generalized version of) the Lasry-Lions monotonicity condition on the data (see for instance condition (1.6) in [2] imposed on non-separable Hamiltonians). This will be studied in a future work.

The equation in the center of our focus for $U : [0,T] \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d) \to \mathbb{R}$ reads as

$$
\begin{cases}
-\partial_t U(t,x,m) - \Delta U(t,x,m) + \mathcal{H}(t,x,\nabla U(t,x,m),m) \\
\qquad - \int_{\mathbb{T}^d} \nabla_y \cdot (\nabla_w U(t,x,m,y))\, dm(y) \\
\qquad + \int_{\mathbb{T}^d} \nabla_w U(t,x,m,y) \cdot D_p \mathcal{H}(t,y,\nabla U(t,y,m))\, dm(y) = 0, \\
(t,x,m) \in (0,T) \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d), \\
U(T,x,m) = G(x,m), \\
(x,m) \in \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d).
\end{cases} \tag{1.1}
$$

The precise assumptions on the data $\mathcal{H} : [0,T] \times \mathbb{T}^d \times \mathbb{R}^d \times [0,+\infty) \to \mathbb{R}$ and $G : \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ will be given in the next section. In the previous equation the special notation $\nabla_w U$ stands for the Wasserstein gradient of U with respect to the measure variable (see for instance in [6,11,21]). All other notations for derivatives are understood with respect to the time and spacial variables.

In our analysis we use PDE techniques in $H^s(\mathbb{T}^d)$ – when we look at the finite dimensional mean field games system as characteristics of the master equation – and similarly as in [11] and [10], it is natural to work at the level of the linearized system to obtain the necessary regularity estimates on the master function. Our approach, at the technical level, fundamentally differs from the ones in [11] and [10] (where the regularization effect of the Hamiltonian in the measure variable was indispensable): at the level of this linearized system, in lack of regularization effect of the Hamiltonian in the measure variable, we perform a careful analysis using the energy method adapted to forward-backward problems. These techniques were used previously by the first author in [4], to show the well-posedness of the underlying finite dimensional mean field games system. As a result of this, when showing the regularity of the master function with respect to the measure variable, it turns out to be very convenient to work in the metric space $(\mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d), \|\cdot\|_{H^{-1}})$. In fact, by the Sobolev embedding theorem, all the regularity estimates on the master function in this metric space will imply the corresponding estimates also in $(\mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d), \|\cdot\|_{H^s})$.

Because of the comparison between the H^{-1} and W_2 metrics (cf. [34]), we expect our results to be useful in studying the convergence problem of games with finitely many players to the mean field limit, in the presence of non-separable, local Hamiltonians. A breakthrough on the convergence problem in the direction of Hamiltonians depending locally on the measure variable has been recently achieved in [9]. Here the proof of the convergence is carried out via the solution of the master equation in which the dependence of the Hamiltonian on the measure variable becomes more and more singular as the number of agents is increasing. Because of this, a careful analysis had to be combined with special structural and monotonicity assumptions on the Hamiltonian. In seems unclear whether such an argument could work directly in our setting, because of the difference in nature of the assumptions on the Hamiltonian. In a future work we aim to pursue a more 'classical route' by first showing that the MFG system (for a class of Hamiltonians that depend locally on the measure variable) is well-posed in Sobolev spaces, for general measure initial data. Having such a result in hand (for which some initial investigation seems already promising), would hopefully translate to the level of the master equation and to the convergence problem. We remark again that some of the crucial estimates on the master function are carried out in Sobolev spaces of negative order, which give us hope to be able to extend the solution of the master equation to rougher probability measures.

Finally, let us remark that (finite dimensional) mean field games systems involving non-separable Hamiltonians that depend locally on the measure variable, while they appear very naturally in models coming from economics (cf. [1]), are poorly understood in the literature in general. Beside the works [4,5], models involving so-called congestion effects have been studied in [2,14,17,18,24,25].

The structure of the paper is as follows. In Section 2 we collect all our standing assumptions and some preliminary results from the literature. This section recalls the notions of derivatives of functions defined on measures that we use in the rest of the paper. In the same time, here we describe the roadmap of our analysis, with the precise steps that lead to the proof of our main theorem. Then in the upcoming Sections 3, 4 and 5 we provide all the arguments to fill the necessary details on the steps prescribed in Section 2. We end the main text with Subsection 5.1 where

we collect the necessary regularity estimates on the master function and conclude with the existence and uniqueness of a solution to (1.1). Finally, in Appendix A we provide a technical stability result on the mean field game system, which is the consequence of the results in [4].

2. Preliminaries, main assumptions and our main theorems

Let $\mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z}^d)$ stand for the flat torus embedded in $\mathbb{R}^d$. We denote by $\mathscr{P}(\mathbb{T}^d)$ the space of Borel probability measures on the flat torus $\mathbb{T}^d$, for $s > 1$ we set $\mathscr{Q} := \mathscr{P}(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$. For $R > 0$, we denote by $\mathscr{Q}_R$ the elements of $m \in \mathscr{Q}$ such that $\|m - \overline{m}\|_{H^s} \leq R$, where $\overline{m} := \frac{1}{(2\pi)^d}$. These stand for R-balls in $H^s(\mathbb{T}^d)$ centered at the uniform density on $\mathbb{T}^d$. Let $T > 0$.

Let $\mathcal{H} : [0,T] \times \mathbb{T}^d \times \mathbb{R}^d \times [0,+\infty) \to \mathbb{R}$ be a given Hamiltonian function and let $G : \mathbb{T}^d \times \mathscr{Q} \to \mathbb{R}$ be a given final cost. We assume the following hypotheses on the data.

Standing assumptions.

(H1) $\quad s > \max\{\lceil (d+5)/2 \rceil + 1; 4\lceil d/2 \rceil + 1\}$ is a given real number.

Since our results are built upon the ones in [4], first we recall all the assumptions present in [4]. First, let $\beta \in (\mathbb{N} \cup \{0\})^{2d+1}$ be a multi-index, associated to the arguments (t,x,p,q) of the Hamiltonian $\mathcal{H}$. The first d coordinates of β correspond to $(x_1,\dots,x_d)$, the following d coordinates correspond to $(p_1,\dots,p_2)$ (which is the placeholder for ∇u), while the last coordinate corresponds to q (which is the placeholder for the m variable). Suppose that there exists $\tilde{F} : [0,+\infty) \to [0,+\infty)$ non-decreasing such that

(H2) $$\left|\partial^\beta \mathcal{H}(\cdot,\cdot,\nabla u, m)\right|_\infty \leq \tilde{F}\left(|\nabla u|_\infty + |m|_\infty\right), \quad \forall\, \beta \in \mathbb{N}^{2d+1},\ |\beta| \leq s+2.$$

Suppose that for any $B \subset \mathbb{R}^{d+1}$ bounded set and $\beta \in \mathbb{N}^{2d+1}$ with $0 \leq |\beta| \leq 2$, $\exists c > 0$ such that if $(p^1,q^1),(p^2,q^2)$, then

(H3) $$|\partial^\beta \mathcal{H}(t,x,p^1,q^1) - \partial^\beta \mathcal{H}(t,x,p^2,q^2)| \leq c\left(\sum_{i=1}^d |p_i^1 - p_i^2| + |q^1 - q^2|\right), \quad \forall (t,x) \in [0,T] \times \mathbb{T}^d.$$

In Section 5 we will need a precise estimate on $D^2\mathcal{H}$ and $D^3\mathcal{H}$. To describe this, let us define the following quantities.

$$\begin{aligned} F_1 &= \mathcal{H}(t,x,\nabla\tilde{u},\tilde{m}) - \mathcal{H}(t,x,\nabla u,m) - D_p\mathcal{H}(t,x,\nabla u,m)\cdot\nabla(\tilde{u}-u) \\ &\quad - \partial_q\mathcal{H}(t,x,\nabla u,m)(\tilde{m}-m), \\ F_2 &= \operatorname{div}(mD_p\mathcal{H}) - \operatorname{div}(\tilde{m}\widetilde{D_p\mathcal{H}}) - \operatorname{div}((\tilde{m}-m)D_p\mathcal{H}) - \operatorname{div}(m(D^2_{pp}\mathcal{H})\nabla(\tilde{u}-u)) \\ &\quad - \operatorname{div}(m(D_p\partial_q\mathcal{H})(m-\tilde{m})), \end{aligned}$$

where we used the shorthand notations $\widetilde{D_p\mathcal{H}} = D_p\mathcal{H}(t,x,\nabla\tilde{u},\tilde{m})$ and $D_p\mathcal{H} = D_p\mathcal{H}(t,x,\nabla u,m)$ in the last line.

Let us notice that under the assumption that s satisfies (H1), we have that there exists $r > 0$ such that

(2.1) $$r > \lceil d/2 \rceil \quad \text{and} \quad s \geq 4r+1 > \max\{\lceil (d+5)/2 \rceil + 1; 4\lceil d/2 \rceil + 1\}.$$

Let $r > 0$ satisfy (2.1). We assume that for any $R > 0$ then there exists $c > 0$ such that

$$\|F_1\|^2_{H^r} \leq c(\|u - \tilde{u}\|^4_{H^{r+1}} + \|m - \tilde{m}\|^4_{H^r}), \tag{H3}$$

$$\|F_2\|^2_{H^{r-1}} \leq c(\|u - \tilde{u}\|^4_{H^r} + \|m - \tilde{m}\|^4_{H^{r-1}}), \tag{H4}$$

for all $m, \tilde{m}, u, \tilde{u}$ such that $\|m\|_{H^r}, \|\tilde{m}\|_{H^r}, \|u\|_{H^{r+1}}, \|\tilde{u}\|_{H^{r+1}} \leq R$.

Here and afterwards we simply use the notation $D_p\mathcal{H}$ and $\partial_q\mathcal{H}$ to refer to the derivatives of $\mathcal{H}$ with respect to the third (the placeholder for ∇u) and fourth (the placeholder for m) variables, respectively. Let us underline that $\partial_q\mathcal{H}$ is used instead of $\partial_m\mathcal{H}$ to avoid possible confusions with the Wasserstein derivative or L^2-Gâteaux derivative with respect to the measure variable and to emphasize that this is a standard 'local' derivative of the Hamiltonian function.

Remark 2.1. The Hamiltonians of the form

$$\mathcal{H}(t, x, p, q) = a(t, x)P(p)Q(q),$$

where $a : [0, +\infty) \times \mathbb{T}^d$ is smooth with bounded derivatives and $P : \mathbb{R}^d \to \mathbb{R}, Q : \mathbb{R} \to \mathbb{R}$ are polynomial functions, satisfy our standing assumptions (H2) through (H4). Involving transcendental functions in the form of $\mathcal{H}$ is also allowable, as $\mathcal{H}(t, x, p, q) = \sin(|p|^2)\ln(1 + q^2)$ or $\mathcal{H}(t, x, p, q) = \exp\{\cos(|p|^2q^3)\}$ also satisfy the assumptions.

Functions of finite regularity, such as fractional powers, are also admissible, but must be sufficiently regular. For instance, if we have $d = 1$ then we find $s = 6$ is acceptable in (H1). Then in (H2), we must have $\mathcal{H}$ differentiable with bounded derivatives at least 8 times; therefore, a Hamiltonian of the form $\mathcal{H}(t, x, p, q) = |p|^{17/2}m^2$ would be admissible.

The prior work [4] includes a result on Hamiltonians applicable to congestion problems; in such Hamiltonians, a power of the measure, m, appears in the denominator. This naturally carries with it an assumption that the initial measure be bounded away from zero. We expect that such Hamiltonians (e.g. $\mathcal{H}(t, x, p, q) = |p|^2/q^{1/2}$) could be treated from the point of view of the present results on the master equation, at the cost of this further restriction on the class of measures considered. We do not do so in the present work, but could pursue this direction in the future.

Finally, we mention that there may be Hamiltonians for which the arguments of [4] and the present work apply, even though the Hamiltonian may not satisfy exactly (H2) through (H4). The assumptions are in place to work for a very general class of Hamiltonians, and it is likely that for some particular Hamiltonians, fewer derivatives may be required.

We consider the final cost function G to be a nonlocal smoothing operator applied to a function of m. This is what was done also in [14] and [10]. The reason for the regularization is that the solutions (u, m) to (2.2) have the following regularity: $u(t, \cdot) \in H^s$ and $m(t, \cdot) \in H^{s-1}$. If u has data $G(m(T, \cdot))$ and G is not regularizing, this is a problem since $m(T, \cdot)$ is not in the space that u should be in. Thus we take G to be smoothing. This issue was discussed also in [4], and [4, Section 5.1] presents the corresponding well-posedness result on the MFG system in the case of

regularizing final cost functions (we note that the non-regularizing case is addressed there as well). We assume the following conditions on G.

There exists $\Upsilon > 0$ such that for all $m_1, m_2 \in H^s(\mathbb{T}^d)$,

$$\|G(m_1,\cdot) - G(m_2,\cdot)\|_{H^1}^2 \leq \Upsilon \|m_1 - m_2\|_{L^2}^2. \tag{H5}$$

We assume that there exists $\kappa > 0$ such that for any s satisfying (H1), we have

$$\left\| \frac{\delta G}{\delta m}(\cdot, m)\mu \right\|_{H^s}^2 \leq \kappa \|\mu\|_{H^{s-1}}^2, \ \forall m \in \mathscr{Q}_R. \tag{H6}$$

It will be convenient for us to take $\kappa \geq 1$, so we make this assumption. Here, $\frac{\delta G}{\delta m}$ stands for the Gâteaux derivative of G, that we define below. In the same time, by $\frac{\delta G}{\delta m}(\cdot, m)\mu$ we denote the action of $\frac{\delta G}{\delta m}(\cdot, m)$ on μ.

Finally, we assume that G is Gâteaux differentiable on $H^s(\mathbb{T}^d)$ (for s satisfying (H1)) and for any $R > 0$, there exists $c > 0$ such that

$$\left\| G(\cdot, \tilde{m}) - G(\cdot, m) - \frac{\delta G}{\delta m}(\cdot, m)(\tilde{m} - m) \right\|_{H^s} \leq c\|\tilde{m} - m\|_{H^s}^2, \quad \forall \tilde{m}, m \in \mathscr{Q}_R. \tag{H7}$$

Remark 2.2. The function

$$G(x, m) := [W * (g(m))](x),$$

where $W : \mathbb{T}^d \to \mathbb{R}$ is smooth and $g : \mathbb{R} \to \mathbb{R}$ is bounded and of class $C^{1,1}$ with bounded derivative, satisfies the assumptions (H5)–(H7). We note that the assumption (H5) could be modified to consider m_i in a bounded set in $H^s(\mathbb{T}^d)$, and then the function g would only need to be locally $C^{1,1}$.

2.1. **Notions of derivatives of functions defined on $\mathscr{P}(\mathbb{T}^d)$.** Let $V : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ and $m_0 \in \mathscr{P}(\mathbb{T}^d)$. We say that V is L^2-differentiable in Gâteaux sense at m_0 if there exists $\frac{\delta V}{\delta m}(m_0) : \mathbb{T}^d \to \mathbb{R}$ continuous such that

$$\lim_{\varepsilon \to 0} \frac{V(m_0 + \varepsilon\chi) - V(m_0)}{\varepsilon} =: \int_{\mathbb{T}^d} \frac{\delta V}{\delta m}(m_0)(y)\, d\chi(y)$$

for all χ signed Borel measure such that $\chi(\mathbb{T}^d) = 0$, independently on χ. Notice that whenever $\frac{\delta V}{\delta m}(m_0)$ exists, it is unique up to additive constants. In what follows it will be convenient to fix such a constant that $\int_{\mathbb{T}^d} \frac{\delta V}{\delta m}(m_0)(y) dm_0(y) = 0$. For notational convenience, we use the notation $\frac{\delta V}{\delta m}(m_0)\mu := \int_{\mathbb{T}^d} \frac{\delta V}{\delta m}(m_0)(y)\, d\mu(y)$.

Given $m_0 \in \mathscr{P}(\mathbb{T}^d)$, we denote as $L^2(m_0)$ the subset of Borel fields $v : \mathbb{T}^d \to \mathbb{R}^d$ which are m_0–square integrable. The Wasserstein tangent space at m_0, denoted as $\mathrm{Tan}_{m_0} \mathscr{P}(\mathbb{T}^d)$, is the closure of $\nabla C^\infty(\mathbb{T}^d)$ in $L^2(m_0)$. We say that $V : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ is differentiable at m_0 if it is sub- and super differentiable in the Wasserstein sense at m_0 (see for instance [6, 21, 22]). In this case we denote by $\nabla_w V(m_0) : \mathbb{T}^d \to \mathbb{R}^d$ its so-called *Wasserstein gradient*, which is the unique element in the intersection of $\mathrm{Tan}_{m_0} \mathscr{P}(\mathbb{T}^d)$ and the sub- and super differentials. Notice that if $\frac{\delta V}{\delta m}(m_0)$ exists and it is in $C^1(\mathbb{T}^d)$, we simply have that $\nabla_w V(m_0)(y) = \nabla_y \frac{\delta V}{\delta m}(m_0, y)$.

Definition 2.3. Let $s > 0$ and let $R > 0$. We say that a function $V : \mathscr{Q}_R \to \mathbb{R}$ is Gâteaux differentiable on $H^s(\mathbb{T}^d)$ if for all $m \in \mathscr{Q}_R$, $\frac{\delta V(m)}{\delta m} : \mathbb{T}^d \to \mathbb{R}$ exists and

$$\left| V(\tilde{m}) - V(m) - \int_{\mathbb{T}^d} \frac{\delta V}{\delta m}(m)(y) d(\tilde{m} - m)(y) \right| = o\left(\|m_0 - \tilde{m}_0\|_{H^s} \right),$$

for all $\tilde{m}, m \in \mathscr{Q}_R$.

Definition 2.4. Let s satisfy (H1) and let $R > 0$. We say that $U : [0,T] \times \mathbb{T}^d \times \mathscr{Q}_R \to \mathbb{R}$ is a solution to the master equation (1.1) if for all $m \in \mathscr{Q}_R$, $U(\cdot, \cdot, m) \in C^1([0,T] \times \mathbb{T}^d) \cap C([0,T]; C^{2,\alpha}(\mathbb{T}^d))$, $U(t,x,\cdot)$ is Gâteaux differentiable on $H^s(\mathbb{T}^d)$ (in the sense of Definition 2.3), with $\frac{\delta U}{\delta m}(t,x,m)(\cdot) \in H^s(\mathbb{T}^d)$, uniformly with respect to $(t,x) \in [0,T] \times \mathbb{T}^d$ and $\mathscr{Q}_R \ni m \mapsto \frac{\delta U}{\delta m}(t,\cdot,m)(y) \in H^{-s}(\mathbb{T}^d)$ is continuous, for all $(t,y) \in (0,T) \times \mathbb{T}^d$. Moreover, (1.1) is satisfied pointwise for all $(t,x,m) \in (0,T) \times \mathbb{T}^d \times \mathscr{Q}_R$.

Our main theorem of this paper can be formulated as follows.

Theorem 2.1. *Let $s, \mathcal{H}$ and G satisfy (H1)–(H7) and let $R > 0$. Then, there exists $T_{***} > 0$ (depending on the data $\mathcal{H}$ and G and R) such that the master equation (1.1) has a unique solution on $[0, T_{***}] \times \mathbb{T}^d \times \mathscr{Q}_R$ in the sense of Definition 2.4.*

2.2. **The strategy of the proof of our main theorem.** In order to prove Theorem 2.1, similarly as in [11, Chapter 3], we rely on the well-posedness of a finite dimensional MFG system, its linearization and the smoothness of the value function with respect to the initial measure. Let us mention, however, that even though the roadmap leading to the main results is following [11], our analysis is fundamentally different from the one present in this reference.

Let us recall the main result on the well-posedness of the MFG system, which is the basis of our analysis. This can be found in [4, Theorem 7].

Theorem 2.2. *Suppose that the assumptions (H1)–(H7) take place and let $R > 0$. There exists $\tilde{T} > 0$ (depending on R and the data) such that for all $0 \le t_0 < \tilde{T}$ the mean field games system*

$$\begin{cases} -\partial_t u - \Delta u + \mathcal{H}(t,x,\nabla u, m) = 0, & (t,x) \in (t_0, \tilde{T}) \times \mathbb{T}^d, \\ \partial_t m - \Delta m - \nabla \cdot (m D_p \mathcal{H}(t,x,\nabla u, m)) = 0, & (t,x) \in (t_0, \tilde{T}) \times \mathbb{T}^d, \\ m(t_0,x) = m_0(x), \quad u(\tilde{T},x) = G(x, m_{\tilde{T}}(x)), & x \in \mathbb{T}^d \end{cases} \tag{2.2}$$

has a unique classical solution (u,m) for any $m_0 \in \mathscr{Q}_R$. Moreover, this solution has the regularity

$$u \in L^\infty([t_0,\tilde{T}]; H^s(\mathbb{T}^d)) \cap L^2([t_0,\tilde{T}]; H^{s+1}(\mathbb{T}^d)) \cap C([t_0,\tilde{T}]; H^{s'}(\mathbb{T}^d)), \quad \forall s' \in [0,s)$$

and

$$m \in L^\infty([t_0,\tilde{T}]; H^{s-1}(\mathbb{T}^d)) \cap L^2([t_0,\tilde{T}]; H^s(\mathbb{T}^d)) \cap C([t_0,\tilde{T}]; H^{s'-1}(\mathbb{T}^d)), \quad \forall s' \in [0,s)$$

and (u,m) is uniformly bounded in the corresponding spaces by a constant depending only on $R > 0$, $\mathcal{H}$ and G.

Corollary 2.3. *By the Sobolev embedding theorem, one has*

$$u \in C([t_0,\tilde{T}]; C^{2,\alpha}(\mathbb{T}^d)) \quad \text{and} \quad m \in C([t_0,\tilde{T}]; C^{2,\alpha}(\mathbb{T}^d)).$$

for some $\alpha \in [0,1)$. *Using the fact that* (u,m) *is a solution to the system* (2.2), *this further yields that one also has* $u, m \in C^1([t_0, \tilde{T}] \times \mathbb{T}^d)$.

The candidate for the solution to (1.1) can be defined as

$$U(t_0, x, m_0) := u(t_0, x), \tag{2.3}$$

where (u,m) is the unique solution to (2.2) with initial measure m_0 and initial time t_0. We aim to show U solves (1.1) in the sense of Definition (2.4).

For this, we follow the following major steps. Let (u,m) be a solution to (2.2).

Step 1. We linearize (2.2) around this solution (u,m) allowing perturbations of the form $m_0 + \varepsilon\mu_0$ on the initial data (here μ_0 is typically a finite signed measure with 0 mass). This system reads as

(2.4)
$$\begin{cases} -\partial_t v - \Delta v + D_p\mathcal{H}(t,x,\nabla u, m)\cdot\nabla v + \partial_q\mathcal{H}(t,x,\nabla u,m)\mu = 0, & (t,x)\in(t_0,T)\times\mathbb{T}^d, \\ \partial_t\mu - \Delta\mu - \nabla\cdot\left[\mu D_p\mathcal{H}(t,x,\nabla u,m) + mD_{pp}^2\mathcal{H}(t,x,\nabla u,m)\nabla v\right] \\ \qquad + \nabla\cdot[mD_p\partial_q\mathcal{H}(t,x,\nabla u,m)\mu] = 0, & (t,x)\in(t_0,T)\times\mathbb{T}^d, \\ \mu(t_0,\cdot) = \mu_0, \quad v(T,x) = \frac{\delta G}{\delta m}(x, m_T)\mu_T, & x\in\mathbb{T}^d. \end{cases}$$

We show that this linearized system has a solution (v,μ) in suitable Sobolev spaces. In particular we have Theorem 2.4.

Theorem 2.4. *Let* $R > 0$, $m_0 \in \mathscr{Q}_R$ *and let* (u,m) *be the solution to* (2.2) *(given in Theorem* 2.2*), let* $T > 0$. *Then, there exists* $T_* > 0$ *such that*

1. *if* $\mu_0 \in H^{s-1}(\mathbb{T}^d)$ *and* $0 < T < T_*$, *then there exist a unique pair* $(v,\mu) \in L^\infty([0,T]; H^s(\mathbb{T}^d)) \times L^\infty([0,T]; H^{s-1}(\mathbb{T}^d))$ *that solves* (2.4);
2. *if* $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$ *and* $0 < T < T_*$, *then there exists a unique solution* (v,μ) *to* (2.4) *such that* $v \in L^\infty([0,T]; H^{-s}(\mathbb{T}^d))$ *and* $\mu \in L^\infty([0,T]; H^{-s-1}(\mathbb{T}^d))$. *This solution is understood in the sense of distributions.*

The proof of part (1) of this theorem is provided in Section 3. Part (2) is a consequence of results in Section 4 (we detail this in Corollary 4.2).

Step 2. We show that if (v,μ) is a solution to (2.4) with initial condition μ_0, then the operator $\mu_0 \mapsto v$ is continuous and linear and there exists $K(t_0, x, m_0, \cdot) : \mathbb{T}^d \to \mathbb{R}$ such that

$$v(t_0, x) = \int_{\mathbb{T}^d} K(t_0, x, m_0, y)\, d\mu_0(y).$$

This way we have a candidate for the L^2-Gâteaux derivative of the master function, i.e.

$$K(t_0, x, m_0, \cdot) = \frac{\delta U}{\delta m}(t_0, x, m_0, \cdot). \tag{2.5}$$

This result is obtained as a consequence of a Riesz-type representation theorem. Let us comment on the regularity of $K(t_0, \cdot, m_0, \cdot)$. As a consequence of Theorem 2.4(2), K will have H^s regularity in the last variable. However, let us underline that for $(t_0, m_0, y) \in (0, T_*) \times \mathscr{Q}_R \times \mathbb{T}^d$ fixed, the application $x \mapsto K(t_0, x, m_0, y)$ will a priori have only H^{-s} regularity. Therefore, a special care is needed in the derivation of the master equation. This is in contrast with the corresponding results from [11], where as a consequence of the regularizing effect of the Hamiltonian in the measure variable, $\frac{\delta U}{\delta m}$ is shown to be smooth also in the second variable.

We provide the details on these arguments in Section 4 and Subsection 5.1.

Step 3. Lastly, by a Taylor expansion argument, we show that the kernel K obtained in the previous step is indeed corresponding to the L^2-Gâteaux derivative $\frac{\delta U}{\delta m}$, i.e. (2.5) is shown rigorously in a suitable sense. To achieve this, we argue as follows. Let $u(t_0,x)$ be the value function in (2.2) with initial measure m_0 and $\tilde{u}(t_0,x)$ the value function in (2.2) with initial measure $\tilde{m}_0$. Let moreover $v(t_0,x)$ be the solution of the first equation in (2.4) with $\mu_0 := \tilde{m}_0 - m_0$. Then we show that

$$\|\tilde{u}(t_0,\cdot) - u(t_0,\cdot) - v(t_0,\cdot)\|_{H^r} = o\left(\|m_0 - \tilde{m}_0\|_{H^{-1}}\right),$$

for $r > 0$ satisfying (2.1), uniformly with respect to t_0.

Actually this further implies that

$$\left\| U(t_0,\cdot,\tilde{m}_0) - U(t_0,\cdot,m_0) - \int_{\mathbb{T}^d} K(t_0,\cdot,m_0,y)\, d(\tilde{m}_0 - m_0) \right\|_{H^r} = o\left(\|m_0 - \tilde{m}_0\|_{H^{-1}}\right),$$

uniformly with respect to t_0. So in particular, since $r > 0$ satisfies (2.1), the Sobolev embedding theorem yields

$$\sup_{t_0\in[0,T],x\in\mathbb{T}^d} \left| U(t_0,x,\tilde{m}_0) - U(t_0,x,m_0) - \int_{\mathbb{T}^d} K(t_0,x,m_0,y)\, d(\tilde{m}_0 - m_0) \right| = o\left(\|m_0 - \tilde{m}_0\|_{H^{-1}}\right),$$

i.e. the necessary differentiability property of U with respect to the measure variable.

To perform the analysis in this step, we will rely on an important observation. Using the notation

$$z(t,x) := \tilde{u}(t,x) - u(t,x) - v(t,x), \tag{2.6}$$

we will have that z solves the equation

$$\begin{cases} -\partial_t z - \Delta z + D_p\mathcal{H}(t,x,\nabla u,m)\cdot\nabla z = -\mathcal{H}(t,x,\nabla\tilde{u},\tilde{m}) \\ \qquad\qquad + \mathcal{H}(t,x,\nabla u,m) + \partial_q\mathcal{H}(t,x,\nabla u,m)\mu \\ \qquad\qquad + D_p\mathcal{H}(t,x,\nabla u,m)\cdot(\nabla\tilde{u} - \nabla u), \\ z(T,\cdot) = G(x,\tilde{m}_T) - G(x,m_T) - \frac{\delta G}{\delta m}(x,m_T)\mu_T, \end{cases} \tag{2.7}$$

and thus, we essentially show that there exists $C > 0$ such that

$$\|z(t,\cdot)\|_{H^r} \le C\|m_0 - \tilde{m}_0\|^{\frac{5}{4}}_{H^{-1}}.$$

In fact the Sobolev embedding theorem further implies that

$$\|z(t,\cdot)\|_{H^r} \le C\|m_0 - \tilde{m}_0\|^{\frac{5}{4}}_{H^s},$$

so the differentiability property holds true in $H^s(\mathbb{T}^d)$. We provide the details of this step in Section 5, where the previous crucial estimate is provided by Theorem 5.1.

2.3. **Some preliminary estimates and Sobolev norms.** We define the operator Λ to be a Fourier multiplier operator with symbol

$$\mathcal{F}\Lambda(k) = \left(1+|k|^2\right)^{1/2},$$

so that Λ^{-1} is the operator with symbol

$$\mathcal{F}\Lambda^{-1}(k) = \frac{1}{(1+|k|^2)^{1/2}}.$$

For $f \in L^1(\mathbb{T}^d)$ we define

$$\hat{f}(k) = \mathcal{F}[f](k) := \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-ik\cdot x} f(x)dx, \qquad \forall k \in \mathbb{Z}^d.$$

For $l \in \mathbb{R}$ and $f \in H^l(\mathbb{T}^d)$ we define

$$\Lambda^l f(x) = \sum_{k\in\mathbb{Z}^d} |\hat{f}(k)|(1+|k|^2)^{l/2} e^{ik\cdot x}.$$

We have the norm

$$\|f\|_{H^l}^2 = \sum_{k\in\mathbb{Z}^d} |\hat{f}(k)|^2(1+|k|^2)^l = \langle \Lambda^l f, \Lambda^l f\rangle_{L^2} = \|\Lambda^l f\|_{L^2}^2. \tag{2.8}$$

We shall use the following convention: for $v = (v_1, \cdots, v_d) \in (H^l(\mathbb{T}^d))^d$ we set

$$\|v\|_{H^l}^2 := \sum_{j=1}^d \|v_i\|_{H^l}^2.$$

With this in mind, the identity $\widehat{\nabla f}(k) = ik\hat{f}(k)$ allows to obtain that

$$\begin{aligned}\|\Lambda^{-s-1}\nabla f\|_{L^2}^2 + \|\Lambda^{-s-1} f\|_{L^2}^2 &= \sum_{k\in\mathbb{Z}^d} \frac{|k|^2|\hat{f}(k)|^2}{(1+|k|^2)^{s+1}} + \sum_{k\in\mathbb{Z}^d} \frac{|\hat{f}(k)|^2}{(1+|k|^2)^{s+1}} \\ &= \sum_{k\in\mathbb{Z}^d} \frac{|\hat{f}(k)|^2}{(1+|k|^2)^s}.\end{aligned}$$

We read off

$$\|\Lambda^{-s-1}\nabla f\|_{L^2}^2 + \|\Lambda^{-s-1} f\|_{L^2}^2 = \|\Lambda^{-s} f\|_{L^2}^2. \tag{2.9}$$

If $f \in H^{-s}(\mathbb{T}^d)$ and $g \in H^s(\mathbb{T}^d)$ we have

$$\langle f, g\rangle = \langle \Lambda^{-s} f, \Lambda^s g\rangle_{L^2} \tag{2.10}$$

and so we may conclude

$$\|f\|_{H^{-s}} = \sup_{\|\phi\|_{H^s}=1} \langle f, \phi\rangle. \tag{2.11}$$

Let us recall some important inequalities that will play important roles in our analysis.

We will need a product estimate for negative-index Sobolev spaces; in particular we want

$$\|fg\|_{H^{-1}} \le \|f\|_{H^{-1}}\|g\|_{H^k} \tag{2.12}$$

for some $k > 0$ large enough. We can establish this through duality and the corresponding estimate for positive-index Sobolev spaces. In particular we have

(2.13)
$$\|fg\|_{H^{-1}} = \sup_{\|\phi\|_{H^1}=1} \langle fg, \phi\rangle_{L^2} = \sup_{\|\phi\|_{H^1}=1} \langle f, g\phi\rangle_{L^2} \leq \sup_{\|\phi\|_{H^1}=1} \|f\|_{H^{-1}} \|g\phi\|_{H^1}.$$

We then use the corresponding estimate for positive-index spaces (see, for example, [7]),

$$\|h\psi\|_{H^r} \leq c\|h\|_{H^r}\|\psi\|_{H^s}, \tag{2.14}$$

as long as $0 \leq r \leq s$ and $s > d/2$. So then (2.13) can be estimated further as

$$\|fg\|_{H^{-1}} \leq \sup_{\|\phi\|_{H^1}=1} c\|f\|_{H^{-1}}\|\phi\|_{H^1}\|g\|_{H^k} = c\|f\|_{H^{-1}}\|g\|_{H^k},$$

where our conditions on k are $k \geq 1$ and $k > d/2$. We can take $k = \lceil \frac{d+1}{2} \rceil$.

In the same spirit, for s satisfying (H1), we have the inequality

$$\|fg\|_{H^{-s}} \leq c\|f\|_{H^{-s}}\|g\|_{H^s}. \tag{2.15}$$

We also need the most standard Sobolev interpolation inequality. Let $0 < \alpha < \beta$ be given, and let $f \in H^\beta$. Then the following holds:

$$\|f\|_{H^\alpha} \leq c\|f\|_{L^2}^{1-\alpha/\beta}\|f\|_{H^\beta}^{\alpha/\beta}. \tag{2.16}$$

The proof of this can be found many places, including in [3].

3. Well-posedness of the linearized system

Theorem 3.1. *Let s satisfy (H1). Let $T > 0$ be given, and let $m_0 \in H^{s-1}$ be given. Let the payoff function G satisfy (H6). Let $u \in L^\infty([0,T]; H^s)$ and $m \in L^\infty([0,T]; H^{s-1})$ solve (2.2). Let $\mu_0 \in H^{s-1}$ be given. There exists $T_* > 0$ such that if $0 < T < T_*$, then there exist $v \in L^\infty([0,T]; H^s)$ and $\mu \in L^\infty([0,T]; H^{s-1})$ which satisfy (2.4).*

Proof. We set up an iterative scheme for v and μ,

$$\partial_t v^{n+1} = -\Delta v^{n+1} + \mathcal{J}_\varepsilon(D_p\mathcal{H} \cdot \nabla v^n) + \mathcal{J}_\varepsilon(\partial_q \mathcal{H} \mu^n), \tag{3.1}$$

(3.2)
$$\partial_t \mu^{n+1} = \Delta \mu^{n+1} + \mathcal{J}_\varepsilon \mathrm{div}(\mu^n D_p \mathcal{H}) + \mathcal{J}_\varepsilon \mathrm{div}(m D^2_{pp}\mathcal{H} \nabla v^n) + \mathcal{J}_\varepsilon \mathrm{div}(m D_p \partial_q \mathcal{H} \mu),$$

with mollified data

$$v^{n+1}(T,\cdot) = \mathcal{J}_\varepsilon\left(\frac{\delta G}{\delta m}\mu^n(T,\cdot)\right), \qquad \mu^{n+1}(0,\cdot) = \mathcal{J}_\varepsilon \mu_0.$$

For parameter $\varepsilon > 0$, the operator $\mathcal{J}_\varepsilon$ is a Friedrichs mollifier. We have suppressed the arguments of $\mathcal{H}$, but of course $\mathcal{H}$ depends on the underlying solution ∇u and m rather than the linearized quantities ∇v and μ. We intialize the iteration with $v^0 = 0$ and $\mu^0 = 0$. Note that v^0 and μ^0 are (trivially) infinitely smooth, and and v^{n+1} and μ^{n+1} satisfy forced heat equations with infinitely smooth data and forcing. Therefore v^{n+1} and μ^{n+1} exist and are infinitely smooth. Note that while of course the solutions depend on both regularization parameters n and ε, we suppress the dependence on ε for the time being.

We now make energy estimates for v^{n+1} and μ^{n+1}. Define E_v and E_μ to be

$$E_v = \frac{1}{2}\|v^{n+1}\|_{H^s}^2 = \frac{1}{2}\int_{\mathbb{T}^d} (\Lambda^s v^{n+1})^2 \, dx,$$

$$E_\mu = \frac{1}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 = \frac{1}{2}\int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})^2 \, dx.$$

Then the time derivative of E_v satisfies

$$\begin{aligned}
\frac{dE_v}{dt} &= \int_{\mathbb{T}^d} (\Lambda^s v^{n+1})(\Lambda^s \partial_t v^{n+1}) \, dx \\
&= -\int_{\mathbb{T}^d} (\Lambda^s v^{n+1})(\Lambda^s \Delta v^{n+1}) \, dx + \int_{\mathbb{T}^d} (\Lambda^s v^{n+1})(\mathcal{J}_\varepsilon \Lambda^s (D_p \mathcal{H} \cdot \nabla v^n)) \, dx \\
&\quad + \int_{\mathbb{T}^d} (\Lambda^s v^{n+1})(\mathcal{J}_\varepsilon \Lambda^s (\partial_q \mathcal{H} \mu^n) \, dx \\
&= I_v + II_v + III_v.
\end{aligned}$$

We integrate this in time from time t to time T, arriving at (after slight rearranging of terms)

$$\|v^{n+1}(t,\cdot)\|_{H^s}^2 = \|v^{n+1}(T,\cdot)\|_{H^s}^2 - 2\int_t^T I_v \, d\tau - 2\int_t^T II_v \, d\tau - 2\int_t^T III_v \, d\tau.$$

The time derivative of E_μ satisfies

$$\begin{aligned}
\frac{dE_\mu}{dt} &= \int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})(\Lambda^{s-1} \partial_t \mu^{n+1}) \, dx \\
&= \int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})(\Lambda^{s-1} \Delta \mu^{n+1}) \, dx + \int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})(\Lambda^{s-1} \mathcal{J}_\varepsilon \mathrm{div}(\mu^n D_p \mathcal{H})) \, dx \\
&\quad + \int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})(\Lambda^{s-1} \mathcal{J}_\varepsilon \mathrm{div}(m D_{pp}^2 \mathcal{H} \cdot \nabla v^n)) \, dx \\
&\quad + \int_{\mathbb{T}^d} (\Lambda^{s-1} \mu^{n+1})(\Lambda^{s-1} \mathcal{J}_\varepsilon \mathrm{div}(m D_p \partial_q \mathcal{H} \mu^n)) \, dx \\
&= I_\mu + II_\mu + III_\mu + IV_\mu.
\end{aligned}$$

We integrate this in time from time zero to time t, finding

$$\begin{aligned}
&\|\mu^{n+1}(t,\cdot)\|_{H^{s-1}}^2 \\
&= \|\mu^{n+1}(0,\cdot)\|_{H^{s-1}}^2 + 2\int_0^t I_\mu \, d\tau + 2\int_0^t II_\mu \, d\tau + 2\int_0^t III_\mu \, d\tau + 2\int_0^t IV_\mu \, d\tau.
\end{aligned}$$

We begin estimating with the terms involving I_v, II_v, and III_v. For I_v we integrate by parts:

$$I_v = \int_{\mathbb{T}_d} (\Lambda^s \nabla v^{n+1})^2 \, dx,$$

and integrating in time we have

$$-\int_t^T I_v = -\int_t^T \|\nabla v^{n+1}\|_{H^s}^2 \, d\tau.$$

For the term with II_v we use Young's inequality with positive parameter σ_1 and (2.14) to obtain

$$|II_v| \le \frac{\sigma_1}{2}\|v^{n+1}\|_{H^s}^2 + \frac{1}{2\sigma_1}\|D_p \mathcal{H} \cdot \nabla v^n\|_{H^s}^2 \le \frac{\sigma_1}{2}\|v^{n+1}\|_{H^s}^2 + \frac{c}{\sigma_1}\|\nabla v^n\|_{H^s}^2.$$

Integrating in time and making some elementary manipulations, this becomes

$$-\int_t^T II_v \, d\tau \leq \frac{\sigma_1 T}{2} \sup_{t\in[0,T]} \|v^{n+1}\|_{H^s}^2 + \frac{c}{\sigma_1}\int_t^T \|\nabla v^n\|_{H^s}^2 \, d\tau.$$

We treat the term with III_v similarly:

$$|III_v| \leq \frac{\sigma_2}{2}\|v^{n+1}\|_{H^s}^2 + \frac{1}{2\sigma_2}\|\partial_q \mathcal{H}\mu^n\|_{H^s}^2 \leq \frac{\sigma_2}{2}\|v^{n+1}\|_{H^s}^2 + \frac{c}{\sigma_2}\|\mu^n\|_{H^s}^2,$$

and integrating in time,

$$-\int_0^T III_v \, d\tau \leq \frac{\sigma_2 T}{2} \sup_{t\in[0,T]} \|v^{n+1}\|_{H^s}^2 + \frac{c}{\sigma_2}\int_t^T \|\mu^n\|_{H^s}^2 \, d\tau.$$

We next turn to the terms coming from E_μ. We begin with the I_μ term, integrating by parts:

$$I_\mu = -\int_{\mathbb{T}^d} (\Lambda^{s-1}\nabla\mu^{n+1})^2 \, dx.$$

Integrating this with respect to time, we find

$$\int_0^t I_\mu \, d\tau = -\int_0^t \|\nabla\mu^{n+1}\|_{H^{s-1}}^2 \, d\tau.$$

For II_μ, we again use Young's inequality, now with positive parameter σ_3 :

$$|II_\mu| \leq \frac{\sigma_3}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{1}{2\sigma_3}\|\mu^n D_p\mathcal{H}\|_{H^s}^2 \leq \frac{\sigma_3}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{c}{\sigma_3}\|\mu^n\|_{H^s}^2.$$

We use the inequality

$$\|\mu^n\|_{H^s}^2 \leq \|\mu^n\|_{H^{s-1}}^2 + c\|\nabla\mu^n\|_{H^{s-1}}^2, \tag{3.3}$$

and we integrate in time, finding

$$\int_0^t |II_\mu| \, d\tau \leq \frac{\sigma_3 T}{2} \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{cT}{\sigma_3} \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + \frac{c}{\sigma_3}\int_0^t \|\nabla\mu^n\|_{H^{s-1}}^2 \, d\tau.$$

We treat III_μ similarly:

$$|III_\mu| \leq \frac{\sigma_4}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{1}{2\sigma_4}\|mD_{pp}^2\mathcal{H}\cdot\nabla v^n\|_{H^s}^2 \leq \frac{\sigma_4}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{c}{\sigma_4}\|\nabla v^n\|_{H^s}^2.$$

Integrated in time, this becomes

$$\int_0^t |III_\mu| \, d\tau \leq \frac{\sigma_4 T}{2} \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{c}{\sigma_4}\int_0^t \|\nabla v^n\|_{H^s}^2 \, d\tau.$$

Finally, we work with IV_μ along similar lines,

$$|IV_\mu| \leq \frac{\sigma_5}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{1}{2\sigma_5}\|mD_p\partial_q\mathcal{H}\mu^n\|_{H^s}^2 \leq \frac{\sigma_5}{2}\|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{c}{\sigma_5}\|\mu^n\|_{H^s}^2.$$

We again use (3.3), and integrate in time to find

$$\int_0^t |IV_\mu| \, d\tau \leq \frac{\sigma_5 T}{2} \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 + \frac{cT}{\sigma_5} \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + \frac{c}{\sigma_5}\int_0^t \|\nabla\mu^n\|_{H^{s-1}}^2 \, d\tau.$$

We summarize what we have concluded so far, and we use (H6):

(3.4)
$$\begin{aligned}\|v^{n+1}(t,\cdot)\|_{H^s}^2 + \int_t^T \|\nabla v^{n+1}\|_{H^s}^2 \, d\tau \le{}& \kappa \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 \\ &+ (\sigma_1+\sigma_2)T \sup_{t\in[0,T]} \|v^{n+1}\|_{H^s}^2 \\ &+ \frac{cT}{\sigma_2} \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + \frac{c}{\sigma_1}\int_0^T \|\nabla v^n\|_{H^s}^2 \, d\tau \\ &+ \frac{c}{\sigma_2}\int_0^T \|\nabla \mu^n\|_{H^{s-1}}^2 \, d\tau,\end{aligned}$$

(3.5)
$$\begin{aligned}&\|\mu^{n+1}(t,\cdot)\|_{H^{s-1}}^2 + \int_0^t \|\nabla \mu^{n+1}\|_{H^{s-1}}^2 \, d\tau \\ &\quad\le \|\mu_0\|_{H^{s-1}}^2 + (\sigma_3+\sigma_4+\sigma_5)T \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 \\ &\qquad + cT\left(\frac{1}{\sigma_3}+\frac{1}{\sigma_5}\right) \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + c\left(\frac{1}{\sigma_3}+\frac{1}{\sigma_5}\right)\int_0^T \|\nabla \mu^n\|_{H^{s-1}}^2 \, d\tau \\ &\qquad + \frac{c}{\sigma_4}\int_0^T \|\nabla v^n\|_{H^s}^2 \, d\tau.\end{aligned}$$

We take $\sigma_1 = \sigma_2 = \frac{1}{8T}$, and we take $\sigma_3 = \sigma_4 = \sigma_5 = \frac{1}{12T}$. We multiply (3.4) by $\frac{1}{8\kappa}$, and we add the result to (3.5), taking into account these choices of our parameters. Neglecting the integral terms on the left-hand sides of (3.4) and (3.5), we are able to take the supremum in time. The result after some slight rearranging is

(3.6)
$$\begin{aligned}&\frac{3}{32\kappa} \sup_{t\in[0,T]} \|v^{n+1}\|_{H^s}^2 + \frac{5}{8} \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 \\ &\quad\le \|\mu_0\|_{H^{s-1}}^2 + cT^2 \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + cT\int_0^T \|\nabla v^n\|_{H^s}^2 + \|\nabla \mu^n\|_{H^{s-1}}^2 \, d\tau.\end{aligned}$$

If we instead neglect the other terms on the left-hand side of (3.4) and (3.5), we can instead find that

(3.7)
$$\begin{aligned}&\frac{1}{8\kappa}\int_0^T \|\nabla v^{n+1}\|_{H^s}^2 \, d\tau + \int_0^T \|\nabla \mu^{n+1}\|_{H^{s-1}}^2 \, d\tau \\ &\quad\le \frac{1}{32\kappa} \sup_{t\in[0,T]} \|v^{n+1}\|_{H^s}^2 + \frac{3}{8} \sup_{t\in[0,T]} \|\mu^{n+1}\|_{H^{s-1}}^2 + \|\mu_0\|_{H^{s-1}}^2 \\ &\qquad + cT^2 \sup_{t\in[0,T]} \|\mu^n\|_{H^{s-1}}^2 + cT\int_0^T \|\nabla v^n\|_{H^s}^2 + \|\nabla \mu^n\|_{H^{s-1}}^2 \, d\tau.\end{aligned}$$

Adding (3.6) and (3.7), and slightly rearranging, we conclude

$$\begin{aligned}&\frac{1}{16\kappa}\sup_{t\in[0,T]}\|v^{n+1}\|_{H^s}^2+\frac{1}{4}\sup_{t\in[0,T]}\|\mu^{n+1}\|_{H^{s-1}}^2+\frac{1}{8\kappa}\int_0^T\|\nabla v^{n+1}\|_{H^s}^2\,d\tau\\&+\int_0^T\|\nabla\mu^{n+1}\|_{H^{s-1}}^2\,d\tau\\&\le 2\|\mu_0\|_{H^{s-1}}^2+cT^2\sup_{t\in[0,T]}\|\mu^n\|_{H^{s-1}}^2+cT\int_0^T\|\nabla v^n\|_{H^s}^2+\|\nabla\mu^n\|_{H^{s-1}}^2\,d\tau.\end{aligned}\tag{3.8}$$

We can now perform our induction. We wish to prove that for all n, we have

$$\begin{aligned}&\frac{1}{16\kappa}\sup_{t\in[0,T]}\|v^n\|_{H^s}^2+\frac{1}{4}\sup_{t\in[0,T]}\|\mu^n\|_{H^{s-1}}^2+\frac{1}{8\kappa}\int_0^T\|\nabla v^n\|_{H^s}^2\,d\tau+\int_0^T\|\nabla\mu^n\|_{H^{s-1}}^2\,d\tau\\&\le 3\|\mu_0\|_{H^{s-1}}^2.\end{aligned}\tag{3.9}$$

Clearly (3.9) holds when $n=0$ since $v^0=\mu^0=0$. We assume (3.9) holds for some value of n. Then from (3.8), we see that for T sufficiently small, (3.9) holds for the value $n+1$. This restriction on the size of T is independent of n. We therefore have established (3.9) for all values of n, and this is a uniform bound on our iterates.

We now pass to the limit. Recall that our solutions depend on two parameters, since there is a mollification parameter ε which we have suppressed, in addition to the iteration parameter n. Note that the above bound (3.9) which is uniform with respect to n is also uniform with respect to ε. We will now take $\varepsilon=\frac{1}{n}$ and simply pass to the limit as $n\to\infty$. Our sequence (v^n,μ^n) is uniformly bounded in $L^\infty([0,T];H^s(\mathbb{T}^d))\times L^\infty([0,T];H^{s-1}(\mathbb{T}^d))$, with $s>3+\frac{d}{2}$. By Sobolev imbedding this implies that $\|\nabla v^n\|_{L^\infty}$, $\|\nabla\mu^n\|_{L^\infty}$ and $\|\Delta v^n\|_{L^\infty}$, $\|\Delta\mu^n\|_{L^\infty}$ are also uniformly bounded. Inspection of the iterated equations (3.1) and (3.2) together with the uniform bound and Sobolev embedding further implies that $\|\partial_t v^n\|_{L^\infty}$ and $\|\partial_t\mu^n\|_{L^\infty}$ are uniformly bounded. We conclude that v^n forms a bounded equicontinuous family, as does μ^n. By the Arzelà-Ascoli theorem, there is a uniformly convergent subsequence, which we do not relabel. We thus have a limit, (v,μ).

Our sequence (v^n,μ^n) converges to (v,μ) in $(C([0,T]\times\mathbb{T}^d))^2$, and thus also in $C([0,T];L^2(\mathbb{T}^d)\times L^2(\mathbb{T}^d))$. With the uniform bound (3.9) and the Sobolev interpolation inequality (2.16), we may then conclude that the convergence actually holds in $C([0,T];H^{s'}(\mathbb{T}^d)\times H^{s'-1}(\mathbb{T}^d))$, for any $s'<s$.

Now we may conclude that (v,μ) satisfies the appropriate system, which is (2.4). Integrating (3.1) with respect to time, we have

$$v^{n+1}(t,\cdot)=\mathcal{J}_{1/n}\left(\frac{\delta G}{\delta m}\mu^n(T,\cdot)\right)+\int_t^T\left(\Delta v^{n+1}-\mathcal{J}_{1/n}(D_p\mathcal{H}\cdot\nabla v^n)-\mathcal{J}_{1/n}(\partial_q\mathcal{H}\mu^n)\right)d\tau.\tag{3.10}$$

(Recall that we have set $\varepsilon=1/n$.) We have established uniform convergence in sufficiently regular spaces to be able to take the limit in every term in (3.10). We therefore have

$$v(t,\cdot)=\frac{\delta G}{\delta m}\mu(T,\cdot)+\int_t^T\left(\Delta v-D_p\mathcal{H}\cdot\nabla v-\partial_q\mathcal{H}\mu\right)d\tau.$$

We can treat μ analagously; we conclude that (2.4) is satisfied by (v,μ).

Finally, we note that since the unit ball of a Hilbert space is weakly compact, the sequences $v^n(t,\cdot)$ and $\mu^n(t,\cdot)$ have weak limits in $H^s(\mathbb{T}^d) \times H^{s-1}(\mathbb{T}^d)$ at every time. This limit must be the same as the limit previously found, therefore we may conclude $(v,\mu) \in L^\infty([0,T]; H^s(\mathbb{T}^d) \times H^{s-1}(\mathbb{T}^d))$ as well. □

Remark 3.1.

(1) By optimizing the time horizon in (3.7) and (3.6) in the above proof, we find that

$$T_* \geq c_0 \frac{1}{c\kappa},$$

where $c_0 > 0$ is a universal constant, κ is constant appearing in (H6), while

$$c = c\left(\|D_p\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^s}, \|\partial_q\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^s}, \right.$$
$$\left. \|mD^2_{pp}\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^s}, \|mD_p\partial_q\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^s}\right).$$

(2) In the same time, we have also that $\|v\|_{L^\infty(H^s)} \leq \tilde{c}_0\|\mu_0\|_{H^{s-1}}$, for some universal constant $\tilde{c}_0 > 0$.

Corollary 3.2. *Under the assumption that $s > \lceil \frac{d+5}{2} \rceil$, by Sobolev embedding theorem, the previous theorem yields $v \in C([0,T]; C^{2,\alpha}(\mathbb{T}^d))$ for some $\alpha \in [0,1)$.*

4. More estimates on the solution of the linearized system

To get the desired regularity on the kernel K with respect to the y variable, we will use a Riesz type representation theorem. To conclude that K is in H^s we will use Theorem 4.1, which shows that v is related to a continuous linear operator on H^{-s}.

Theorem 4.1. *Let v and μ_0 be as in Theorem 3.1. There exists $T_{**} > 0$ with $T_{**} \leq T_*$ and $c > 0$ such that if $T \in (0, T_{**})$, then for s satisfying (H1)*

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{H^{-s}} \leq c\|\mu_0\|_{H^{-s-1}}. \tag{4.1}$$

Proof. We define an energy, beginning with the piece

$$E_v = \frac{1}{2}\|v\|^2_{H^{-s}} = \int_{\mathbb{T}^d} \left(\Lambda^{-s}v\right)^2 \, dx.$$

We take the time derivative of E_v, finding

$$\frac{dE_v}{dt} = \int_{\mathbb{T}^d} \left(\Lambda^{-s}v\right)\left(\Lambda^{-s}\partial_t v\right) \, dx.$$

We now substitute from the $\partial_t v$ equation in (2.4), finding

$$\frac{dE_v}{dt} = -\int_{\mathbb{T}^d} \left(\Lambda^{-s}v\right)\left(\Lambda^{-s}\Delta v\right) dx + \int_{\mathbb{T}^d} \left(\Lambda^{-s}v\right)\left(\Lambda^{-s}(D_p\mathcal{H}\cdot\nabla v)\right) dx$$
$$+ \int_{\mathbb{T}^d} \left(\Lambda^{-s}v\right)\left(\Lambda^{-s}(\partial_q\mathcal{H}\mu)\right) dx = I_v + II_v + III_v,$$

where we have suppressed the arguments in $\mathcal{H}$. We then integrate in time from time t to time T, finding

$$E_v(t) = E_v(T) - \int_t^T I_v \, d\tau - \int_t^T II_v \, d\tau - \int_t^T III_v \, d\tau.$$

We integrate by parts with respect to the spatial variables in the term involving I_v :

$$-\int_t^T I_v \, d\tau = -\int_t^T \int_{\mathbb{T}^d} (\Lambda^{-s}\nabla v)^2 \, dx d\tau = -\int_t^T \|\Lambda^{-s}\nabla v\|_{L^2}^2 \, d\tau. \tag{4.2}$$

This term gives us a parabolic gain of regularity and will be useful as we continue our estimates.

For the term involving II_v, we begin by estimating with Cauchy-Schwarz, and we recall from (2.8) that for any ϕ, we have $\|\phi\|_{H^{-s}} = \|\Lambda^{-s}\phi\|_{L^2}$:

$$|II_v| \leq \|v\|_{H^{-s}} \|D_p\mathcal{H} \cdot \nabla v\|_{H^{-s}}.$$

We then use (2.15), finding

$$|II_v| \leq c\|v\|_{H^{-s}} \|D_p\mathcal{H}\|_{H^s} \|\nabla v\|_{H^{-s}}.$$

We use Young's inequality with positive parameter σ_1 :

$$|II_v| \leq \frac{c}{\sigma_1}\|v\|_{H^{-s}}^2 \|D_p\mathcal{H}\|_{H^s}^2 + \frac{\sigma_1}{2}\|\Lambda^{-s}\nabla v\|_{L^2}^2.$$

We then integrate in time:

$$-\int_t^T II_v \, d\tau \leq \frac{c}{\sigma_1}\int_t^T \|v\|_{H^{-s}}^2 \|D_p\mathcal{H}\|_{H^s}^2 d\tau + \frac{\sigma_1}{2}\int_t^T \|\Lambda^{-s}\nabla v\|_{L^2}^2 d\tau. \tag{4.3}$$

We now turn to III_v, estimating similarly to before:

$$|III_v| \leq \|v\|_{H^{-s}} \|\partial_q\mathcal{H}\mu\|_{H^{-s}} \leq c\|v\|_{H^{-s}} \|\partial_q\mathcal{H}\|_{H^s} \|\mu\|_{H^{-s}}.$$

We next use Young's inequality with positive parameter σ_2 :

$$|III_v| \leq \frac{c}{\sigma_2}\|v\|_{H^{-s}}^2 \|\partial_q\mathcal{H}\|_{H^s}^2 + \frac{\sigma_2}{2}\|\mu\|_{H^{-s}}^2.$$

We integrate this in time:

$$-\int_t^T III_v \, d\tau \leq \frac{c}{\sigma_2}\int_t^T \|v\|_{H^{-s}}^2 \|\partial_q\mathcal{H}\|_{H^s}^2 d\tau + \frac{\sigma_2}{2}\int_t^T \|\mu\|_{H^{-s}}^2 d\tau. \tag{4.4}$$

We next define the other piece of the energy, E_μ, which is

$$E_\mu = \frac{1}{2}\|\mu\|_{H^{-s-1}}^2 = \frac{1}{2}\int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)^2 dx.$$

The time derivative of this is

$$\frac{dE_\mu}{dt} = \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)\left(\Lambda^{-s-1}\partial_t\mu\right) dx.$$

We substitute from the $\partial_t\mu$ equation in (2.4), finding

$$\begin{aligned}
\frac{dE_\mu}{dt} &= \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)\left(\Lambda^{-s-1}\Delta\mu\right) dx - \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)\left(\Lambda^{-s-1}\nabla\cdot(\mu D_p\mathcal{H})\right) dx \\
&\quad - \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)\left(\Lambda^{-s-1}\nabla\cdot(m(D_{pp}^2\mathcal{H})\nabla v)\right) dx \\
&\quad - \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\mu\right)\left(\Lambda^{-s-1}\nabla\cdot(m(D_p\partial_q\mathcal{H})\mu)\right) dx \\
&= I_\mu + II_\mu + III_\mu + IV_\mu.
\end{aligned}$$

We integrate in time from time zero until time t,

$$E_\mu(t) = E_\mu(0) + \int_0^t I_\mu \, d\tau + \int_0^t II_\mu \, dt + \int_0^t III_\mu \, d\tau + \int_0^t IV_\mu \, d\tau.$$

We integrate the term involving I_μ by parts, to find

$$\int_0^t I_\mu \, d\tau = -\int_0^t \int_{\mathbb{T}^d} \left(\Lambda^{-s-1}\nabla\mu\right)^2 \, dx d\tau. \tag{4.5}$$

This is a parabolic term which will be helpful to us, as the term I_v was as well.

We next note that for any $f \in H^{-s}$, we have $\|\Lambda^{-s-1}\nabla f\|_{L^2} \leq c\|f\|_{H^{-s}}$. The H^{-s}-norm of f, though, is not equivalent to the L^2-norm of $\Lambda^{-s-1}\nabla f$, because of the $k = 0$ mode. We instead will use the inequality

$$\|f\|_{H^{-s}}^2 \leq c\|f\|_{H^{-s-1}}^2 + c\|\Lambda^{-s-1}\nabla f\|_{L^2}^2. \tag{4.6}$$

We begin to estimate the rest of the terms, starting with the II_μ term,

$$|II_\mu| \leq c\|\Lambda^{-s-1}\mu\|_{L^2}\|D_p\mathcal{H}\|_{H^s}\|\mu\|_{H^{-s}}.$$

We use Young's inequality with positive parameter σ_3,

$$|II_\mu| \leq \frac{c}{\sigma_3}\|D_p\mathcal{H}\|_{H^s}^2\|\mu\|_{H^{-s-1}}^2 + \frac{\sigma_3}{2}\|\mu\|_{H^{-s}}^2.$$

We finally use (4.6) with this, yielding

$$|II_\mu| \leq \frac{c}{\sigma_3}\|D_p\mathcal{H}\|_{H^s}^2\|\mu\|_{H^{-s-1}}^2 + \frac{c\sigma_3}{2}\|\mu\|_{H^{-s-1}}^2 + \frac{c\sigma_3}{2}\|\Lambda^{-s-1}\nabla\mu\|_{L^2}^2.$$

We integrate in time, finding

$$\begin{aligned}\int_0^t |II_\mu| d\tau \leq \frac{c}{\sigma_3}\int_0^t \|D_p\mathcal{H}\|_{H^s}^2\|\mu\|_{H^{-s-1}}^2 \, d\tau + \frac{c\sigma_3}{2}\int_0^t \|\mu\|_{H^{-s-1}}^2 \, d\tau \\ + \frac{c\sigma_3}{2}\int_0^t \|\Lambda^{-s-1}\nabla\mu\|_{L^2}^2 \, d\tau.\end{aligned} \tag{4.7}$$

We next consider III_μ :

$$|III_\mu| \leq c\|\Lambda^{-s-1}\mu\|_{L^2}\|mD_{pp}^2\mathcal{H}\|_{H^s}\|\nabla v\|_{H^{-s}}. \tag{4.8}$$

We use Young's inequality with positive parameter σ_4 :

$$|III_\mu| \leq \frac{c}{\sigma_4}\|\Lambda^{-s-1}\mu\|_{L^2}^2\|mD_{pp}^2\mathcal{H}\|_{H^s}^2 + \frac{\sigma_4}{2}\|\Lambda^{-s}\nabla v\|_{L^2}^2.$$

Integrated in time, this becomes

$$\int_0^t |III_\mu| d\tau \leq \frac{c}{\sigma_4}\int_0^t \|\Lambda^{-s-1}\mu\|_{L^2}^2\|mD_{pp}^2\mathcal{H}\|_{H^s}^2 \, d\tau + \frac{\sigma_4}{2}\int_0^t \|\Lambda^{-s}\nabla v\|_{L^2}^2 \, d\tau. \tag{4.9}$$

Our final term to estimate is IV_μ, and we use Cauchy-Schwarz to bound this as

$$|IV_\mu| \leq c\|\Lambda^{-s-1}\mu\|_{L^2}\|mD_p\partial_q\mathcal{H}\|_{H^s}\|\mu\|_{H^{-s}}.$$

We use Young's inequality with positive parameter σ_5 :

$$|IV_\mu| \leq \frac{c}{\sigma_5}\|\Lambda^{-s-1}\mu\|_{L^2}^2\|mD_p\partial_q\mathcal{H}\|_{H^s}^2 + \frac{\sigma_5}{2}\|\mu\|_{H^{-s}}^2.$$

Next we use (4.6), yielding

$$|IV_\mu| \leq \frac{c}{\sigma_5}\|\Lambda^{-s-1}\mu\|_{L^2}^2\|mD_p\partial_q\mathcal{H}\|_{H^s}^2 + \frac{c\sigma_5}{2}\|\mu\|_{H^{-s-1}}^2 + \frac{c\sigma_5}{2}\|\Lambda^{-s-1}\nabla\mu\|_{L^2}^2.$$

Integrated in time, this is

$$\int_0^t |IV_\mu|\ d\tau \le \frac{c}{\sigma_5}\int_0^t \|\Lambda^{-s-1}\mu\|_{L^2}^2 \|mD_p\partial_q\mathcal{H}\|_{H^s}^2\ d\tau + \frac{c\sigma_5}{2}\Big(\int_0^t \|\mu\|_{H^{-s-1}}^2\ d\tau + \int_0^t \|\Lambda^{-s-1}\nabla\mu\|_{L^2}^2\ d\tau\Big). \tag{4.10}$$

We choose our parameters from Young's inequality as follows:

$$\frac{\sigma_1}{2}=\frac{1}{8},\qquad \frac{\sigma_2}{2}=\frac{1}{12},\qquad \frac{c\sigma_3}{2}=\frac{1}{12},\qquad \frac{\sigma_4}{2}=\frac{1}{8\kappa},\qquad \frac{c\sigma_5}{2}=\frac{1}{12}.$$

Then, our findings from our energy estimates can be summarized in the following inequalities. First by (4.2)–(4.4) we have

$$E_v(t)+\frac{7}{8}\int_t^T \|\Lambda^{-s}\nabla v\|_{L^2}^2\ d\tau \le E_v(T)+4c\int_t^T \|v\|_{H^{-s}}^2\|D_p\mathcal{H}\|_{H^s}^2 d\tau + 6c\int_t^T \|v\|_{H^{-s}}^2\|\partial_q\mathcal{H}\|_{H^s}^2 d\tau+\frac{1}{12}\int_0^T \|\mu\|_{H^{-s}}^2\ d\tau.$$

Increasing the value of c if necessary we have

$$E_v(t)+\frac{7}{8}\int_t^T \|\Lambda^{-s}\nabla v\|_{L^2}^2\ d\tau \le E_v(T)+cT\|v\|_{H^{-s}}^2+\frac{1}{12}\int_0^T \|\mu\|_{H^{-s}}^2\ d\tau.$$

We use (H6) and the fact that $v(T,x)=D_qG(x,m_T)\mu_T$ to obtain

$$E_v(t)+\frac{7}{8}\int_t^T \|\Lambda^{-s}\nabla v\|_{L^2}^2\ d\tau \le \frac{\kappa}{2}\|\mu(T,\cdot)\|_{H^{-s-1}}^2+cT\sup_{t\in[0,T]}\|v\|_{H^{-s}}^2 + \frac{1}{12}\int_0^T \|\mu\|_{H^{-s}}^2\ d\tau. \tag{4.11}$$

Since each one of the two terms at the left-hand side of (4.11) is nonnegative we conclude that they are less than or equal to the expression at the right-hand side of (4.11). Hence,

$$\max\Big\{\frac{1}{2}\sup_{t\in[0,T]}\|v\|_{H^{-s}}^2,\ \frac{7}{8}\int_0^T \|\Lambda^{-s}\nabla v\|_{L^2}^2\ d\tau\Big\} \le \frac{\kappa}{2}\|\mu(T,\cdot)\|_{H^{-s-1}}^2+cT\sup_{t\in[0,T]}\|v\|_{H^{-s}}^2+\frac{1}{12}\int_0^T \|\mu\|_{H^{-s-1}}^2\ d\tau + \frac{1}{12}\int_0^T \|\nabla\mu\|_{H^{-s-1}}^2\ d\tau. \tag{4.12}$$

Arguing as above, we combine (4.5) (4.7) (4.9) (4.10) to obtain

$$E_\mu(t)+\frac{5}{6}\int_0^T \|\Lambda^{-s-1}\nabla\mu\|_{L^2}^2\ d\tau \le E_\mu(0)+cT\sup_{t\in[0,T]}\|\mu\|_{H^{-s-1}}^2+\frac{1}{8\kappa}\int_0^T \|\Lambda^{-s}\nabla v\|_{L^2}^2\ d\tau.$$

Taking into account the fact that each one of the expressions at the left and side of the previous inequality are nonnegative, we conclude that

$$\begin{aligned}\max\left\{\frac{1}{2}\sup_{t\in[0,T]}\|\mu\|^2_{H^{-s-1}},\ \frac{5}{6}\int_0^T\|\Lambda^{-s-1}\nabla\mu\|^2_{L^2}d\tau\right\}\\ \leq\frac{1}{2}\|\mu_0\|^2_{H^{-s-1}}+cT\sup_{t\in[0,T]}\|\mu\|^2_{H^{-s-1}}+\frac{1}{8\kappa}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau.\end{aligned}\tag{4.13}$$

This gives an upper bound on $\kappa\|\mu(T,\cdot)\|^2_{H^{-s-1}}$ which we can use in (4.12) to obtain

$$\begin{aligned}\max\left\{\frac{1}{2}\sup_{t\in[0,T]}\|v\|^2_{H^{-s}},\ \frac{7}{8}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau\right\}\\ \leq\frac{\kappa}{2}\|\mu(0,\cdot)\|^2_{H^{-s-1}}+cT\sup_{t\in[0,T]}\|v\|^2_{H^{-s}}+\frac{1}{12}\int_0^T\|\mu\|^2_{H^{-s-1}}\,d\tau\\ +\frac{1}{12}\int_0^T\|\nabla\mu\|^2_{H^{-s-1}}\,d\tau+\frac{1}{8}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau.\end{aligned}\tag{4.14}$$

We combine (4.13) and (4.14) to conclude that

$$\begin{aligned}&\frac{1}{2}\sup_{t\in[0,T]}\|\mu\|^2_{H^{-s-1}}+\frac{5}{6}\int_0^T\|\nabla\mu\|^2_{H^{-s-1}}d\tau+\frac{1}{2}\sup_{t\in[0,T]}\|v\|^2_{H^{-s}}+\frac{7}{8}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau\\ &\leq\|\mu_0\|^2_{H^{-s-1}}+cT\sup_{t\in[0,T]}\|\mu\|^2_{H^{-s-1}}+\frac{1}{4\kappa}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau\\ &+\kappa\|\mu_0\|^2_{H^{-s-1}}+cT\sup_{t\in[0,T]}\|v\|^2_{H^{-s}}+\frac{1}{6}\int_0^T\|\mu\|^2_{H^{-s-1}}\,d\tau+\frac{1}{6}\int_0^T\|\nabla\mu\|^2_{H^{-s-1}}\,d\tau\\ &+\frac{1}{4}\int_0^T\|\Lambda^{-s}\nabla v\|^2_{L^2}\,d\tau.\end{aligned}$$

Thus,

$$\begin{aligned}&\left(\frac{1}{3}-cT\right)\sup_{t\in[0,T]}\|\mu\|^2_{H^{-s-1}}+\frac{2}{3}\int_0^T\|\nabla\mu\|^2_{H^{-s-1}}d\tau+\left(\frac{1}{2}-cT\right)\sup_{t\in[0,T]}\|v\|^2_{H^{-s}}\\ &\qquad+\left(\frac{5}{8}-\frac{1}{4\kappa}\right)\int_0^T\|\nabla v\|^2_{H^{-s}}\,d\tau\\ &\leq(\kappa+1)\|\mu_0\|^2_{H^{-s-1}}.\end{aligned}$$

Since $\kappa\geq 1$ we complete the proof when $3cT<1$. □

Remark 4.1.

(1) We remark that the constant c in (4.1) depends on

$$\|D_p\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^s},\|\partial_q\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^s},\|mD^2_{pp}\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^s}$$
$$\text{and }\|mD_p\partial_q\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^s}.$$

(2) The time horizon appearing in the previous theorem satisfies $T_{**}<\frac{1}{3c\kappa}$.

(3) In fact, we proved that

$$\max\left\{\sup_{t\in[0,T]}\|v(t,\cdot)\|_{H^{-s}};\ \sup_{t\in[0,T]}\|\mu\|_{H^{-s-1}};\ \int_0^T\|\nabla\mu\|^2_{H^{-s-1}}d\tau;\ \int_0^T\|\nabla v\|^2_{H^{-s}}\,d\tau\right\} \leq c\|\mu_0\|_{H^{-s-1}},$$

which is a stronger conclusion than (4.1).

Remark 4.2. We remark that in the proof of Theorem 4.1 we have only used that $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$ (i.e. we do not require for these estimates $\mu_0 \in H^{s-1}(\mathbb{T}^d)$) and the regularity of solutions (m,u) on the mean field game system (2.2). Therefore, we can conclude that in fact Theorem 4.1 (combined with the techniques of the proof of Theorem 3.1) provides an existence and uniqueness result for (2.4) for initial data $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$.

Corollary 4.2. *Under the assumptions of Theorem 3.1 and $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$ (that replaces the assumption $\mu_0 \in H^{s-1}(\mathbb{T}^d)$), there exists $T_{**} > 0$ and a unique distributional solution (v,μ) to (2.4) such that if $T \in (0, T_{**})$ then*

$$(v,\mu) \in L^\infty([0,T]; H^{-s}(\mathbb{T}^d)) \times L^\infty([0,T]; H^{-s-1}(\mathbb{T}^d)).$$

The proof of this result is straightforward, since the system (2.4) is linear with smooth coefficients (and it follows from the estimates in the proof of Theorem 4.1), so we omit it.

The estimate in Theorem 4.1 yields the following result.

Lemma 4.3. *The v component of the solution to the linearized system (2.4) (given by Theorem 2.4) can be represented as*

$$v(t,x) = \langle \mu_0, K(t,x,m_0,\cdot)\rangle_{H^{-s},H^s},\ \forall (t,x) \in [0,T_*)\times\mathbb{T}^d, \tag{4.15}$$

for a unique $K : [0,T_*)\times\mathbb{T}^d\times\mathscr{Q}_R\times\mathbb{T}^d \to \mathbb{R}$. In particular $K(t,x,m_0,\cdot) \in H^s(\mathbb{T}^d)$.

Proof. Let us denote by $T : H^{-s}(\mathbb{T}^d) \to H^{-s}(\mathbb{T}^d)$ the bounded linear operator such that $T(\mu_0) = v(t,\cdot)$ and let $T' : H^s(\mathbb{T}^d) \to H^s(\mathbb{T}^d)$ stand for its adjoint. Let $Q : H^{-s}(\mathbb{T}^d) \to H^s(\mathbb{T}^d)$ canonically, so that $Q^{-1} : H^s \to H^{-s}$ canonically. (The operator Q can be expressed simply in terms of the operator Λ we have introduced previously; since $f \in H^{-s}$ if and only if $\Lambda^{-s}f \in L^2$, and similarly for H^s, we may define $Qf = \Lambda^{-2s}f$ for any $f \in H^{-s}$. Then $\Lambda^s Qf = \Lambda^{-s}f \in L^2$, and we see $Qf \in H^s$. The inverse is similarly given by $Q^{-1} = \Lambda^{2s}$.)

For any $\mu_0 \in H^{-s}(\mathbb{T}^d)$, we have $QT\mu_0 \in H^s$. Since $s > \frac{d}{2}$, if we let x be given, then $\mu_0 \mapsto (QT\mu_0)(x)$ is a bounded linear functional on H^{-s}. Thus this bounded linear functional can be represented as the L^2 inner product with an element of H^s. We write this as

$$(QT\mu_0)(x) = \langle \mu_0(y), K(x,y)\rangle_{(H^{-s},H^s)_y},$$

where we have established $K(x,y) \in H^s_y(\mathbb{T}^d)$ for any x.

If we fix y_0 and let μ_0 be the Dirac mass supported at y_0, then we have

$$(QT\mu_0)(x) = K(x,y_0).$$

We know that $QT\mu_0 \in H^s_x(\mathbb{T}^d)$, so we have $K(x,y_0) \in H^s_x(\mathbb{T}^d)$ for all y_0.

We let $\tilde{K} = Q_x^{-1} K$, and we see that $\tilde{K}(\cdot, y) \in H_x^{-s}$ for all y and $\tilde{K}(x, \cdot) \in H_y^s$ for all x. So far, we have the following:

$$\begin{aligned}\langle T\mu_0(x), h(x)\rangle_{(H^{-s},H^s)_x} &= \langle Q_x^{-1}\langle \mu_0(y), K(x,y)\rangle_{(H^{-s},H^s)_y}, h(x)\rangle_{(H^{-s},H^s)_x} \\ &= \langle \langle \mu_0(y), Q_x^{-1}K(x,y)\rangle_{(H^{-s},H^s)_y}, h(x)\rangle_{(H^{-s},H^s)_x} \\ &= \langle \langle \mu_0(y), \tilde{K}(x,y)\rangle_{(H^{-s},H^s)_y}, h(x)\rangle_{(H^{-s},H^s)_x}.\end{aligned}$$

We can then write this as

$$\langle T\mu_0(x), h(x)\rangle_{(H^{-s},H^s)_x} = \int\int \mu_0(y)\tilde{K}(x,y)h(x)\, dydx.$$

Considering the subset of $H^{-s}(\mathbb{T}^d) \times H^s(\mathbb{T}^d)$ of pairs of functions (μ_0, h) each of which has exponentially decaying Fourier series, we see that we may change the order of integration, finding

$$\begin{aligned}\langle T\mu_0(x), h(x)\rangle_{(H^{-s},H^s)_x} &= \int\int \mu_0(y)\tilde{K}(x,y)h(x)\, dxdy \\ &= \langle \mu_0(y), \langle \tilde{K}(x,y), h(x)\rangle_{(H^{-s},H^s)_x}\rangle_{(H^{-s},H^s)_y} = \langle \mu_0(x), T'h(x)\rangle_{(H^{-s},H^s)_x}.\end{aligned} \tag{4.16}$$

Since (4.16) holds on a dense subset, this representation holds for all of $H^{-s}(\mathbb{T}^d) \times H^s(\mathbb{T}^d)$. The thesis of the lemma follows by setting $K(t, x, m_0, \cdot) := \tilde{K}(x, \cdot)$. □

5. The first variation of the master function

In addition to the problem (2.7) written for z, we need to develop the corresponding problem for $\nu := \tilde{m} - m - \mu$. Note that the initial data for ν is a signed measure that has zero total mass, but is not equal to zero. An equation for ν is needed, since the PDE in (2.7) involves the term $\partial_q \mathcal{H}\mu$, to which we need to add and subtract $\tilde{m} - m$. One of the resulting terms will help to complete the Taylor expansion of $\mathcal{H}$, but the other resulting term will include ν, which then must be estimated.

We take (2.7) and perform the mentioned adding and subtracting:

$$\begin{aligned}-\partial_t z - \Delta z + D_p\mathcal{H}(t,x,\nabla u, m)\cdot\nabla z = &-\mathcal{H}(t,x,\nabla\tilde{u},\tilde{m}) + \mathcal{H}(t,x,\nabla u, m) \\ &+ D_p\mathcal{H}(t,x,\nabla u,m)\cdot\nabla(\tilde{u}-u) \\ &- \partial_q\mathcal{H}(t,x,\nabla u,m)(\tilde{m}-m-\mu) \\ &+ \partial_q\mathcal{H}(t,x,\nabla u,m)(\tilde{m}-m).\end{aligned}$$

We rearrange this to put the term with ν on the left-hand side:

$$\begin{aligned}&-\partial_t z - \Delta z + D_p\mathcal{H}(t,x,\nabla u,m)\cdot\nabla z + \partial_q\mathcal{H}(t,x,\nabla u,m)\nu \\ &\quad = -\mathcal{H}(t,x,\nabla\tilde{u},\tilde{m}) + \mathcal{H}(t,x,\nabla u,m) \\ &\qquad + D_p\mathcal{H}(t,x,\nabla u,m)\cdot\nabla(\tilde{u}-u) + \partial_q\mathcal{H}(t,x,\nabla u,m)(\tilde{m}-m).\end{aligned} \tag{5.1}$$

We next need to form the ν equation. Clearly we have $\partial_t\nu = \partial_t\tilde{m} - \partial_t m - \partial_t\mu$, and we use (2.2) for the $\tilde{m}$ and m equations, and we use (2.4) for the μ equation. For brevity we will write $D_p\mathcal{H} = D_p\mathcal{H}(t,x,\nabla u,m)$ and $\widetilde{D_p\mathcal{H}} = D_p\mathcal{H}(t,x,\nabla\tilde{u},\tilde{m})$, and similarly for $D^2_{pp}\mathcal{H}$ and $D_p\partial_q\mathcal{H}$. To begin, just substituting for $\partial_t\tilde{m}$, $\partial_t m$, and $\partial_t\mu$, we have

$$\begin{aligned}\partial_t\nu = \Delta\nu + \operatorname{div}(\tilde{m}\widetilde{D_p\mathcal{H}}) - \operatorname{div}(mD_p\mathcal{H}) - \operatorname{div}(\mu D_p\mathcal{H}) - \operatorname{div}(mD^2_{pp}\mathcal{H}\nabla v) \\ + \operatorname{div}(mD_p\partial_q\mathcal{H}\mu).\end{aligned}$$

We next add and subtract a few times; for each of the two times on the right-hand side that μ appears, we add and subtract $\tilde{m} - m$ to make ν appear, and in the one instance that v appears on the right-hand side, we add and subtract $\tilde{u} - u$ so as to bring out z. These considerations yield the following:

$$\partial_t \nu = \Delta \nu - \operatorname{div}(\tilde{m}\widetilde{D_p \mathcal{H}}) + \operatorname{div}(m D_p \mathcal{H}) + \operatorname{div}(\nu D_p \mathcal{H}) - \operatorname{div}((\tilde{m} - m) D_p \mathcal{H}) \tag{5.2}$$
$$+\operatorname{div}(m D^2_{pp} \mathcal{H} \nabla z) - \operatorname{div}(m D^2_{pp} \mathcal{H} \nabla (\tilde{u} - u)) - \operatorname{div}(m D_p \partial_q \mathcal{H} \nu) + \operatorname{div}(m D_p \partial_q \mathcal{H} (\tilde{m} - m)).$$

We restate the equations (5.1), (5.2) so that the terms which do not explicitly involve z or ν are rephrased as forcing terms,

$$-\partial_t z - \Delta z + D_p \mathcal{H}(t, x, \nabla u, m) \cdot \nabla z + \partial_q \mathcal{H}(t, x, \nabla u, m) \nu + F_1 = 0, \tag{5.3}$$

$$\partial_t \nu = \Delta \nu + \operatorname{div}(\nu D_p \mathcal{H}) + \operatorname{div}(m D^2_{pp} \mathcal{H} \nabla z) - \operatorname{div}(m D_p \partial_q \mathcal{H} \nu) + F_2. \tag{5.4}$$

Here, we recall the formulas for F_1 and F_2 from (H3) and (H4).

Theorem 5.1. *Let s satisfy (H1) and let r be given as in (2.1). Let $R > 0$. Let $(u, m), (\tilde{u}, \tilde{m})$ be solutions to (2.2), with initial data $m_0, \tilde{m}_0 \in \mathcal{Q}_R$, respectively, given by Theorem 2.2. Let (v, ν) be the solution to (2.4) with initial data $\mu_0 := \tilde{m}_0 - m_0$, given by Theorem 2.4. There exists $c > 0$ and $T_{***} > 0$ such that if $T \in (0, T_{***})$, then*

$$\sup_{t \in [0,T]} \|z\|^2_{H^r} \leq c \|m_0 - \tilde{m}_0\|^{5/2}_{H^{-1}}. \tag{5.5}$$

Proof. We will make energy estimates for $z \in H^r(\mathbb{T}^d)$ and $\nu \in H^{r-1}(\mathbb{T}^d)$, for the given r. We therefore define

$$E_z = \frac{1}{2} \int_{\mathbb{T}^d} (\Lambda^r z)^2 \, dx,$$

$$E_\nu = \frac{1}{2} \int_{\mathbb{T}^d} (\Lambda^{r-1} \nu)^2 \, dx.$$

Let us notice that by the assumptions on s in (H1) and the choice of r in (2.1), $H^r(\mathbb{T}^d)$ is an algebra.

Then we have

$$\frac{dE_z}{dt} = \int_{\mathbb{T}^d} (\Lambda^r z)(\Lambda^r \partial_t z) \, dx.$$

We substitute from (5.3), to get

$$\frac{dE_z}{dt} = -\int_{\mathbb{T}^d} (\Lambda^r z)(\Lambda^r \Delta z) \, dx + \int_{\mathbb{T}^d} (\Lambda^r z)(\Lambda^r (D_p \mathcal{H} \cdot \nabla z)) \, dx$$
$$+ \int_{\mathbb{T}^d} (\Lambda^r z)(\Lambda^r (\partial_q \mathcal{H} \nu)) \, dx + \int_{\mathbb{T}^d} (\Lambda^r z)(\Lambda^r F_1) \, dx = I_z + II_z + III_z + IV_z.$$

We integrate in time from time t to time T to find

$$E_z(t) = E_z(T) - \int_t^T I_z \, d\tau - \int_t^T II_z \, d\tau - \int_t^T III_z \, d\tau - \int_t^T IV_z \, d\tau. \tag{5.6}$$

For the term involving I_z, we integrate by parts with respect to the spatial variable,

$$-\int_t^T I_z \, d\tau = -\int_t^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau. \tag{5.7}$$

For the term involving II_z, we use Young's inequality with parameter σ_1,

$$-\int_t^T II_z \, d\tau \leq \int_t^T \frac{1}{2\sigma_1}\|z\|_{H^r}^2 \, d\tau + \int_t^T \frac{\sigma_1}{2}\|(D_p\mathcal{H})\nabla z\|_{H^r}^2 \, d\tau.$$

We continue by bounding $D_p\mathcal{H}$ by a constant, since this depends only on the solution (u,m) of the original problem. We also bound the integrals on the right-hand side, using a supremum in time for the first of these, and taking a larger domain of integration for the second of these:

$$-\int_t^T II_z \, d\tau \leq \frac{T}{2\sigma_1}\sup_{t\in[0,T]}\|z\|_{H^r}^2 + \frac{c\sigma_1}{2}\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2 \, dxd\tau. \tag{5.8}$$

We treat III_z similarly, applying Young's inequality with positive parameter σ_2,

$$-\int_t^T III_z \, d\tau \leq \int_t^T \frac{1}{2\sigma_2}\|z\|_{H^r}^2 \, d\tau + \int_t^T \frac{\sigma_2}{2}\|(\partial_q\mathcal{H})\nu\|_{H^r}^2 \, d\tau.$$

Again, continuing as we have for the previous term, we find

$$-\int_t^T III_z \, d\tau \leq \frac{T}{2\sigma_2}\sup_{t\in[0,T]}\|z\|_{H^r}^2 + \frac{c\sigma_2}{2}\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nu)^2 \, dxd\tau.$$

We next wish to change the form of the final integral on the right-hand side, to have $(\Lambda^{r-1}\nabla\nu)^2$ rather than $(\Lambda^r\nu)^2$. Using the inequality

$$\int_{\mathbb{T}^d}(\Lambda^r\nu)^2dx \leq c\left(\|\nu\|_{H^{r-1}} + \int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2dx\right)$$

we obtain the bound

$$\begin{aligned}-\int_t^T III_z \, d\tau \leq{}& \frac{T}{2\sigma_2}\sup_{t\in[0,T]}\|z\|_{H^r}^2 + c\sigma_2 T\sup_{t\in[0,T]}\|\nu\|_{H^{r-1}}^2 \\ &+ c\sigma_2\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2 \, dxd\tau.\end{aligned} \tag{5.9}$$

For the term involving IV_z, we use Young's inequality without parameter,

$$-\int_t^T IV_z \, d\tau \leq \frac{T}{2}\sup_{t\in[0,T]}\|z\|_{H^r}^2 + \frac{T}{2}\sup_{t\in[0,T]}\|F_1\|_{H^r}^2.$$

We then use (H3), finding

$$-\int_t^T IV_z \, d\tau \leq \frac{T}{2}\sup_{t\in[0,T]}\|z\|_{H^r}^2 + c\sup_{t\in[0,T]}(\|m-\tilde{m}\|_{H^r}^2 + \|u-\tilde{u}\|_{H^{r+1}}^2). \tag{5.10}$$

We combine (5.6), using the definition of E_z, with (5.7), (5.8), (5.9), and (5.10). This yields

$$\begin{aligned}&\frac{1}{2}\|z(t,\cdot)\|_{H^r}^2 + \int_t^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2 \, dxd\tau \\ &\quad\leq \frac{1}{2}\|z(T,\cdot)\|_{H^r}^2 + \frac{T}{2}\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}+1\right)\sup_{t\in[0,T]}\|z\|_{H^r}^2 + c\sigma_2 T\sup_{t\in[0,T]}\|\nu\|_{H^{r-1}}^2 \\ &\qquad + c\sigma_1\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2 \, dxd\tau + c\sigma_2\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2 \, dxd\tau \\ &\qquad\qquad + c\sup_{t\in[0,T]}\left(\|m-\tilde{m}\|_{H^r}^2 + \|u-\tilde{u}\|_{H^{r+1}}^2\right).\end{aligned}$$

Since by (H7) the data $z(T,\cdot)$ is also quadratic in $m-\tilde{m}$, this simplifies slightly as

$$\begin{aligned}\frac{1}{2}\|z(t,\cdot)\|^2_{H^r} + \int_t^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau \\ \leq \frac{T}{2}\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}+1\right) \sup_{t\in[0,T]} \|z\|^2_{H^r} + c\sigma_2 T \sup_{t\in[0,T]} \|\nu\|^2_{H^{r-1}} \\ + c\sigma_1 \int_0^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau + c\sigma_2 \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla \nu)^2 \, dx d\tau \\ + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|^2_{H^r} + \|u-\tilde{u}\|^2_{H^{r+1}}\right).\end{aligned}$$

We then treat the terms on the left-hand side separately, and take the supremum with respect to time, finding

$$\begin{aligned}\frac{1}{2}\sup_{t\in[0,T]} \|z\|^2_{H^r} \leq \frac{T}{2}\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}+1\right) \sup_{t\in[0,T]} \|z\|^2_{H^r} + c\sigma_2 T \sup_{t\in[0,T]} \|\nu\|^2_{H^{r-1}} \\ + c\sigma_1 \int_0^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau + c\sigma_2 \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla \nu)^2 \, dx d\tau \\ + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|^2_{H^r} + \|u-\tilde{u}\|^2_{H^{r+1}}\right),\end{aligned} \tag{5.11}$$

as well as

$$\begin{aligned}\int_0^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau \leq \frac{T}{2}\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}+1\right) \sup_{t\in[0,T]} \|z\|^2_{H^r} + c\sigma_2 T \sup_{t\in[0,T]} \|\nu\|^2_{H^{r-1}} \\ + c\sigma_1 \int_0^T \int_{\mathbb{T}^d} (\Lambda^r \nabla z)^2 \, dx d\tau + c\sigma_2 \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla \nu)^2 \, dx d\tau \\ + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|^2_{H^r} + \|u-\tilde{u}\|^2_{H^{r+1}}\right).\end{aligned} \tag{5.12}$$

We next write out the time derivative of E_ν,

$$\frac{dE_\nu}{dt} = \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}\partial_t \nu) \, dx.$$

We substitute from (5.4) to find

$$\begin{aligned}\frac{dE_\nu}{dt} = \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}\Delta\nu) \, dx + \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}\mathrm{div}(\nu D_p \mathcal{H})) \, dx \\ + \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}\mathrm{div}(mD^2_{pp}\mathcal{H}\nabla z)) \, dx - \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}\mathrm{div}(mD_p\partial_q \mathcal{H}\nu)) \, dx \\ + \int_{\mathbb{T}^d} (\Lambda^{r-1}\nu)(\Lambda^{r-1}F_2) \, dx = I_\nu + II_\nu + III_\nu + IV_\nu + V_\nu.\end{aligned}$$

We integrate in time from time zero until time T, and we use $E_\nu(0)=0$, finding

$$E_\nu(t) = \int_0^t I_\nu \, d\tau + \int_0^t II_\nu \, d\tau + \int_0^t III_\nu \, d\tau + \int_0^t IV_\nu \, d\tau + \int_0^t V_\nu \, d\tau. \tag{5.13}$$

For the term involving I_ν, we integrate by parts with respect to the spatial variable,

$$\int_0^t I_\nu \, d\tau = -\int_0^t \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau. \tag{5.14}$$

For the term involving II_ν, we use Young's inequality with paramter σ_3, and we proceed as we did above for the terms involving II_z and III_z,

$$\int_0^t |II_\nu| \, d\tau \le \frac{T}{2\sigma_3} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 + \frac{c\sigma_3}{2} \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau. \tag{5.15}$$

For the term involving III_ν, we use Young's inequality with parameter σ_4,

$$\int_0^t |III_\nu| \, d\tau \le \frac{T}{2\sigma_4} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 + \frac{c\sigma_4}{2} \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r}\nabla z)^2 \, dx d\tau. \tag{5.16}$$

For the term involving IV_ν we continue in the same manner, using Young's inequality with positive parameter σ_5,

$$\int_0^t |IV_\nu| \, d\tau \le \frac{T}{2\sigma_5} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 + \frac{c\sigma_5}{2} \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau. \tag{5.17}$$

Finally, we estimate the term involving V_ν in the same manner we estimated the above term involving IV_z,

$$\int_0^t |V_\nu| \, d\tau \le \frac{T}{2} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 + \frac{T}{2} \sup_{t\in[0,T]} \|F_2\|_{H^{r-1}}^2.$$

We use (H4) with this, yielding

$$\int_0^t |V_\nu| \, d\tau \le \frac{T}{2} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|_{H^{r-1}}^4 + \|u-\tilde{u}\|_{H^r}^4 \right). \tag{5.18}$$

We use the definition of E_ν in (5.13), and we substitute from (5.14), (5.15), (5.16), (5.17), and (5.18). This yields

$$\begin{aligned}
&\frac{1}{2}\|\nu(t,\cdot)\|_{H^{r-1}}^2 + \int_0^t \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau \\
&\qquad \le \frac{T}{2}\left(\frac{1}{\sigma_3}+\frac{1}{\sigma_4}+\frac{1}{\sigma_5}+1\right) \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 \\
&\quad + c(\sigma_3+\sigma_5) \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau + c\sigma_4 \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r}\nabla z)^2 \, dx d\tau \\
&\qquad + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|_{H^{r-1}}^4 + \|u-\tilde{u}\|_{H^r}^4 \right).
\end{aligned}$$

Treating the terms on the left-hand side separately, and taking the supremum with respect to time, we thus find

$$\begin{aligned}
&\frac{1}{2} \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 \le \frac{T}{2}\left(\frac{1}{\sigma_3}+\frac{1}{\sigma_4}+\frac{1}{\sigma_5}+1\right) \sup_{t\in[0,T]} \|\nu\|_{H^{r-1}}^2 \\
&\quad + c(\sigma_3+\sigma_5) \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r-1}\nabla\nu)^2 \, dx d\tau + c\sigma_4 \int_0^T \int_{\mathbb{T}^d} (\Lambda^{r}\nabla z)^2 \, dx d\tau \\
&\qquad + c \sup_{t\in[0,T]} \left(\|m-\tilde{m}\|_{H^{r-1}}^4 + \|u-\tilde{u}\|_{H^r}^4 \right),
\end{aligned} \tag{5.19}$$

and also

$$\int_0^T \int_{\mathbb{T}}^d (\Lambda^{r-1}\nabla\nu)^2 \, dxd\tau \le \frac{T}{2}\left(\frac{1}{\sigma_3}+\frac{1}{\sigma_4}+\frac{1}{\sigma_5}+1\right)\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}} + c(\sigma_3+\sigma_5)\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2\,dxd\tau + c\sigma_4\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2\,dxd\tau + c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^4_{H^{r-1}}+\|u-\tilde u\|^4_{H^r}\right). \tag{5.20}$$

We add (5.12) and (5.20), finding

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2\,dxd\tau+\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2\,dxd\tau\\ &\le\frac{T}{2}\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}+1\right)\sup_{t\in[0,T]}\|z\|^2_{H^r}\\ &\quad+\left(\frac{T}{2\sigma_3}+\frac{T}{2\sigma_4}+\frac{T}{2\sigma_5}+\frac{T}{2}+c\sigma_2T\right)\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}}\\ &\quad+c(\sigma_1+\sigma_4)\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2\,dxd\tau+c(\sigma_2+\sigma_3+\sigma_5)\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2\,dxd\tau\\ &\quad+c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^4_{H^r}+\|u-\tilde u\|^4_{H^{r+1}}\right).\end{aligned}$$

We choose the constants so that

$$c(\sigma_1+\sigma_4)=\frac{1}{2},\qquad c(\sigma_2+\sigma_3+\sigma_5)=\frac{1}{2}.$$

We then conclude

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d}(\Lambda^r\nabla z)^2\,dxd\tau+\int_0^T\int_{\mathbb{T}^d}(\Lambda^{r-1}\nabla\nu)^2\,dxd\tau\\ &\le cT\sup_{t\in[0,T]}\|z\|^2_{H^r}+cT\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}}+c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^4_{H^r}+\|u-\tilde u\|^4_{H^{r+1}}\right).\end{aligned} \tag{5.21}$$

We next use the bound (5.21) on the right-hand sides of (5.11) and (5.19), finding

$$\sup_{t\in[0,T]}\|z\|^2_{H^r}+\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}}\le cT\sup_{t\in[0,T]}\|z\|^2_{H^r}+cT\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}} + c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^4_{H^r}+\|u-\tilde u\|^4_{H^{r+1}}\right).$$

Taking T sufficiently small, we have

$$\sup_{t\in[0,T]}\|z\|^2_{H^r}+\sup_{t\in[0,T]}\|\nu\|^2_{H^{r-1}}\le c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^4_{H^r}+\|u-\tilde u\|^4_{H^{r+1}}\right).$$

Neglecting the ν-term on the left-hand side, this establishes

$$\sup_{t\in[0,T]}\|z\|^2_{H^r}\le c\sup_{t\in[0,T]}\left(\|m-\tilde m\|^2_{H^r}+\|u-\tilde u\|^2_{H^{r+1}}\right). \tag{5.22}$$

We will use the Sobolev interpolation inequality (2.16), and we will do this in two steps. First we take $\alpha=r$ and $\beta=4r$. Then we have

$$\|f\|^4_{H^r}\le c\|f\|^3_{L^2}\|f\|_{H^{4r}}. \tag{5.23}$$

We do this because, Theorem A.1 yields

$$\sup_{t\in[0,T]} \left(\|m-\tilde{m}\|_{L^2}^4 + \|u-\tilde{u}\|_{H^1}^4\right) \leq c\|m_0-\tilde{m}_0\|_{L^2}^4. \tag{5.24}$$

Although this estimate is the consequence of results from [4], for completeness, we provide its proof in Appendix A. So, combining (5.22) with (5.23) and (5.24), we have

$$\sup_{t\in[0,T]} \|z\|_{H^r}^2 \leq c\left(\sup_{t\in[0,T]} \|m-\tilde{m}\|_{L^2}^3 + \|u-\tilde{u}\|_{H^1}^3\right) \leq c\|m_0-\tilde{m}_0\|_{L^2}^3,$$

where we have used that m and $\tilde{m}$ remain bounded in H^{4r}, and u and $\tilde{u}$ remain bounded in H^{4r+1}, throughout the time interval $[0,T]$. We then make another use of (2.16), with $\alpha=1$ and $\beta=6$,

$$\|f\|_{L^2}^3 = \|\Lambda^{-1}f\|_{H^1}^3 \leq \|\Lambda^{-1}f\|_{L^2}^{5/2}\|\Lambda^{-1}f\|_{H^6}^{1/2} = c\|f\|_{H^{-1}}^{5/2}\|f\|_{H^5}^{1/2}.$$

With m_0 and $\tilde{m}_0$ bounded in H^5, this implies

$$\sup_{t\in[0,T]} \|z\|_{H^r}^2 \leq c\|m_0-\tilde{m}_0\|_{H^{-1}}^{5/2}.$$

As desired, we have established $z = o(\|m_0-\tilde{m}_0\|_{H^{-1}})$. This completes the proof. □

Remark 5.1.

(1) We notice that the constant c in the estimate (5.5) is depending on

$$\|D_p\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^r}, \|\partial_q\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^r}, \|mD_{pp}^2\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^r},$$
$$\|mD_p\partial_q\mathcal{H}(t,\cdot,\nabla u,m)\|_{H^r}, \|m\|_{H^{4r}}, \|\tilde{m}\|_{H^{4r}}, \|u\|_{H^{4r+1}}, \|\tilde{u}\|_{H^{4r+1}}$$

and on the constants from the assumptions (H3)–(H4). These last constants depend in particular on

$$\|D^2\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^r} \quad \text{and} \quad \|mD^3\mathcal{H}(t,\cdot,m,\nabla u)\|_{H^r}.$$

(2) The time horizon appearing in Theorem 5.1 satisfies $T_{***} < \frac{1}{c}$.

5.1. **Regularity of the master function.** We show now how the results from Sections 3, 4 and 5 will yield the necessary regularity of the master functions U.

Let us recall our standing assumption (H1). We recall r from (2.1) which in particular satisfies $r > \lceil \frac{d}{2} \rceil$.

Theorem 5.2. *Let $R>0$ and let $U:[0,T]\times\mathbb{T}^d\times\mathscr{Q}_R\to\mathbb{R}$ be defined in (2.3). Then*

(i) *$U(\cdot,\cdot,m_0)\in L^\infty([0,T];H^s(\mathbb{T}^d))\cap C([0,T];C^{2,\alpha}(\mathbb{T}^d))\cap C^1([0,T]\times\mathbb{T}^d)$ for all $m_0\in\mathscr{Q}_R$ and $t\in[0,T]$.*

(ii) *For all $(t,x)\in[0,T]\times\mathbb{T}^d$, $y\in\mathbb{T}^d$ and $m_0\in\mathscr{Q}_R$, the function $\frac{\delta U}{\delta m}(t,x,m_0)(y)$ is well defined. Moreover, $\frac{\delta U}{\delta m}(t,x,m_0)(\cdot)$ defines a continuous linear operator on $H^{-s}(\mathbb{T}^d)$.*

(iii) *The map $m\mapsto U(t,x,m)$ is Gâteaux differentiable in $H^s(\mathbb{T}^d)$, uniformly with respect to $t,x\in[0,T]\times\mathbb{T}^d$. In particular, there exists $C>0$ such that*

$$\sup_{t\in[0,T],x\in\mathbb{T}^d}\left|U(t,x,\tilde{m}_0)-U(t,x,m_0)-\int_{\mathbb{T}^d}\frac{\delta U}{\delta m}(t,x,m_0)(y)\,d(\tilde{m}_0-m_0)(y)\right|$$
$$\leq C\|m_0-\tilde{m}_0\|_{H^s}^{\frac{5}{4}},$$

for all $\tilde{m}_0, m_0 \in \mathscr{Q}_R$.

(iv) $H^s(\mathbb{T}^d) \ni m \mapsto \frac{\delta U}{\delta m}(t,\cdot,m)(y) \in H^{-s}(\mathbb{T}^d)$ *is continuous uniformly with respect to* $(t,y) \in [0,T]\times\mathbb{T}^d$.

(v) *The functions* $\mathbb{T}^d \ni y \mapsto \frac{\delta U}{\delta m}(t,\cdot,m_0)(y) \in H^{-s}(\mathbb{T}^d)$, $\mathbb{T}^d \ni y \mapsto \nabla_y \frac{\delta U}{\delta m}(t,\cdot,m_0)(y) \in H^{-s}(\mathbb{T}^d)$ *and* $\mathbb{T}^d \ni y \mapsto D^2_{yy}\frac{\delta U}{\delta m}(t,\cdot,m_0)(y) \in H^{-s}(\mathbb{T}^d)$ *are Lipschitz continuous, for any* $(t,m_0)\in[0,T]\times\mathscr{Q}_R$.

Proof. (i) By the definition of U in (2.3), for each fixed $m_0 \in \mathscr{Q}_R$, $U(\cdot,\cdot,m_0)$ inherits the regularity properties of u. The result follows from Corollary 2.3.

(ii–iii) Let us recall K from the Riesz type representation, provided in Lemma 4.3.

By (2.6), (2.3) and the estimates in Theorem 5.1, for $m_0, \tilde{m}_0 \in \mathscr{Q}_R$, we see that there exists a constant $C>0$ such that

$$\sup_{t\in[0,T]} \Big\| U(t,\cdot,\tilde{m}_0) - U(t,\cdot,m_0) - \int_{\mathbb{T}^d} K(t,\cdot,m_0)(y)\, d(\tilde{m}_0 - m_0)(y) \Big\|_{H^r} \leq C\|m_0-\tilde{m}_0\|^{\frac{5}{4}}_{H^{-1}}.$$

If $r > \lceil \frac{d}{2} \rceil$, we have in addition

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{T}^d} \Big| U(t,x,\tilde{m}_0) - U(t,x,m_0) - \int_{\mathbb{T}^d} K(t,x,m_0)(y)\, d(\tilde{m}_0 - m_0)(y) \Big| \leq C\|m_0-\tilde{m}_0\|^{\frac{5}{4}}_{H^{-1}} \leq C\|m_0-\tilde{m}_0\|^{\frac{5}{4}}_{H^{s}}.$$

By the regularity of K, this means that $U(t,x,\cdot)$ is Gâteaux differentiable in $H^{-s}(\mathbb{T}^d)\cap H^s(\mathbb{T}^d) = H^s(\mathbb{T}^d)$, uniformly with respect to $(t,x)\in[0,T]\times\mathbb{T}^d$.

Now, let us fix $y_0 \in \mathbb{T}^d$. Since $\mathscr{P}(\mathbb{T}^d) \hookrightarrow H^{-s-1}(\mathbb{T}^d)$, we have that $\mu_0 = \delta_{y_0} \in H^{-s-1}(\mathbb{T}^d)$. And so, by the previous representation one obtains

$$v(t,\cdot) = \int_{\mathbb{T}^d} K(t,x,m_0)(y)\, d\mu_0(y) = K(t,x,m_0)(y_0)$$

and so, we can set

$$\frac{\delta U}{\delta m}(t,x,m_0)(y_0) := K(t,x,m_0)(y_0),$$

and this expression is meaningful pointwise.

(iv) By Corollary 4.2 for $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$ we have that $v \in L^\infty([0,T];H^{-s}(\mathbb{T}^d))$. Therefore, (4.15) yields that

$$H^{-s}(\mathbb{T}^d) \ni v(t,\cdot) = \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t,\cdot,m_0)(y)\, d\mu_0(y),$$

for all $t\in[0,T]$. Since the solution (u,m) of (2.2) depends continuously on $m_0 \in \mathscr{Q}_R$, and so the coefficient functions in (2.4) depend also continuously on m_0, we have that for each fixed $\mu_0 \in H^{-s-1}(\mathbb{T}^d)$, $v(t,\cdot)$ depends continuously on $m_0\in\mathscr{Q}_R$, uniformly in $t\in[0,T]$. Now, let us fix $y_0\in\mathbb{T}^d$ and consider $\mu_0 = \delta_{y_0} \in H^{-s-1}(\mathbb{T}^d)$. Again, by the previous representation one has

$$v(t,\cdot) = \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t,\cdot,m_0)(y)\, d\mu_0(y) = \frac{\delta U}{\delta m}(t,\cdot,m_0)(y_0).$$

Thus, the previous reasonings imply that $H^s(\mathbb{T}^d)\ni m_0 \mapsto \frac{\delta U}{\delta m}(t,\cdot,m_0)(y_0)$ is continuous, uniformly with respect to $(t,y_0)\in[0,T]\times\mathbb{T}^d$, and so the claim follows.

(v) Now, let us define $v(\cdot,\cdot,y)$ to be the first component of the solution of (2.4) with initial condition δ_y. By linearity of the system (2.4), we find

$$v(t,x,y)-v(t,x,z)=\frac{\delta U}{\delta m}(t,x,m_0)(y)-\frac{\delta U}{\delta m}(t,x,m_0)(z).$$

Thus, (4.1) yields

$$\begin{aligned}\|v(t,\cdot,y)-v(t,\cdot,z)\|_{H^{-s}}&=\left\|\frac{\delta U}{\delta m}(t,\cdot,m_0)(y)-\frac{\delta U}{\delta m}(t,\cdot,m_0)(z)\right\|_{H^{-s}}\\&\le C\|\delta_y-\delta_z\|_{H^{-s-1}}\le C|y-z|.\end{aligned}$$

This shows the Lipschitz continuity of $\mathbb{T}^d\ni y\mapsto\frac{\delta U}{\delta m}(t,\cdot,m_0)(y)\in H^{-s}(\mathbb{T}^d)$.

Now, for $\varepsilon\in\mathbb{R}$ small and $e_i\in\mathbb{R}^d$ the i^{th} canonical basis element we set $y^i_\varepsilon:=y+\varepsilon e_i$. Again, by linearity of the system (2.4), we find

$$v(t,x,y^i_\varepsilon)-v(t,x,y)=\frac{\delta U}{\delta m}(t,x,m_0)(y^i_\varepsilon)-\frac{\delta U}{\delta m}(t,x,m_0)(y).$$

By dividing this identity by ε and taking the limit as $\varepsilon\to 0$, one obtains in the sense of distributions

$$\partial_{y_i}v(t,x,y)=\partial_{y_i}\frac{\delta U}{\delta m}(t,x,m_0)(y).$$

This last object also corresponds to

$$\left\langle\frac{\delta U}{\delta m}(t,x,m_0)(y),-\partial_{y_i}\delta_y\right\rangle_{H^s,H^{-s}},$$

where $\partial_{y_i}\delta_y$ stands for the distributional derivative of δ_y. This is the same as solving (2.4) with $\mu_0=-\partial_{y_i}\delta_y$.

Therefore, fixing $y,y'\in\mathbb{T}^d$, the linearity of the system (2.4) (4.1) yields

$$\left\|\nabla_y\frac{\delta U}{\delta m}(t,\cdot,m_0)(y)-\nabla_y\frac{\delta U}{\delta m}(t,\cdot,m_0)(y')\right\|_{H^{-s}}\le C\,\|D\delta_y-D\delta_{y'}\|_{H^{-s-1}}\le C|y-y'|.$$

In the last inequality we have relied on the following observation. Since $\widehat{D\delta_y}(k)=ik\widehat{\delta_y}(k)=\frac{ik}{(2\pi)^d}e^{-ik\cdot y}$, we have

$$\widehat{D\delta_y}(k)-\widehat{D\delta_{y'}}(k)=\frac{ik}{(2\pi)^d}(y-y')\cdot(-ik)\int_0^1 e^{-ik\cdot(sy+(1-s)y')}\,ds,$$

and so

$$\|D\delta_y-D\delta_{y'}\|^2_{H^{-s-1}}\le\frac{1}{(2\pi)^{2d}}|y-y'|^2\sum_{k\in\mathbb{Z}^d}\frac{|k|^4}{(1+|k|^2)^{s+1}}\le C|y-y'|^2,$$

for some $C>0$ depending only on d. Here, we also used the fact that s satisfies (H1).

So, by fixing $(t,m_0)\in[0,T]\times H^s(\mathbb{T}^d)$, we find that $y\mapsto\nabla_y\frac{\delta U}{\delta m}(t,\cdot,m_0)(y)\in H^{-s}(\mathbb{T}^d)$ is a Lipschitz continuous function. Using similar reasoning, the same conclusion can be made for $y\mapsto D^2_{yy}\frac{\delta U}{\delta m}(t,\cdot,m_0)(y)$. □

6. Well-posedness of the master equation

In this section we conclude by showing that the master equation (1.1) has a unique solution. This proof closely follows the proof of [11, Theorem 2.4.2].

Proof of Theorem 2.1. **Existence.**

Let $t_0 \in (0, T_{**})$ and $h \in (-1, 1)$, $h \neq 0$ such that $t_0 + h \in (0, T_{***})$. Let (u, m) be the solution to the MFG system (2.2) initiated at some $m_0 \in \mathcal{Q}_R$ at t_0 and let U be defined in (2.3). We have

$$\frac{U(t_0+h,x,m_0)-U(t_0,x,m_0)}{h} = \frac{U(t_0+h,x,m_0)-U(t_0+h,x,m(t_0+h))}{h} + \frac{U(t_0+h,x,m(t_0+h))-U(t_0,x,m_0)}{h}.$$

Let us denote $m_s := (1-s)m_0 + sm(t_0+h)$. Now, we can write

$$\begin{aligned}
&U(t_0+h,x,m(t_0+h)) - U(t_0+h,x,m_0)\\
&= \int_0^1 \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t_0+h,x,m_s)(y)(m(t_0+h,y)-m(t_0,y))\,dy\,ds\\
&= \int_0^1 \int_{\mathbb{T}^d} \int_{t_0}^{t_0+h} \frac{\delta U}{\delta m}(t_0+h,x,m_s)(y)\partial_t m(t,y)\,dt\,dy\,ds\\
&= \int_0^1 \int_{\mathbb{T}^d} \int_{t_0}^{t_0+h} \frac{\delta U}{\delta m}(t_0+h,x,m_s)(y)\,[\Delta m(t,y)\\
&\qquad +\nabla\cdot(m(t,y)D_p\mathcal{H}(t,y,Du(t,y),m(t,y)))]\,dt\,dy\,ds.
\end{aligned}$$

Now, by dividing this expression by h and taking the limit as $h \to 0$, we find that

$$\begin{aligned}
&\lim_{h\to 0} \frac{U(t_0+h,x,m(t_0+h))-U(t_0+h,x,m_0)}{h}\\
&= \int_0^1 \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t_0,x,m_{t_0})(y)\,[\Delta m(t_0,y)\\
&\quad +\nabla\cdot(m(t_0,y)D_p\mathcal{H}(t_0,y,Du(t_0,y),m(t_0,y)))]\,dy\,ds\\
&= \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t_0,x,m_{t_0})(y)\,[\Delta m(t_0,y)\\
&\quad +\nabla\cdot(m(t_0,y)D_p\mathcal{H}(t_0,y,Du(t_0,y),m(t_0,y)))]\,dy\\
&= \int_{\mathbb{T}^d} \nabla_y\cdot[\nabla_w U](t_0,x,m_0)(y)m_0(y)\,dy\\
&\quad - \int_{\mathbb{T}^d} \nabla_w U(t_0,x,m_0)(y)\cdot D_p\mathcal{H}(t_0,y,Du(t_0,y),m_0(y))m_0(y)\,dy.
\end{aligned}$$

In the first equality, we have used the continuity of the integrand (see Theorem 5.2(iv)), while in the last equality, we used the fact that $\frac{\delta U}{\delta m}(t_0,x,m_{t_0})(\cdot) \in H^s(\mathbb{T}^d)$ (see Theorem 5.2(ii)).

Second, by the time regularity of u (see Corollary 2.3), we have

$$U(t_0+h,x,m(t_0+h))-U(t_0,x,m_0) = u(t_0+h,x)-u(t_0,x) = h\partial_t u(t_0,x)+o(h).$$

Therefore, we have

$$\begin{aligned}\lim_{h\to 0}\frac{U(t_0+h,x,m_0)-U(t_0,x,m_0)}{h} &= \partial_t u(t_0,x)\\ &- \int_{\mathbb{T}^d}\nabla_y\cdot[\nabla_w U](t_0,x,m_0)(y)m_0(y)\,dy\\ &+ \int_{\mathbb{T}^d}\nabla_w U(t_0,x,m_0)(y)\cdot D_p\mathcal{H}(t_0,y,Du(t_0,y),m_0(y))m_0(y)\,dy\end{aligned}$$

And since u solves the first equation from (2.2), we get

$$\begin{aligned}&-\partial_t U(t_0,x,m_0)-\Delta U(t_0,x,m_0)+\mathcal{H}(t,x,\nabla U,m)\\ &-\int_{\mathbb{T}^d}\nabla_y\cdot[\nabla_w U](t_0,x,m_0)(y)m_0(y)\,dy\\ &+\int_{\mathbb{T}^d}\nabla_w U(t_0,x,m_0)(y)\cdot D_p\mathcal{H}(t_0,y,\nabla U(t_0,y),m_0(y))m_0(y)\,dy=0.\end{aligned}\tag{6.1}$$

Let us remark that for $(t_0,m_0)\in[0,T]\times H^s(\mathbb{T}^d)$ fixed, by Theorem 5.2, the mapping

$$\begin{aligned}x\mapsto &-\int_{\mathbb{T}^d}\nabla_y\cdot[\nabla_w U](t_0,x,m_0)(y)m_0(y)\,dy\\ &+\int_{\mathbb{T}^d}\nabla_w U(t_0,x,m_0)(y)\cdot D_p\mathcal{H}(t_0,y,\nabla U(t_0,y),m_0(y))m_0(y)\,dy\end{aligned}\tag{6.2}$$

has to be regarded a priori as an element of $H^{-s}(\mathbb{T}^d)$. However, by (6.1), since

$$x\mapsto -\partial_t U(t_0,x,m_0)-\Delta U(t_0,x,m_0)+\mathcal{H}(t,x,\nabla U,m)$$

is continuous, the mapping in (6.2) becomes also continuous. Therefore, U is indeed a solution to the master equation.

Uniqueness. Let us suppose now that $V:[0,T]\times\mathbb{T}^d\times\mathscr{Q}_R\to\mathbb{R}$ is another solution to the master equation. Let $m_0\in\mathscr{Q}_R$. By the regularity of V, decreasing the time horizon $T>0$ if necessary, one has that the problem

$$\begin{cases}\partial_t m-\Delta m-\nabla\cdot(mD_p\mathcal{H}(t,x,\nabla_x V(t,x,m),m))=0, & \text{in } (0,T)\times\mathbb{T}^d,\\ m(t_0,\cdot)=m_0, & \text{in } \mathbb{T}^d\end{cases}$$

has a classical solution.

Now, for all $(t,x)\in[t_0,T]\times\mathbb{T}^d$, set $\tilde{u}(t,x):=V(t,x,m_t)$. By the regularity of V and m the computations below are justified:

$$\begin{aligned}\partial_t u(t,x)&=\partial_t V(t,x,m_t)+\int_{\mathbb{T}^d}\frac{\delta V}{\delta m}(t,x,m_t)(y)\partial_t m_t(y)\,dy\\ &=\partial_t V(t,x,m_t)+\int_{\mathbb{T}^d}\operatorname{div}_y\nabla_w V(t,x,m_t)(y)m_t(y)\,dy\\ &-\int_{\mathbb{T}^d}\nabla_w V(t,x,m_t)(y)\cdot D_p\mathcal{H}(t,x,\nabla_x V,m_t)m_t(y)\,dy\\ &=\Delta_x V(t,x,m_t)-\mathcal{H}(t,x,\nabla_x V(t,x,m_t),m_t)\\ &=\Delta_x\tilde{u}(t,x)-\mathcal{H}(t,x,\nabla_x\tilde{u}(t,x),m_t),\end{aligned}$$

where in the penultimate equality we have used the fact that V solves the master equation, while in the last equality, we have used the definition of $\tilde{u}$. Since by

definition we also have $\tilde{u}(T,x) = G(x, m_T)$, we have that $(\tilde{u}, m)$ is a solution to the mean field game system. That system has a unique solution and therefore, we must have that $V(t_0, x, m_0) = U(t_0, x, m_0)$ for m_0 smooth. The uniqueness follows. □

Appendix A. A stability result

For a probability density m, if we write $m = \bar{m} + \mu$, with μ having zero mean (where $\bar{m}$ stands for the uniform measure on $\mathbb{T}^d$), then we introduce $\tilde{G}$ through $\tilde{G}(\mu(x), x) = PG(m(x), x)$, where P is the projection operator that removes the mean of a periodic function. In particular, we also have $Pm = \mu$.

Let us recall the Lipschitz continuity condition on G, assumed in (H5). For $m, \tilde{m}$ probability measures, note that $m - \tilde{m} = \mu - \tilde{\mu}$. We notice that (H5) immediately implies

$$\|\tilde{G}(\mu_1, \cdot) - \tilde{G}(\mu_2, \cdot)\|_1^2 \leq \Upsilon \|\mu_1 - \mu_2\|_0^2.$$

This follows because $\|Pf\|_{H^1}^2 \leq \|f\|_{H^1}^2$ for all f.

We are now able to state our stability theorem.

Theorem A.1. *Let $(u^1, \bar{m}+\mu^1)$ and $(u^2, \bar{m}+\mu^2)$ be two classical solutions of (2.2) on $(0,T) \times \mathbb{T}^d$, with data $m^1(0,\cdot) = m_0^1$, $m^2(0,\cdot) = m_0^2$. Assume that there exists K such that the solutions are each bounded by K :*

$$\|Du^i\|_{H^{s-1}} + \|\mu^i\|_{H^{s-1}} \leq K, \qquad i \in \{1,2\},$$

for some $s > 2 + \frac{d}{2}$ (which is ensured by Theorem 2.2 and the assumption (H1)). Assume that (H3) and (H5) hold. If T is sufficiently small, then there exists a constant $c > 0$ (depending on the data and K) such that

$$\sup_{t\in[0,T]} \|u^1 - u^2\|_{H^1}^2 + \|m^1 - m^2\|_{L^2}^2 \leq c\|m_0^1 - m_0^2\|_{L^2}^2. \tag{A.1}$$

Proof. We use the notation $w^i := Pu^i$ and we notice that $Dw^i = Du^i$. We also use the notation

$$\Theta(t, x, Dw, \mu) := -\mathcal{H}(t, x, Du, m).$$

Let us notice that (2.2) yields

$$\begin{cases} \partial_t w^i + \Delta w^i + P\Theta(t, x, Dw^i, \mu^i) = 0, & (t,x) \in (0,T) \times \mathbb{T}^d, \\ \partial_t \mu^i - \Delta \mu^i + \nabla \cdot (\mu^i D_p \Theta(t, x, Dw^i, \mu^i)) & \\ \quad + \bar{m} \nabla \cdot (D_p \Theta(t, x, Dw^i, \mu^i)) = 0, & (t,x) \in (0,T) \times \mathbb{T}^d, \\ \mu^i(t_0, x) = m_0(x) - \bar{m}, \quad w^i(\tilde{T}, x) = \tilde{G}(x, \mu_T^i(x)), & x \in \mathbb{T}^d. \end{cases} \tag{A.2}$$

We define an energy $E = E_\mu + E_w$, by

$$E_\mu = \frac{1}{2} \int_{\mathbb{T}^d} (\mu^1 - \mu^2)^2 \, dx,$$

$$E_w = \frac{1}{2} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2)^2 \, dx.$$

We will first bound E_μ and E_w. To finish the argument, we will then only need to bound the mean of $u_1 - u_2$.

To estimate the growth of this energy, we begin by taking the following time derivative:

$$\frac{dE_\mu}{dt} = \int_{\mathbb{T}^d} (\mu^1 - \mu^2)(\partial_t \mu^1 - \partial_t \mu^2) \, dx.$$

Substituting for μ_t^1 and μ_t^2 from (A.2), and then adding and subtracting, we arrive at

$$\begin{aligned}\frac{dE_\mu}{dt} &= \int_{\mathbb{T}^d} (\mu^1-\mu^2)\Delta(\mu^1-\mu^2)dx \\ &\quad - \int_{\mathbb{T}^d} (\mu^1-\mu^2)\mathrm{div}\left((\mu^1-\mu^2)D_p\Theta(t,x,Dw^1,\mu^1)\right)dx \\ &\quad - \int_{\mathbb{T}^d} (\mu^1-\mu^2)\mathrm{div}\left(\mu^2\left(D_p\Theta(t,x,Dw^1,\mu^1)-D_p\Theta(t,x,Dw^2,\mu^2)\right)\right)dx \\ &\quad - \bar{m}\int_{\mathbb{T}^d} (\mu^1-\mu^2)\mathrm{div}\left(D_p\Theta(t,x,Dw^1,\mu^1)-D_p\Theta(t,x,Dw^2,\mu^2)\right)dx.\end{aligned}$$

We expand the derivatives on the right-hand side, introducing the notation

$$\frac{dE_\mu}{dt} = \sum_{\ell=1}^{14} V_\ell,$$

where the summands are given by the following expressions:

$$V_1 = \int_{\mathbb{T}^d} (\mu^1-\mu^2)\Delta(\mu^1-\mu^2)\,dx,$$

$$V_2 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)\left(\nabla\left(\mu^1-\mu^2\right)\right)\cdot D_p\Theta(t,x,Dw^1,\mu^1)\,dx,$$

$$V_3 = -\int_{\mathbb{T}^d} \left(\mu^1-\mu^2\right)^2 \mathrm{div}\left(D_p\Theta(t,x,Dw^1,\mu^1)\right)\,dx,$$

$$V_4 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\nabla\mu^2)\cdot\left(D_p\Theta(t,x,Dw^1,\mu^1)-D_p\Theta(t,x,Dw^2,\mu^2)\right)\,dx,$$

$$V_5 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\mu^2)\sum_{i=1}^d\left[\Theta_{p_ix_i}(t,x,Dw^1,\mu^1)-\Theta_{p_ix_i}(t,x,Dw^2,\mu^2)\right]\,dx,$$

$$V_6 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\mu^2)\sum_{i=1}^d\left[\Theta_{p_iq}(t,x,Dw^1,\mu^1)\frac{\partial\mu^1}{\partial x_i}-\Theta_{p_iq}(t,x,Dw^2,\mu^2)\frac{\partial\mu^1}{\partial x_i}\right]\,dx,$$

$$V_7 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\mu^2)\sum_{i=1}^d\left[\Theta_{p_iq}(t,x,Dw^2,\mu^2)\left(\frac{\partial(\mu^1-\mu^2)}{\partial x_i}\right)\right]\,dx,$$

$$\begin{aligned}V_8 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\mu^2)\sum_{i=1}^d\sum_{j=1}^d\Bigg[&\Theta_{p_ip_j}(t,x,Dw^1,\mu^1)\frac{\partial^2 w^1}{\partial x_i\partial x_j} \\ &-\Theta_{p_ip_j}(t,x,Dw^2,\mu^2)\frac{\partial^2 w^1}{\partial x_i\partial x_j}\Bigg]\,dx,\end{aligned}$$

$$V_9 = -\int_{\mathbb{T}^d} (\mu^1-\mu^2)(\mu^2)\sum_{i=1}^d\sum_{j=1}^d\left[\Theta_{p_ip_j}(t,x,Dw^2,\mu^2)\left(\frac{\partial^2(w^1-w^2)}{\partial x_i\partial x_j}\right)\right]\,dx,$$

$$V_{10} = -\bar{m}\int_{\mathbb{T}^d} (\mu^1-\mu^2)\sum_{i=1}^d\left[\Theta_{p_ix_i}(t,x,Dw^1,\mu^1)-\Theta_{p_ix_i}(t,x,Dw^2,\mu^2)\right]\,dx,$$

$$V_{11} = -\bar{m}\int_{\mathbb{T}^d} (\mu^1-\mu^2)\sum_{i=1}^d\left[\Theta_{p_iq}(t,x,Dw^1,\mu^1)\frac{\partial\mu^1}{\partial x_i}-\Theta_{p_iq}(t,x,\mu^2,Dw^1)\frac{\partial\mu^1}{\partial x_i}\right]\,dx,$$

$$V_{12} = -\bar{m}\int_{\mathbb{T}^d}(\mu^1-\mu^2)\sum_{i=1}^{d}\Theta_{p_iq}(t,x,Dw^2,\mu^2)\left(\frac{\partial(\mu^1-\mu^2)}{\partial x_i}\right)\,dx,$$

$$V_{13} = -\bar{m}\int_{\mathbb{T}^d}(\mu^1-\mu^2)\sum_{i=1}^{d}\sum_{j=1}^{d}\left[\Theta_{p_ip_j}(t,x,Dw^1,\mu^1)\frac{\partial^2 w^1}{\partial x_i\partial x_j} - \Theta_{p_ip_j}(t,x,Dw^2,\mu^2)\frac{\partial^2 w^1}{\partial x_i\partial x_j}\right]dx,$$

$$V_{14} = -\bar{m}\int_{\mathbb{T}^d}(\mu^1-\mu^2)\sum_{i=1}^{d}\sum_{j=1}^{d}\left[\Theta_{p_ip_j}(t,x,Dw^2,\mu^2)\left(\frac{\partial^2(w^1-w^2)}{\partial x_i\partial x_j}\right)\right]\,dx.$$

We integrate four of these terms by parts. For V_1 we arrive at

$$V_1 = -\int_{\mathbb{T}^d}\left|\nabla\left(\mu^1-\mu^2\right)\right|^2\,dx. \tag{A.3}$$

Continuing to integrate by parts, we have

$$V_2 = \frac{1}{2}\int_{\mathbb{T}^d}(\mu^1-\mu^2)^2\mathrm{div}\left(D_p\Theta(t,x,Dw^1,\mu^1)\right)\,dx,$$

$$V_7 = \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{T}^d}(\mu^1-\mu^2)^2\frac{\partial}{\partial x_i}\left((\mu^2)\Theta_{p_iq}(t,x,Dw^1,\mu^1)\right)\,dx,$$

$$V_{12} = \frac{\bar{m}}{2}\sum_{i=1}^{d}\int_{\mathbb{T}^d}(\mu^1-\mu^2)^2\frac{\partial}{\partial x_i}\left(\Theta_{p_iq}(t,x,Dw^1,\mu^1)\right)\,dx.$$

These terms, and also V_3, are then bounded in terms of the energy, using the bound on the solutions in terms of K. We introduce a non-decreasing function $\mathcal{G}$ such that

$$V_2 + V_3 + V_7 + V_{12} \le \mathcal{G}(K)E_\mu. \tag{A.4}$$

Note that the we have estimated, for instance, μ^2 in L^∞ here, which we may do since s is large enough. Such estimates will be made several times in the rest of this proof.

For most of the remaining V_i terms, we may estimate them using Lipschitz properties of $D_p\Theta$ and its derivatives; we then have

$$V_4 + V_5 + V_6 + V_8 + V_{10} + V_{11} + V_{13} \le \mathcal{G}(K)(E_\mu + E_\mu^{1/2}E_w^{1/2}), \tag{A.5}$$

where $\mathcal{G}(K)$ is as before.

For the final two terms, V_9 and V_{14}, we use Young's inequality with an eye toward bounding them by a beneficial term arising from E_w. For V_9, we bound $\Theta_{p_ip_j}$ and μ^2 with $\mathcal{G}(K)$:

$$V_9 \le \mathcal{G}(K)\sum_{i=1}^{d}\sum_{j=1}^{d}\int_{\mathbb{T}^d}(\mu^1-\mu^2)\frac{\partial^2(w^1-w^2)}{\partial_{x_i}\partial_{x_j}}\,dx.$$

We introduce a constant $\varpi = \frac{1}{8\Upsilon}$. We then apply Young's inequality, with parameter $4\mathcal{G}(K)/\varpi$:

$$V_9 \leq \mathcal{G}(K) \int_{\mathbb{T}^d} (\mu^1 - \mu^2)^2 \, dx + \frac{1}{8} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx \tag{A.6}$$

$$\leq \mathcal{G}(K) E_\mu + \frac{\varpi}{8} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx.$$

Notice that in the first term on the right-hand side, we have simply incorporated ϖ into $\mathcal{G}$. The final term, V_{14}, is then entirely similar:

$$V_{14} \leq \mathcal{G}(K) E_\mu + \frac{\varpi}{8} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx. \tag{A.7}$$

We add the bounds (A.4), (A.5), (A.6), and (A.7), finding the following conclusion:

$$\sum_{\ell=2}^{14} V_\ell \leq \mathcal{G}(K)(E_w + E_\mu) + \frac{\varpi}{4} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx. \tag{A.8}$$

We next treat E_w, introducing the notation $E_w = \sum_{j=1}^{d} E_w^j$, with

$$\frac{dE_w^j}{dt} = \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \partial_t (\partial_{x_j} w^1 - \partial_{x_j} w^2) \, dx.$$

We further decompose each of these time derivatives as

$$\frac{dE_w^j}{dt} = \sum_{\ell=1}^{6} W_\ell^j,$$

where the W_i^j are given by

$$W_1^j = -\int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \Delta (\partial_{x_j} w^1 - \partial_{x_j} w^2) \, dx,$$

$$W_2^j = -\int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w_2) \left(\Theta_{x_j}(t, x, Dw^1, \mu^1) - \Theta_{x_j}(t, x, Dw^2, \mu^2) \right) \, dx,$$

$$W_3^j = -\int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \left(\Theta_q(t, x, Dw^1, \mu^1) \mu^1_{x_j} - \Theta_q(t, x, Dw^2, \mu^2) \mu^1_{x_j} \right) \, dx,$$

$$W_4^j = -\int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \left(D_p \Theta(t, x, Dw^2, \mu^2) \mu^1_{x_j} - \Theta_q(t, x, Dw^2, \mu^2) \mu^2_{x_j} \right) \, dx,$$

$$W_5^j = -\sum_{i=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \left(\Theta_{p_i}(t, x, Dw^1, \mu^1) \partial^2_{x_i x_j} w^1 \right.$$
$$\left. - \Theta_{p_i}(t, x, Dw^2, \mu^2) \partial^2_{x_i x_j} w^1 \right) \, dx,$$

$$W_6^j = -\sum_{i=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \left(\Theta_{p_i}(t, x, Dw^2, \mu^2) \partial^2_{x_i x_j} w^1 \right.$$
$$\left. - \Theta_{p_i}(t, x, Dw^2, \mu^2) \partial^2_{x_i x_j} w^2 \right) \, dx.$$

As before, we begin by integrating some of the terms by parts. Integrating W_1^j by parts, we have

$$W_1^j = \int_{\mathbb{T}^d} \left|\nabla \partial_{x_j} \left(w^1 - w^2\right)\right|^2 \, dx.$$

We also have the following, upon integrating W_6^j by parts:

$$W_6^j = \frac{1}{2}\sum_{i=1}^d \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2)^2 \partial_{x_i} \left(\Theta_{p_i}(t, x, Dw^2, \mu^2)\right) \, dx.$$

With $\mathcal{G}$ as before, we may then estimate several of the terms:

$$W_2^j + W_3^j + W_5^j + W_6^j \leq \mathcal{G}(K)(E_w + E_w^{1/2} E_\mu^{1/2}). \tag{A.9}$$

The only remaining term to estimate is W_4^j. We first bound $D_p\Theta(t, x, Dw^2, \mu^2)$ in L^∞ using $\mathcal{G}(K)$, arriving at

$$W_4^j \leq \mathcal{G}(K) \int_{\mathbb{T}^d} (\partial_{x_j} w^1 - \partial_{x_j} w^2) \partial_{x_j} (\mu^1 - \mu^2) \, dx.$$

Applying Young's inequality with $2\mathcal{G}(K)$ as the parameter, we have

$$W_4^j \leq \mathcal{G}(K) E_w + \frac{1}{4} \int_{\mathbb{T}^d} (\partial_{x_j} \mu^1 - \partial_{x_j} \mu^2)^2 \, dx.$$

Adding (A.9) to this yields

$$\sum_{\ell=2}^6 \sum_{j=1}^d W_\ell^j \leq \mathcal{G}(K)(E_w + E_\mu) + \frac{1}{4} \int_{\mathbb{T}^d} |D\mu^1 - D\mu^2|^2 \, dx. \tag{A.10}$$

We are now in a position to integrate $\frac{dE_\mu}{dt}$ with respect to time over the interval $[0, t]$; doing so, we arrive at

$$E_\mu(t) = E_\mu(0) + \int_0^t \sum_{\ell=1}^{14} V_\ell.$$

Applying (A.8) then yields

$$E_\mu(t) \leq E_\mu(0) + T\mathcal{G}(K)(E_\mu + E_w) + \int_0^t \left[V_1 + \frac{\varpi}{4} \sum_{j=1}^d \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx\right] d\tau.$$

We similarly integrate $\frac{dE_w}{dt}$ over the temporal interval $[t, T]$, finding

$$E_w(t) = E_w(T) - \int_t^T \sum_{j=1}^d \sum_{\ell=1}^6 W_\ell^j \, d\tau.$$

Applying (A.10), we find

$$E_w(t) \leq E_w(T) + T\mathcal{G}(K)(E_w + E_\mu) - \sum_{j=1}^d \int_t^T W_1^j \, d\tau + \frac{1}{4} \int_t^T \int_{\mathbb{T}^d} |D\mu^1 - D\mu^2|^2 \, dx d\tau. \tag{A.11}$$

We work now with the term $E_w(T)$. A formula for this is

$$E_w(T) = \sum_j \|\partial_{x_j} w_1(T, \cdot) - \partial_{x_j} w_2(T, \cdot)\|_0^2.$$

Using the payoff function, $\tilde{G}$, this can be estimated as

$$E_w(T) \leq \|\tilde{G}(\mu_1(T,\cdot),\cdot) - \tilde{G}(\mu_2(T,\cdot),\cdot)\|_1^2.$$

Using our smoothing assumption on $\tilde{G}$, then, this may again be estimated as

$$E_w(T) \leq \Upsilon \|\mu_1(T,\cdot) - \mu_2(T,\cdot)\|_0^2.$$

We may replace this with a supremum, as

$$E_w(T) \leq \Upsilon \sup_{t\in[0,T]} E_\mu(t).$$

We multiply (A.11) by $\varpi > 0$, apply the definitions of V_1 and W_1^j, and summarize what we have found thus far:

$$\begin{aligned}(\text{A.12})\quad &\varpi E_w(t) + E_\mu(t) + \int_0^t \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \varpi \int_t^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau \\ &\leq \varpi\Upsilon \sup_{t\in[0,T]} E_\mu(t) + E_\mu(0) + T\mathcal{G}(K)(E_w(t) + E_\mu(t)) \\ &\quad + \frac{\varpi}{4}\int_t^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \frac{\varpi}{4}\int_0^t \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau.\end{aligned}$$

We can bound the integrals appearing on the right-hand side by taking a larger domain of integration:

$$\begin{aligned}(\text{A.13})\quad &\varpi E_w(t) + E_\mu(t) + \int_0^t \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \varpi \int_t^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau \\ &\leq \varpi\Upsilon \sup_{t\in[0,T]} E_\mu(t) + E_\mu(0) + T\mathcal{G}(K)(E_w(t) + E_\mu(t)) \\ &\quad + \frac{\varpi}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \frac{\varpi}{4}\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau.\end{aligned}$$

Considering the first integral on the left-hand side of (A.13), we find

$$\begin{aligned}(\text{A.14})\quad &\int_0^t \|D\mu^1 - D\mu^2\|_0^2 \, d\tau \leq \varpi\Upsilon \sup_{t\in[0,T]} E_\mu(t) + E_\mu(0) + T\mathcal{G}(K)(E_w(t) + E_\mu(t)) \\ &\quad + \frac{\varpi}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \frac{\varpi}{4}\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau.\end{aligned}$$

Taking the supremum with respect to t on both sides of (A.14), we have

$$\begin{aligned}(\text{A.15})\quad &\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau \leq \varpi\Upsilon \sup_{t\in[0,T]} E_\mu(t) + E_\mu(0) + T\mathcal{G}(K)\left(\sup_{t\in[0,T]}(E_w(t) + E_\mu(t))\right) \\ &\quad + \frac{\varpi}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \frac{\varpi}{4}\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau.\end{aligned}$$

Now considering instead the second integral on the left-hand side of (A.13), making the corresponding manipulations, we have

$$\begin{aligned}\varpi &\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau \\ &\leq \varpi \Upsilon \sup_{t\in[0,T]} E_\mu(t) + E_\mu(0) + T\mathcal{G}(K)\left(\sup_{t\in[0,T]} (E_w(t) + E_\mu(t))\right) \\ &\quad + \frac{\varpi}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \frac{\varpi}{4}\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau.\end{aligned} \tag{A.16}$$

We assume that ϖ satisfies $\varpi < 1$. Adding (A.15) and (A.16) and rearranging, we have

$$\begin{aligned}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \, d\tau + \varpi &\int_0^T \|D^2 w^1 - D^2 w^2\|_0^2 \, d\tau \\ &\leq 4E_\mu(0) + 4\varpi\Upsilon E_\mu(0) + T\mathcal{G}(K)\left(\sup_{t\in[0,T]} (E_w(t) + E_\mu(t))\right).\end{aligned}$$

Combining this with (A.13), and making some adjustments of factors of ϖ, we have

$$\begin{aligned}\sup_{t\in[0,T]} (\varpi E_w(t) + E_\mu(t)) &\leq 2E_\mu(0) + 2\varpi\Upsilon \sup_{t\in[0,T]} (\varpi E_w(t) + E_\mu(t)) \\ &\quad + T\mathcal{G}(K)\left(\sup_{t\in[0,T]} (\varpi E_w(t) + E_\mu(t))\right).\end{aligned} \tag{A.17}$$

Recall our choice $\varpi = \frac{1}{8\Upsilon}$. We take $T\mathcal{G}(K) < \frac{1}{4}$ (which may be accomplished simply by taking T sufficiently small), concluding

$$\sup_{t\in[0,T]} (E_\mu(t) + \varpi E_w(t)) \leq 8E_\mu(0) = 8\|m^1(0) - m^2(0)\|_0^2.$$

We may eliminate the factor of ϖ on the left, resulting in

$$\sup_{t\in[0,T]} (E_\mu(t) + E_w(t)) \leq 64\Upsilon\|m^1(0) - m^2(0)\|_0^2. \tag{A.18}$$

Let us notice that the estimate on the μ variables directly translates to the estimate on the m variables (since $m^1 - m^2 = \mu^1 = \mu^2$). Therefore, to conclude with the estimate on the u variables from the estimate on the w variables, it remains to estimate the difference of the averages of the u variables.

The equation for u is

$$\partial_t u = -\Delta u + \mathcal{H}(t, x, \nabla u, m).$$

Integrating this over the entire spatial domain, we have

$$\partial_t \left(\int_{\mathbb{T}^d} u \, dx\right) = \int_{\mathbb{T}^d} \mathcal{H}(t, x, \nabla u, m) \, dx.$$

Let (u_1, m_1) and (u_2, m_2) be two solutions. We take the difference in the evolution equations for the mean, recalling that $\nabla u_i = \nabla w_i$:

$$\partial_t \int_{\mathbb{T}^d} u_1 - u_2 \, dx = \int_{\mathbb{T}^d} \mathcal{H}(t, x, \nabla w_1, m_1) - \mathcal{H}(t, x, \nabla w_2, m_2) \, dx.$$

Taking absolute values, adding and subtracting, and using the triangle inequality, we have

$$\left|\partial_t \int_{\mathbb{T}_d} u_1 - u_2 \, dx\right| \leq \int_{\mathbb{T}^d} |\mathcal{H}(t,x,\nabla w_1, m_1) - \mathcal{H}(t,x,\nabla w_2, m_1)| \, dx + \int_{\mathbb{T}^d} |\mathcal{H}(t,x,\nabla w_2, m_1) - \mathcal{H}(t,x,\nabla w_2, m_2)| \, dx.$$

Using the Lipschitz properties of the Hamiltonian, and the boundedness of the solutions (w_i, m_i), we may bound this as

$$\left|\partial_t \int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq c \int_{\mathbb{T}^d} |\nabla w_1 - \nabla w_2| \, dx + c \int_{\mathbb{T}^d} |m_1 - m_2| \, dx.$$

Using $m_1 - m_2 = \mu_1 - \mu_2$, this then implies

$$\left|\partial_t \int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq c \left(\|w_1 - w_2\|_{H^1} + \|\mu_1 - \mu_2\|_{L^2} \right). \tag{A.19}$$

We can write the difference of the means as

$$\int_{\mathbb{T}^d} u_1(t,\cdot) - u_2(t,\cdot) \, dx = \int_{\mathbb{T}^d} u_1(T,\cdot) - u_2(T,\cdot) \, dx - \int_t^T \partial_t \left(\int_{\mathbb{T}^d} u_1 - u_2 \, dx \right) ds.$$

Using the terminal condition for u, this becomes

$$\int_{\mathbb{T}^d} u_1(t,\cdot) - u_2(t,\cdot) \, dx = \int_{\mathbb{T}^d} G(m_1(T,\cdot),\cdot) - G(m_2(T,\cdot),\cdot) \, dx - \int_t^T \partial_t \left(\int_{\mathbb{T}^d} u_1 - u_2 \, dx \right) ds.$$

We may then estimate this as

$$\left|\int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq \int_{\mathbb{T}^d} |G(m_1(T,\cdot),\cdot) - G(m_2(T,\cdot),\cdot)| \, dx + T \sup_{t\in[0,T]} \left|\partial_t \int_{\mathbb{T}^d} u_1 - u_2 \, dx\right|.$$

The Lipschitz property of G then implies

$$\left|\int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq c \int_{\mathbb{T}^d} |m_1(T,\cdot) - m_2(T,\cdot)| \, dx + T \sup_{t\in[0,T]} \left|\partial_t \int_{\mathbb{T}^d} u_1 - u_2 \, dx\right|.$$

Again using $m_1 - m_2 = \mu_1 - \mu_2$, and also using (A.19), and using a supremum in time, we have

$$\sup_{t\in[0,T]} \left|\int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq c \sup_{t\in[0,T]} (\|w_1 - w_2\|_{H^1} + \|\mu_1 - \mu_2\|_{L^2}).$$

We see that (A.18) then implies

$$\sup_{t\in[0,T]} \left|\int_{\mathbb{T}^d} u_1 - u_2 \, dx\right| \leq c \|m_0^1 - m_0^2\|.$$

This completes the proof of (A.1). □

Acknowledgments

The authors wish to thank W. Gangbo his constant interest in this work and for his valuable feedback and comments that he gave at various stages during the preparation of the manuscript. This project was conceived at IPAM, UCLA and a part of it was done while both authors were members of the long program "High Dimensional Hamilton-Jacobi PDEs" in 2020. D.M.A. has been partially supported by the NSF grant DMS-1907684, A.R.M. has been partially supported by the Air Force grant FA9550-18-1-0502 and by the King Abdullah University of Science and Technology Research Funding (KRF) under Award No. ORA-2021-CRG10-4674.2.

References

[1] Yves Achdou, Francisco J. Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, *Partial differential equation models in macroeconomics*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **372** (2014), no. 2028, 20130397, 19, DOI 10.1098/rsta.2013.0397. MR3268061

[2] Yves Achdou and Alessio Porretta, *Mean field games with congestion*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **35** (2018), no. 2, 443–480, DOI 10.1016/j.anihpc.2017.06.001. MR3765549

[3] David M. Ambrose, *Well-posedness of vortex sheets with surface tension*, SIAM J. Math. Anal. **35** (2003), no. 1, 211–244, DOI 10.1137/S0036141002403869. MR2001473

[4] David M. Ambrose, *Existence theory for non-separable mean field games in Sobolev spaces*, Indiana Univ. Math. J. **71** (2022), no. 2, 611–647. MR4420100

[5] David M. Ambrose, *Strong solutions for time-dependent mean field games with non-separable Hamiltonians*, J. Math. Pures Appl. (9) **113** (2018), 141–154, DOI 10.1016/j.matpur.2018.03.003. MR3784807

[6] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR2401600

[7] J. Thomas Beale, *The initial value problem for the Navier-Stokes equations with a free surface*, Comm. Pure Appl. Math. **34** (1981), no. 3, 359–392, DOI 10.1002/cpa.3160340305. MR611750

[8] A. Bensoussan, P. J. Graber, and S. C. P. Yam, *Control on Hilbert spaces and application to mean field type control theory*, arXiv:2005.10770, 2020.

[9] P. Cardaliaguet, *The convergence problem in mean field games with local coupling*, Appl. Math. Optim. **76** (2017), no. 1, 177–215, DOI 10.1007/s00245-017-9434-0. MR3679342

[10] P. Cardaliaguet, M. Cirant, A. Porretta, Splitting methods and short time existence for the master equations in mean field games, *J. Eur. Math. Soc. (JEMS)*, to appear, arXiv:2001.11970, 2020. DOI 10.4171/JEMS/1227.

[11] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions, *The master equation and the convergence problem in mean field games*, Annals of Mathematics Studies, vol. 201, Princeton University Press, Princeton, NJ, 2019, DOI 10.2307/j.ctvckq7qf. MR3967062

[12] René Carmona and François Delarue, *Probabilistic theory of mean field games with applications. II*, Probability Theory and Stochastic Modelling, vol. 84, Springer, Cham, 2018. Mean field games with common noise and master equations. MR3753660

[13] J.-F. Chassagneux, D. Crisan, and F. Delarue, *A probabilistic approach to classical solutions of the master equation for large population equilibria*, Mem. Amer. Math. Soc., *to appear*.

[14] Marco Cirant, Roberto Gianni, and Paola Mannucci, *Short-time existence for a general backward-forward parabolic system arising from mean-field games*, Dyn. Games Appl. **10** (2020), no. 1, 100–119, DOI 10.1007/s13235-019-00311-5. MR4064664

[15] François Delarue, Daniel Lacker, and Kavita Ramanan, *From the master equation to mean field game limit theory: a central limit theorem*, Electron. J. Probab. **24** (2019), Paper No. 51, 54, DOI 10.1214/19-EJP298. MR3954791

[16] François Delarue, Daniel Lacker, and Kavita Ramanan, *From the master equation to mean field game limit theory: large deviations and concentration of measure*, Ann. Probab. **48** (2020), no. 1, 211–263, DOI 10.1214/19-AOP1359. MR4079435

[17] David Evangelista, Rita Ferreira, Diogo A. Gomes, Levon Nurbekyan, and Vardan Voskanyan, *First-order, stationary mean-field games with congestion*, Nonlinear Anal. **173** (2018), 37–74, DOI 10.1016/j.na.2018.03.011. MR3802565

[18] David Evangelista and Diogo A. Gomes, *On the existence of solutions for stationary mean-field games with congestion*, J. Dynam. Differential Equations **30** (2018), no. 4, 1365–1388, DOI 10.1007/s10884-017-9615-1. MR3871606

[19] W. Gangbo and A. R. Mészáros, *Global well-posedness of Master equations for deterministic displacement convex potential mean field games*, Comm. Pure Appl. Math., to appear. arXiv:2004.01660 (2020). DOI 10.1002/cpa.22069.

[20] W. Gangbo, A. R. Mészáros, C. Mou, and J. Zhang, *Mean field games master equations with non-separable Hamiltonians and displacement monotonicity*, Ann. Probab. **50** (2022), no. 6, 2178–2217.

[21] Wilfrid Gangbo and Adrian Tudorascu, *On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi equations* (English, with English and French summaries), J. Math. Pures Appl. (9) **125** (2019), 119–174, DOI 10.1016/j.matpur.2018.09.003. MR3944201

[22] Wilfrid Gangbo and Adrian Tudorascu, *Weak KAM theory on the Wasserstein torus with multidimensional underlying space*, Comm. Pure Appl. Math. **67** (2014), no. 3, 408–463, DOI 10.1002/cpa.21492. MR3158572

[23] Wilfrid Gangbo and Andrzej Święch, *Existence of a solution to an equation arising from the theory of mean field games*, J. Differential Equations **259** (2015), no. 11, 6573–6643, DOI 10.1016/j.jde.2015.08.001. MR3397332

[24] Diogo Gomes and Marc Sedjro, *One-dimensional, forward-forward mean-field games with congestion*, Discrete Contin. Dyn. Syst. Ser. S **11** (2018), no. 5, 901–914, DOI 10.3934/dcdss.2018054. MR3817560

[25] Diogo A. Gomes and Vardan K. Voskanyan, *Short-time existence of solutions for mean-field games with congestion*, J. Lond. Math. Soc. (2) **92** (2015), no. 3, 778–799, DOI 10.1112/jlms/jdv052. MR3431662

[26] Minyi Huang, Roland P. Malhamé, and Peter E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst. **6** (2006), no. 3, 221–251. MR2346927

[27] Jean-Michel Lasry and Pierre-Louis Lions, *Jeux à champ moyen. I. Le cas stationnaire* (French, with English and French summaries), C. R. Math. Acad. Sci. Paris **343** (2006), no. 9, 619–625, DOI 10.1016/j.crma.2006.09.019. MR2269875

[28] Jean-Michel Lasry and Pierre-Louis Lions, *Jeux à champ moyen. II. Horizon fini et contrôle optimal* (French, with English and French summaries), C. R. Math. Acad. Sci. Paris **343** (2006), no. 10, 679–684, DOI 10.1016/j.crma.2006.09.018. MR2271747

[29] Jean-Michel Lasry and Pierre-Louis Lions, *Mean field games*, Jpn. J. Math. **2** (2007), no. 1, 229–260, DOI 10.1007/s11537-007-0657-8. MR2295621

[30] P.-L. Lions, *Lectures at Collège de France*.

[31] Sergio Mayorga, *Short time solution to the master equation of a first order mean field game*, J. Differential Equations **268** (2020), no. 10, 6251–6318, DOI 10.1016/j.jde.2019.11.031. MR4069006

[32] Stéphane Mischler and Clément Mouhot, *Kac's program in kinetic theory*, Invent. Math. **193** (2013), no. 1, 1–147, DOI 10.1007/s00222-012-0422-3. MR3069113

[33] C. Mou and J. Zhang, *Wellposedness of second order master equations for mean field games with nonsmooth data*, Mem. Amer. Math. Soc., to appear. arXiv:1903.09907v3 (2019).

[34] Rémi Peyre, *Comparison between* W_2 *distance and* $\dot{H}^{-1}$ *norm, and localization of Wasserstein distance*, ESAIM Control Optim. Calc. Var. **24** (2018), no. 4, 1489–1501, DOI 10.1051/cocv/2017050. MR3922440

Department of Mathematics, Drexel University, Philadelphia, Pennsylvania
Email address: dma68@drexel.edu

Department of Mathematical Sciences, University of Durham, Durham DH1 3LE, England
Email address: alpar.r.meszaros@durham.ac.uk